Principles of Digital PCR and Measurement Issues:

The certification of Cytomegalovirus Standard Reference Material (SRM 2366) as a model for future SRMs

by Ross Haynes

National Institute of Standards and Technology

October 15, 2012
Digital PCR Applications and Advances
San Diego, CA

Disclaimer

- Views expressed in this presentation are the author's opinion and do not represent the opinion of NIST or the Department of Commerce.
- Any mention of commercial products within this presentation is for information only; it does not imply recommendation or endorsement by NIST, nor does it imply that the products are the best available.

Agenda

- NIST Overview
- Quantitative PCR versus Digital PCR
- Digital PCR Applications
- Poisson Statistics
- Standard Reference Material 2366
  - Cytomegalovirus for DNA Measurements
- Technology Types
- Measurement Issues

NIST Overview

National Institute of Standards and Technology

- Located ~25 miles north west of Washington DC
- Non-regulatory agency
- Part of the US Department of Commerce
- National Metrology Institute for the US
- Mission: To promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life.

NIST Overview

- 1901 National Bureau of Standards (NBS) established
- 1988 name changed to National Institute of Standards and Technology (NIST)
- Focus on standards of current importance
  - Past: railroads, fire hydrants, etc.
  - Current: healthcare, IT security, etc.

NIST Overview

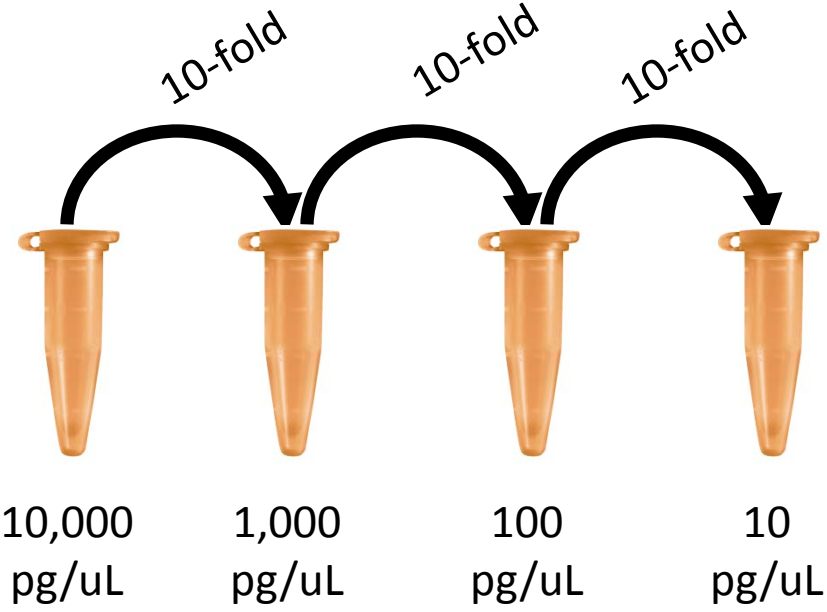
- NIST produces Standard Reference Materials
- Intended to be used to calibrate instrumentation or everyday use calibrants
- Applied Genetics Group
  - Cytomegalovirus (SRM 2366)
  - Huntington's Disease (SRM 2393)
  - Human DNA Quantitation Standard (SRM 2372)
  - Mitochondrial DNA Sequencing (SRM 2392 & SRM 2392-I)
  - Forensic DNA Typing using STRs (SRM 2391c & SRM 2395)

Agenda

- Overview of NIST
- Quantitative PCR versus Digital PCR
- Digital PCR Applications
- Poisson Statistics
- Standard Reference Material 2366
  - Cytomegalovirus for DNA Measurements
- Technology Types
- Measurement Issues

qPCR

- Calibrant concentration is independently determined (e.g. UV spectrophotometer)
- Prepare a dilution curve of calibrant



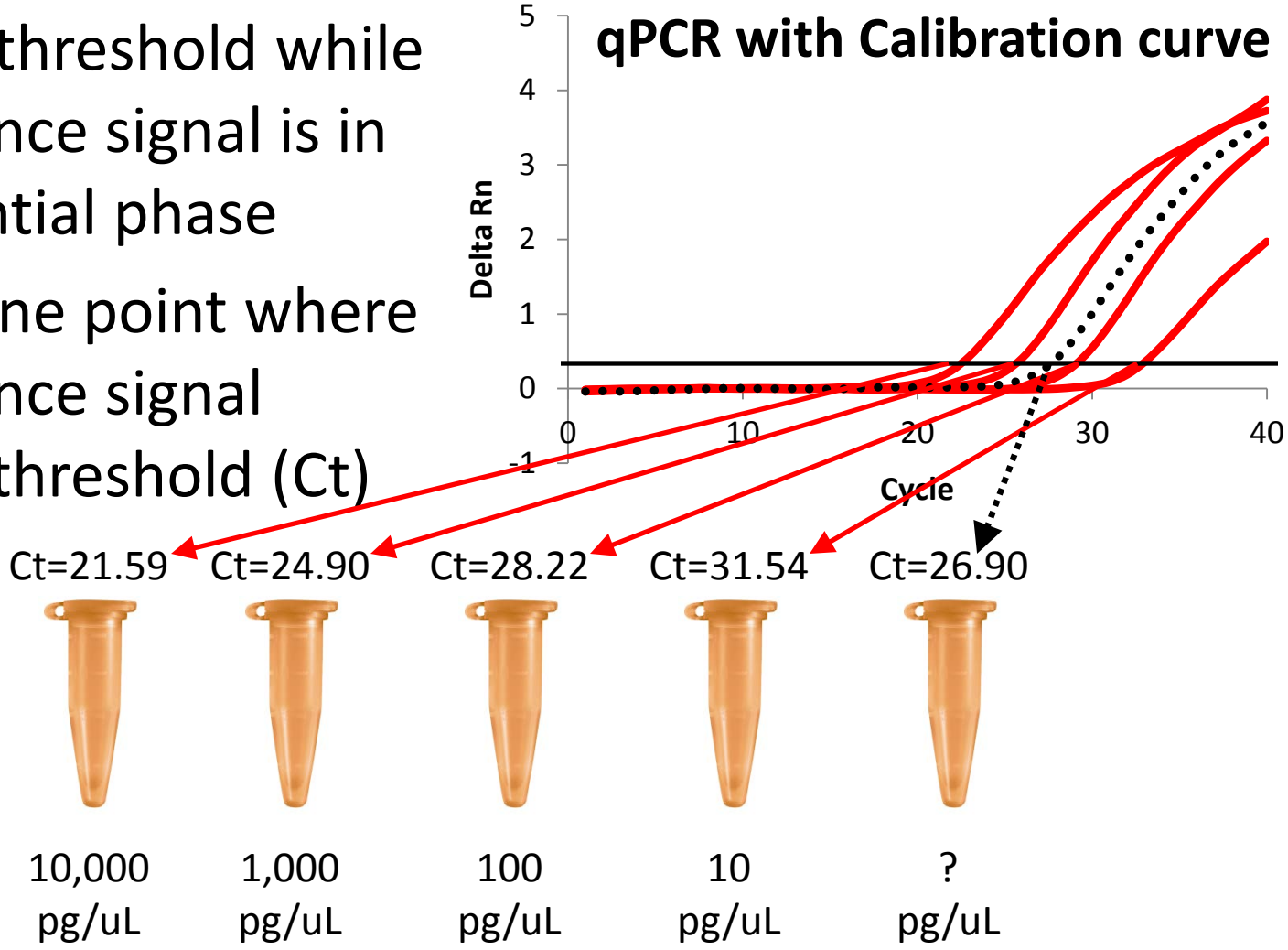
qPCR

- Use calibration dilutions plus samples of unknown concentration as template for qPCR
- Thermal cycle and measure florescence signal after each cycle of PCR

10,000 pg/uL | 1,000 pg/uL | 100 pg/uL | 10 pg/uL | ? pg/uL

qPCR

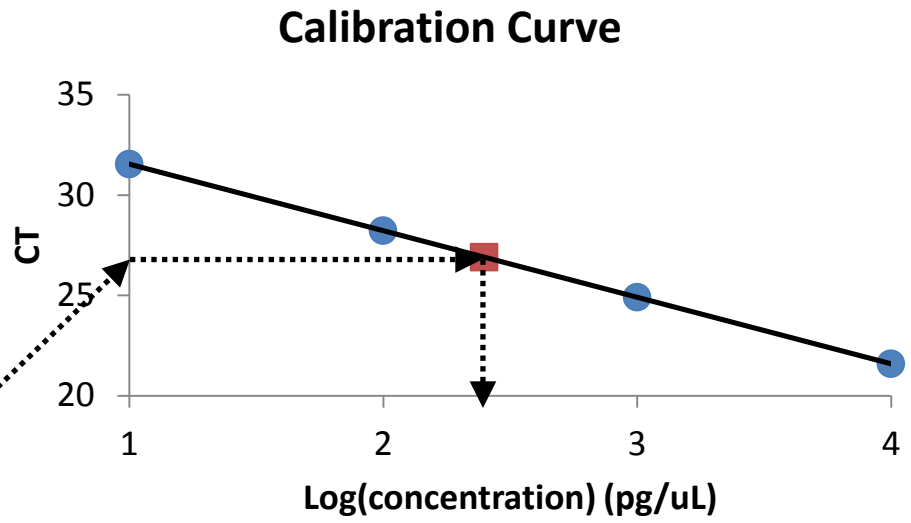
- Apply a threshold while florescence signal is in exponential phase
- Determine point where florescence signal crosses threshold (Ct)



qPCR

- Log transform concentration
- Plot Log(conc.) vs Ct

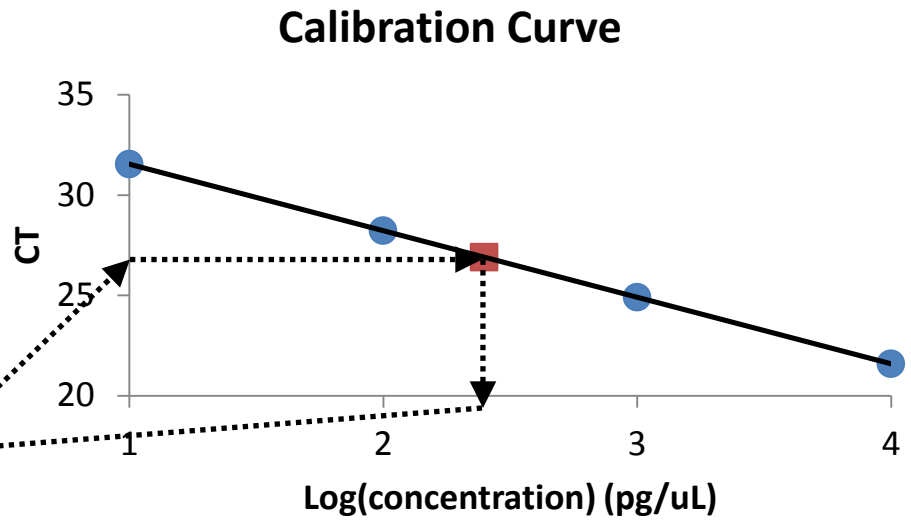
Conc. (pg/µL)	Log(conc.)	Ct
10,000	4	21.59
1,000	3	24.90
100	2	28.22
10	1	31.54
unknown	unknown	26.90



qPCR

- Log transform concentration
- Plot Log(conc.) vs Ct

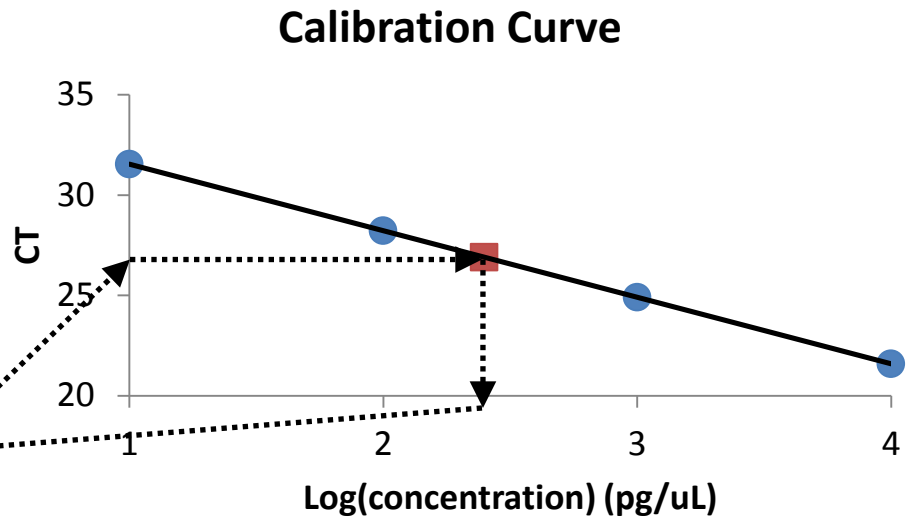
Conc. (pg/µL)	Log(conc.)	Ct
10,000	4	21.59
1,000	3	24.90
100	2	28.22
10	1	31.54
unknown	2.40	26.90



qPCR

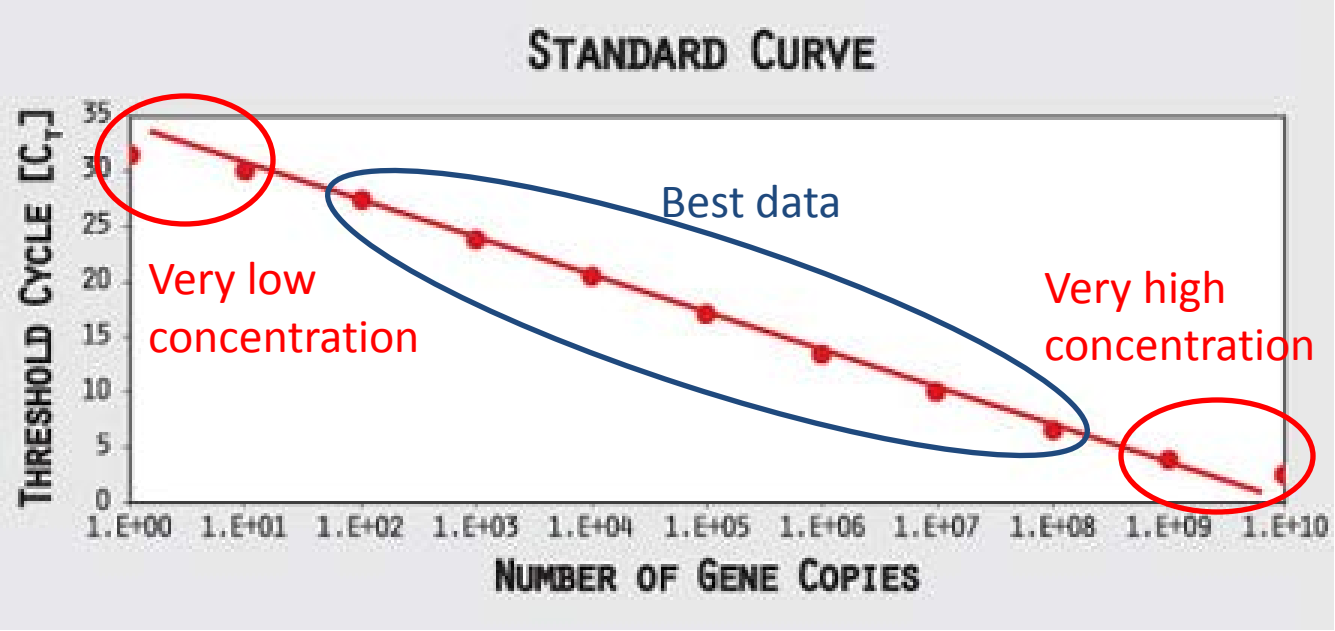
- Log transform concentration
- Plot Log(conc.) vs Ct

Conc. (pg/µL)	Log(conc.)	Ct
10,000	4	21.59
1,000	3	24.90
100	2	28.22
10	1	31.54
251	2.40	26.90



qPCR "Goldilocks Zone"

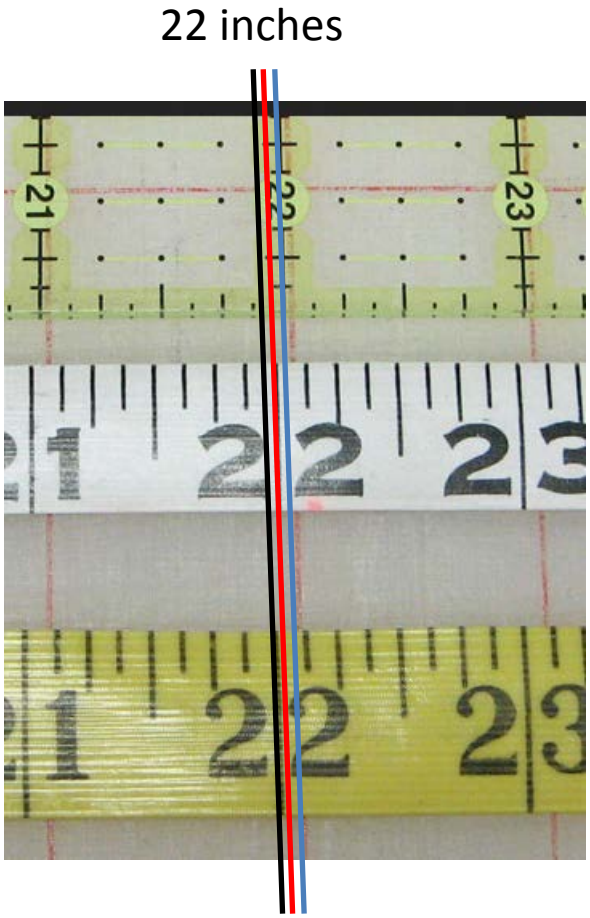
- Very high and very low concentrations do not fit on the line
- Best data obtained from the middle



http://www.sabiosciences.com/pathwaymagazine/pathways7/designing-validating-real-time-pcr-primers.php

qPCR

- Relative quantitation between calibrant of known concentration (aka standard) and samples of unknown concentration
  - Just as using a tape measure is a relative measurement if the calibrant is inaccurate the measurement will be inaccurate
- Spectrophotometer measures everything that absorbs at 260 nm (i.e. DNA, RNA, protein, monomers)

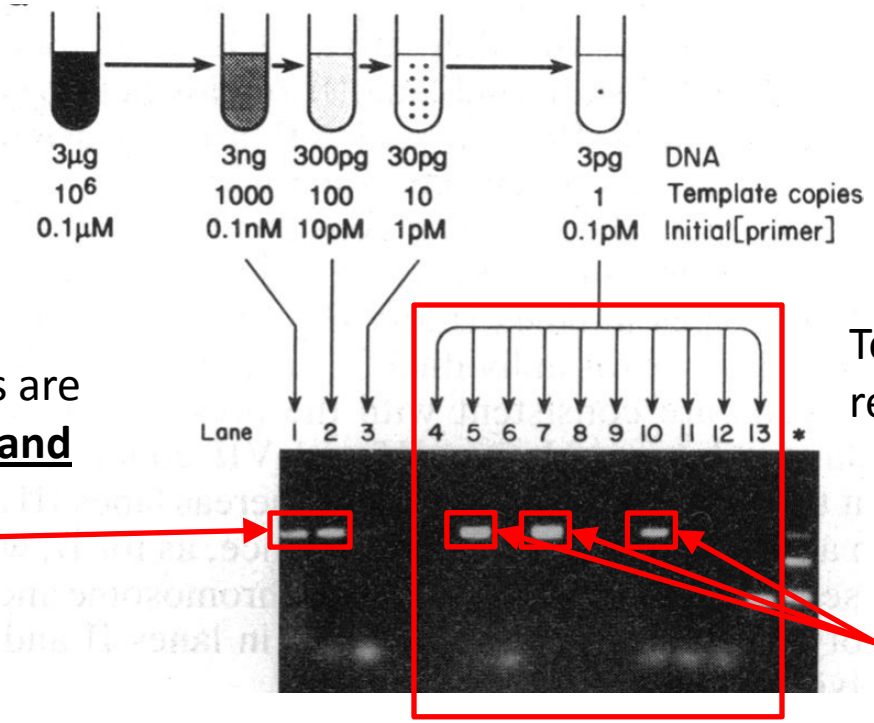


http://vickiwelsh.typepad.com/field_trips_in_fiber/tips/

History of dPCR

- 1990 – Single Molecule Dilution (SMD) PCR
  - Resolve maternal/paternal sequence
  - Poisson statistics to estimate dilution needed

Figure 3a



These PCR fragments are mixture of maternal **and** paternal sequences

Ten replicate PCR reactions

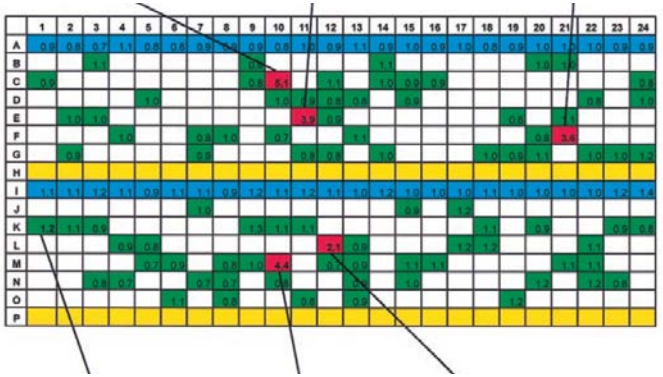
These PCR fragments are likely either maternal **or** paternal sequence.

Ruano et al. Proc Natl Acad Sci U S A. 1990 Aug;87(16):6296-300.

History of dPCR

- 1999 – “Digital PCR”
  - Detect rare mutations (cancer/wt) using SMD PCR
  - Use Poisson statistics to quantitate DNA

Mutant allele (KRAS)
Wild type allele
Positive control
Negative control



- 97 wild type
- 4 KRAS mutants by sequencing
- 1 silent mutation
- 4 % detection KRAS mutants
- 1 % detection of silent mutants

Vogelstein et al. Proc Natl Acad Sci U S A. 1999 Aug 3;96(16):9236-41. Figure 5

History of dPCR

- 1999 – “Digital PCR”
  - Detect rare mutations (cancer/wt) using SMD PCR
  - Use Poisson statistics to quantitate DNA
  - Proposed uses for Digital PCR
    - Investigation of individual alleles (1990)
    - Rare allele detection
    - Relative expression levels
    - Quantitative analysis of PCR products

Vogelstein et al. Proc Natl Acad Sci U S A. 1999 Aug 3;96(16):9236-41.

dPCR

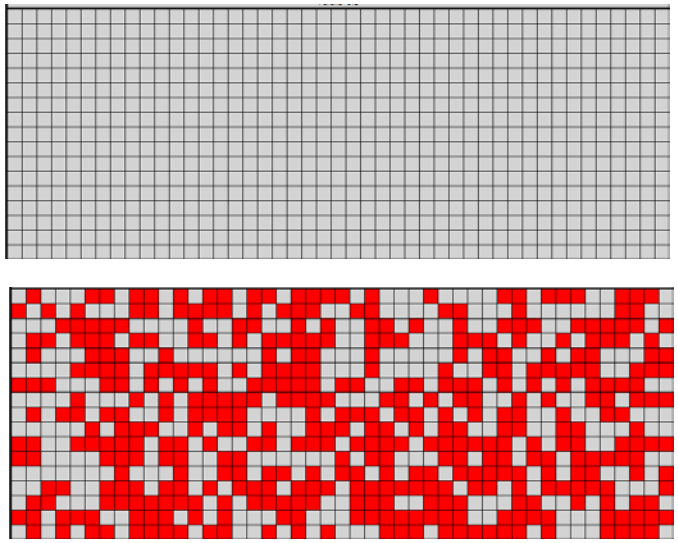
1) Create a PCR mastermix as if for qPCR

2) Aliquot across 100s or 1000s of wells

3) Thermal cycle as if for qPCR & count wells with detectible amplification at any cycle

4) Use Poisson statistics to determine concentration of starting material

?
pg/uL



dPCR

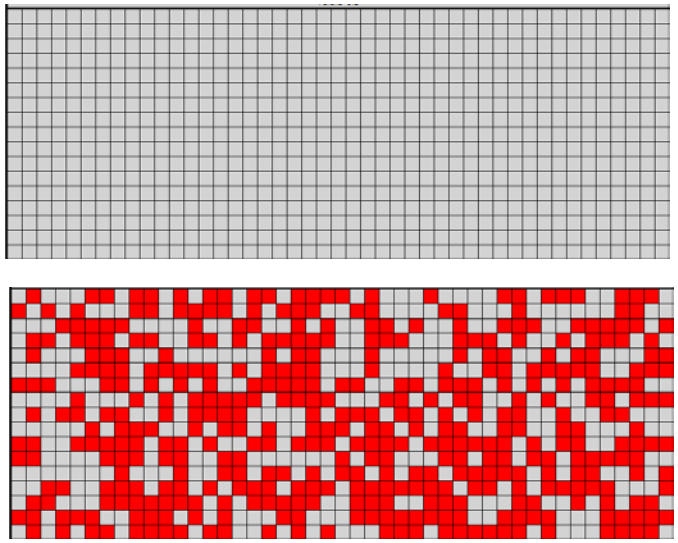
1) Create a PCR mastermix as if for qPCR

2) Aliquot across 100s or 1000s of wells

3) Thermal cycle as if for qPCR & count wells with detectible amplification at any cycle

4) Use Poisson statistics to determine concentration of starting material

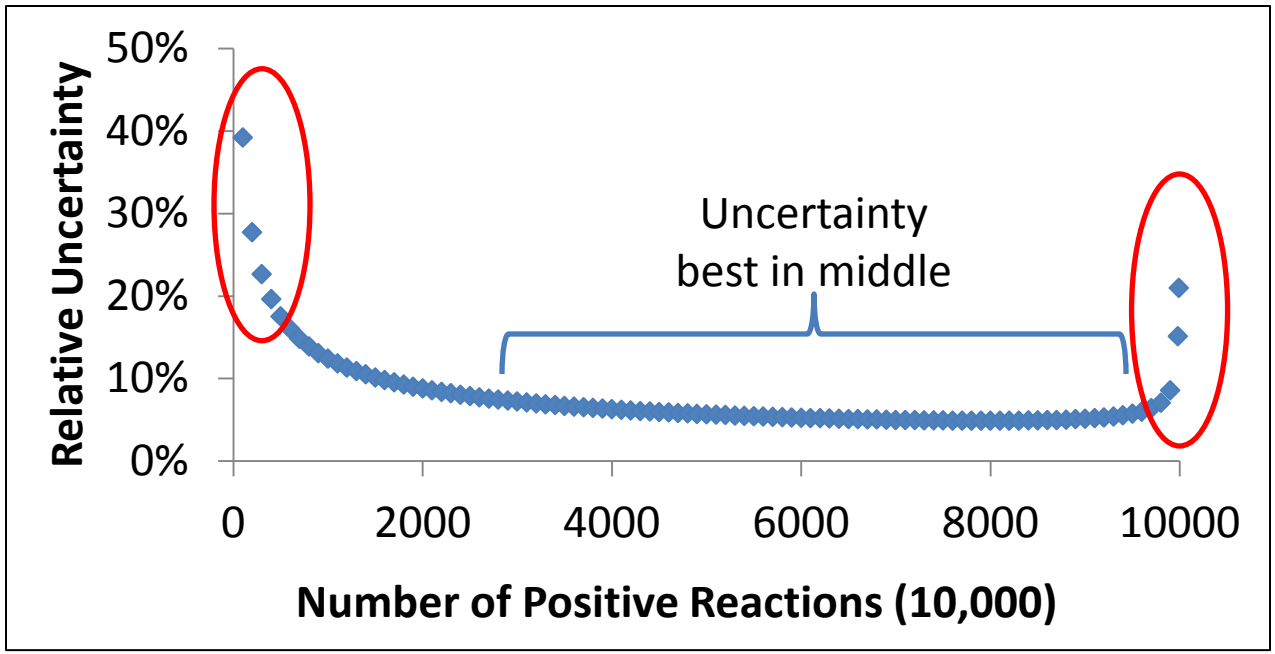
?
pg/uL



$$\text{Concentration } (c/\mu L) = \frac{\cancel{\text{Number of copies}}}{\text{Volume of all PCR reactions}} = \frac{\text{Number of wells}}{\text{Volume of all PCR reactions}} \ln\left(\frac{\text{Number of wells}}{\text{Number of negative wells}}\right)$$

dPCR "Goldilocks Zone"

- Very high and very low proportion of positive PCR reactions give increased uncertainty
- Uncertainty is lowest in the middle



Calculations by Ross Haynes (NIST)

dPCR

- Absolute quantitation of target sequence
- Relies on PCR amplification
  - Only detects specific target DNA or RNA
  - Will not detect proteins or monomers
  - Will not detect fragmented or degraded DNA molecules

Comparison

Quantitative PCR	Digital PCR
Quant is based on amplifiable DNA	Quant is based on amplifiable DNA
Quant is based on a calibrant; as the calibrant goes so will sample values	Quant is based on Poisson sampling statistics (i.e. calibrant free)
Samples must be bracketed by calibrant dilution curve	Samples must be within a range of concentrations
Older technology Widely accepted	New technology Gaining acceptance
Currently less expensive	Currently more expensive

Agenda

- Overview of NIST
- Quantitative PCR versus Digital PCR
- Digital PCR Applications
- Poisson Statistics
- Standard Reference Material 2366
  - Cytomegalovirus for DNA Measurements
- Technology Types
- Measurement Issues

Applications of dPCR

- Absolute quantitation
  - Produce standards (e.g. SRM 2366)
  - Detect level of virus in 1 mL of blood
- Relative quantitation
  - Copy Number Variation (CNV)
  - Relative expression levels
- Rare allele detection
  - Tumor cells in background of wild-type cells
  - Circulating fetal DNA in maternal blood

Absolute Quantitation

- PCR amplify → 1000 reactions
- Count positive wells → 594 reactions amplified
- Poisson stats → 900 copies
- Divide by total volume → 20 µL
- Correct for dilutions → 10 fold dilution
- Concentration → 450 c/µL = (900/20)x10
- Uncertainty is based on binomial statistics → 95 % CI: 415 to 489 c/µL

Relative Quantitation

- Copy Number Variations (CNVs)

$$\text{CNV of gene 1} = \frac{\text{Absolute quant gene 1}}{\text{Absolute quant gene 2}}$$

  - E.g. number of EGFR copies has implications in some cancers*

- Relative Gene Expression Levels

$$\text{Expression of gene 1} = \frac{\text{Absolute quant RNA 1}}{\text{Absolute quant RNA 2}}$$

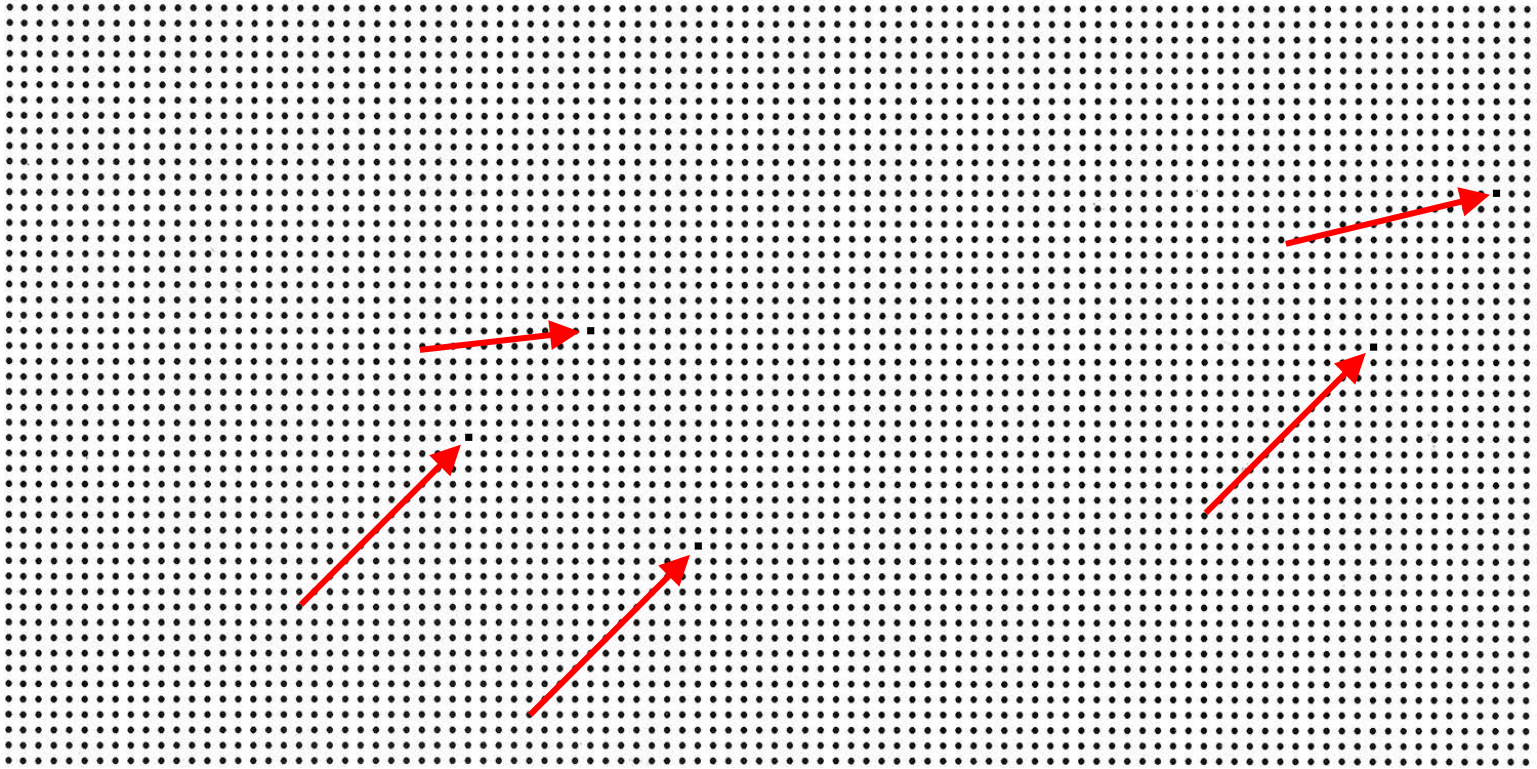
*Cappuzzo et al. J Natl Cancer Inst. 2005 May 4;97(9):643-55.

Rare Allele Detection

- Mutant alleles can be difficult to detect in high background of wild-type
- Mass partitioning of sample reduces the ratio between mutant and wild-type allele in a single PCR reaction
- Can also calculate absolute quant of both mutant and wild type allele

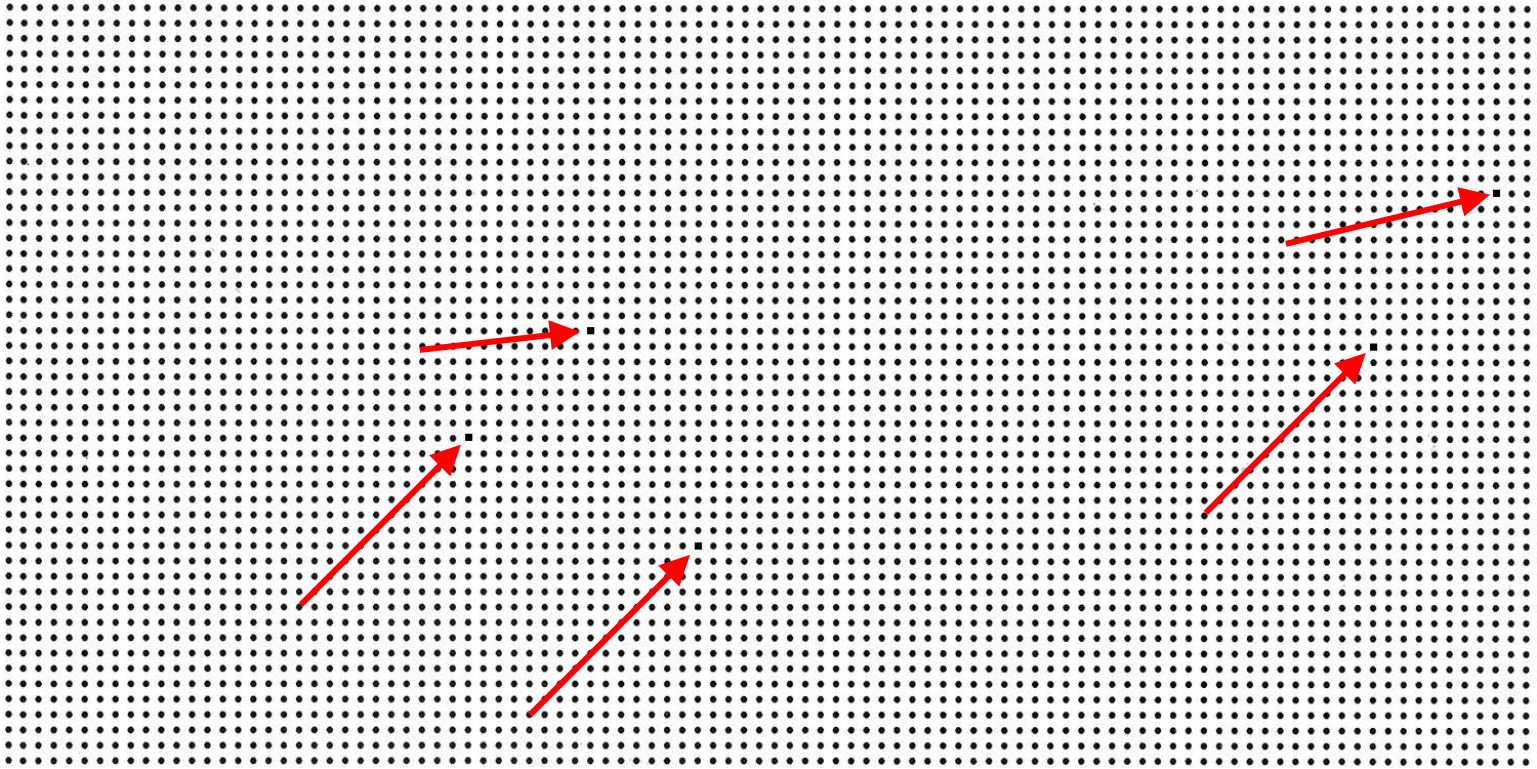
Rare Allele Detection

- There are five (5) squares and 4995 circles
  n = 5000; squares = 0.1 %



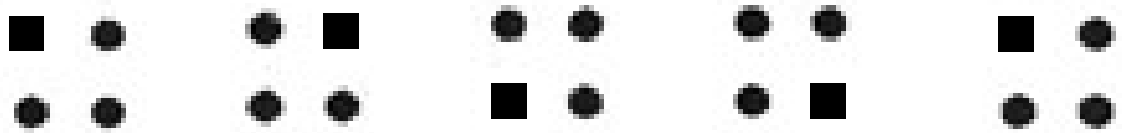
Rare Allele Detection

- If we partition this image into 1250 samples of four shapes each, detection of squares is easier

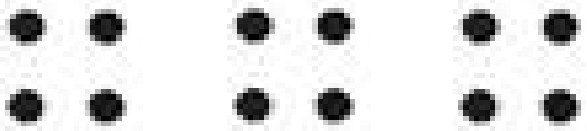


Rare Allele Detection

- These five (5) partitions have one square each
  - Effectively reducing the complexity of the background by subdivision



- The other 1245 partitions have only circles


…

Conclusions for Applications

- dPCR can have a wide variety of applications
  - Absolute quantitation
  - Relative quantitation
    - CNVs
    - Gene Expression
  - Rare allele detection
  - Individual investigation of alleles; haploid typing

Agenda

- Overview of NIST
- Quantitative PCR versus Digital PCR
- Digital PCR Applications
- Poisson Statistics
- Standard Reference Material 2366
  - Cytomegalovirus for DNA Measurements
- Technology Types
- Measurement Issues

Poisson Statistics

dPCR has several assumptions some of which relate to Poisson statistics

Assumptions:

1. Large number of PCR reactions
   - As with most statistics a larger n means more power to discriminate small differences
2. Random distribution
   - It is easier to think in arithmetic division rather than random distribution
3. Independent segregation of molecules
4. Every copy gives a signal
5. Every molecule is dsDNA

Assumption 1
Large number of PCR reactions

- Student's t table n > 120 is considered large*
- Microfluidic technology delivers 100s to 1000s to 1,000,000s of PCR reactions
- Accuracy and precision requirements will dictate number of PCR reactions
- Example of 100 reaction dPCR

*Pagano & Gauvreau "Principles of Biostatistics" 2nd ed. Appendix A pA-10

Assumption 2
Random distribution

- Use Poisson statistics because molecules distribute randomly
- For validating dPCR principles:
  - Random distribution is testable with Ripley's K function
  - Detects clustering or ordering of positive wells
  - See paper below for details

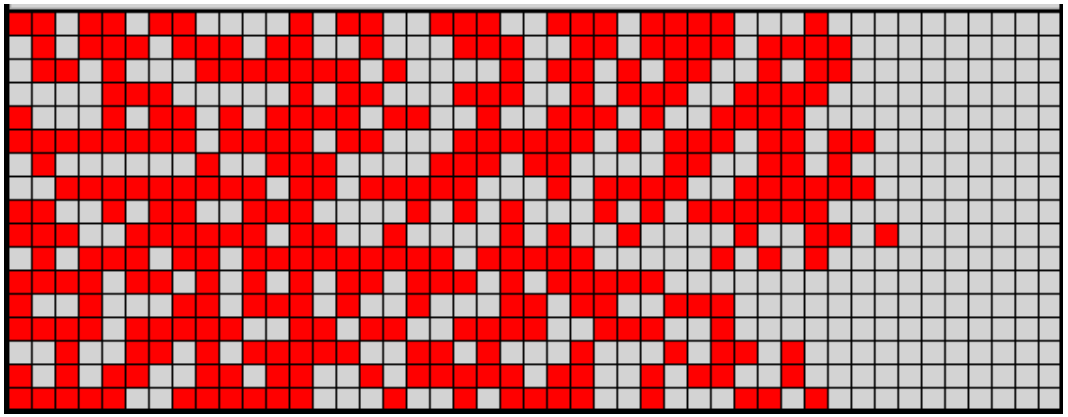
Bhat et al. Anal Bioanal Chem. 2009 May;394(2):457-67.

Assumption 2
Random distribution

- Everyday testing, use eyeball method

This area empty

Clustered



Fluidigm 12.765 digital array 1151-120-151 panel 5

Loaded this direction

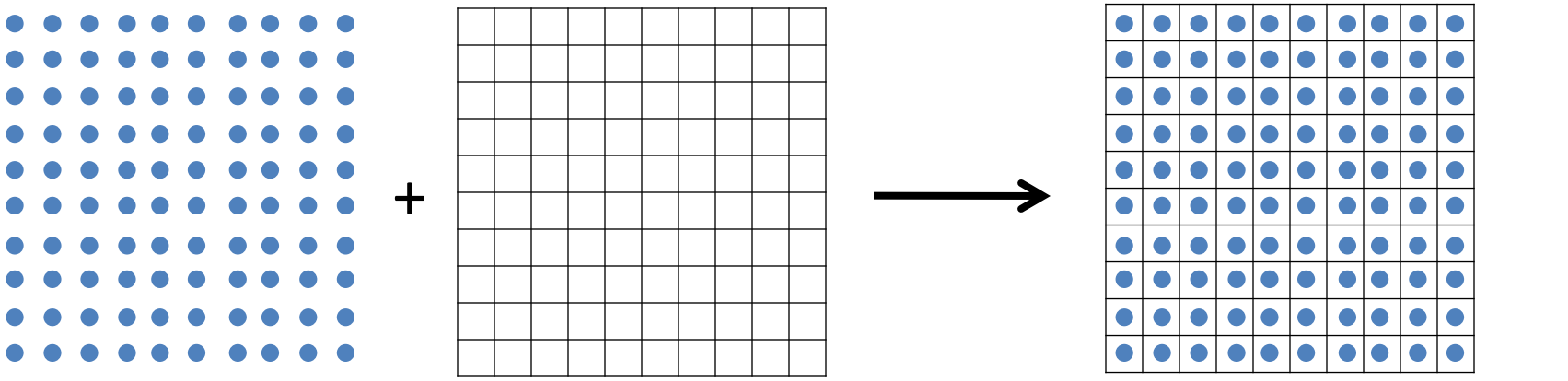
Looks like a loading problem

Ordered

http://www.newretrodining.com/retro_laminates.htm

Even vs. Random Distributions

- Even distribution



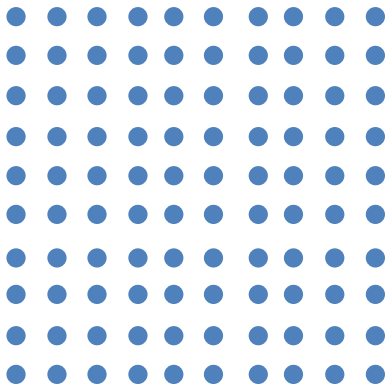
100 molecules

100 wells

All 100 well have exactly one molecule

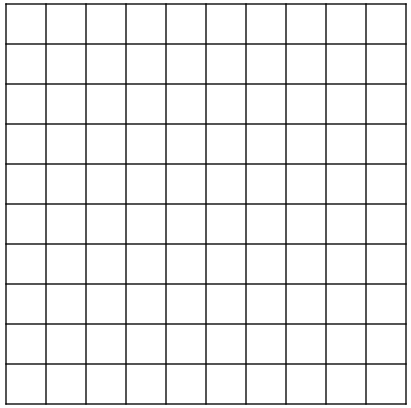
Even vs. Random Distributions

- Random distribution

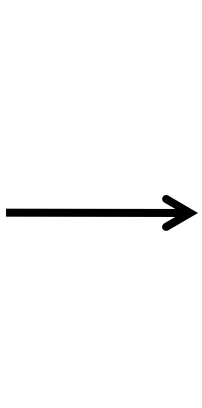


100 molecules

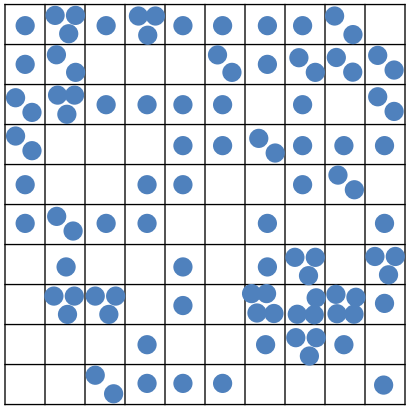
+



100 wells



One possibility

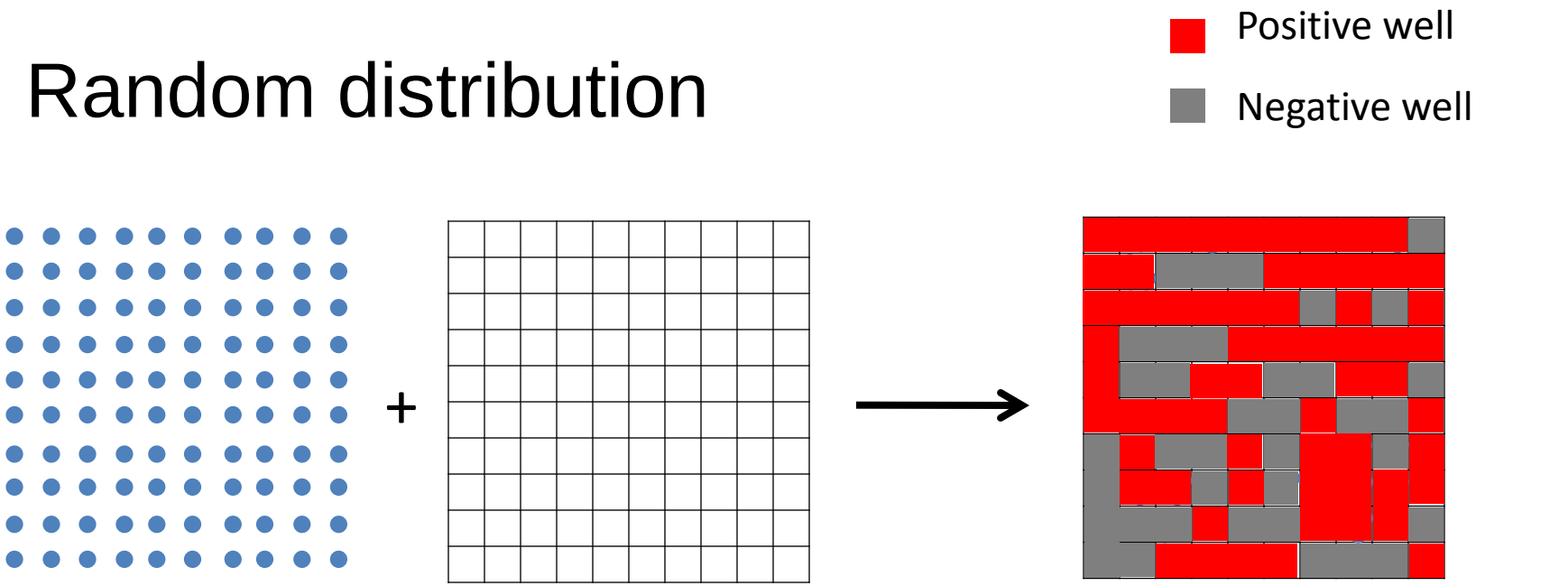


37 wells with 0 molecules
39 wells with 1 molecule
13 wells with 2 molecules
9 wells with 3 molecules
2 wells with 4 molecules
Total = 100 molecules

If we could count the individual molecules we would not need to use Poisson statistics or dPCR

Even vs. Random Distributions

- Random distribution



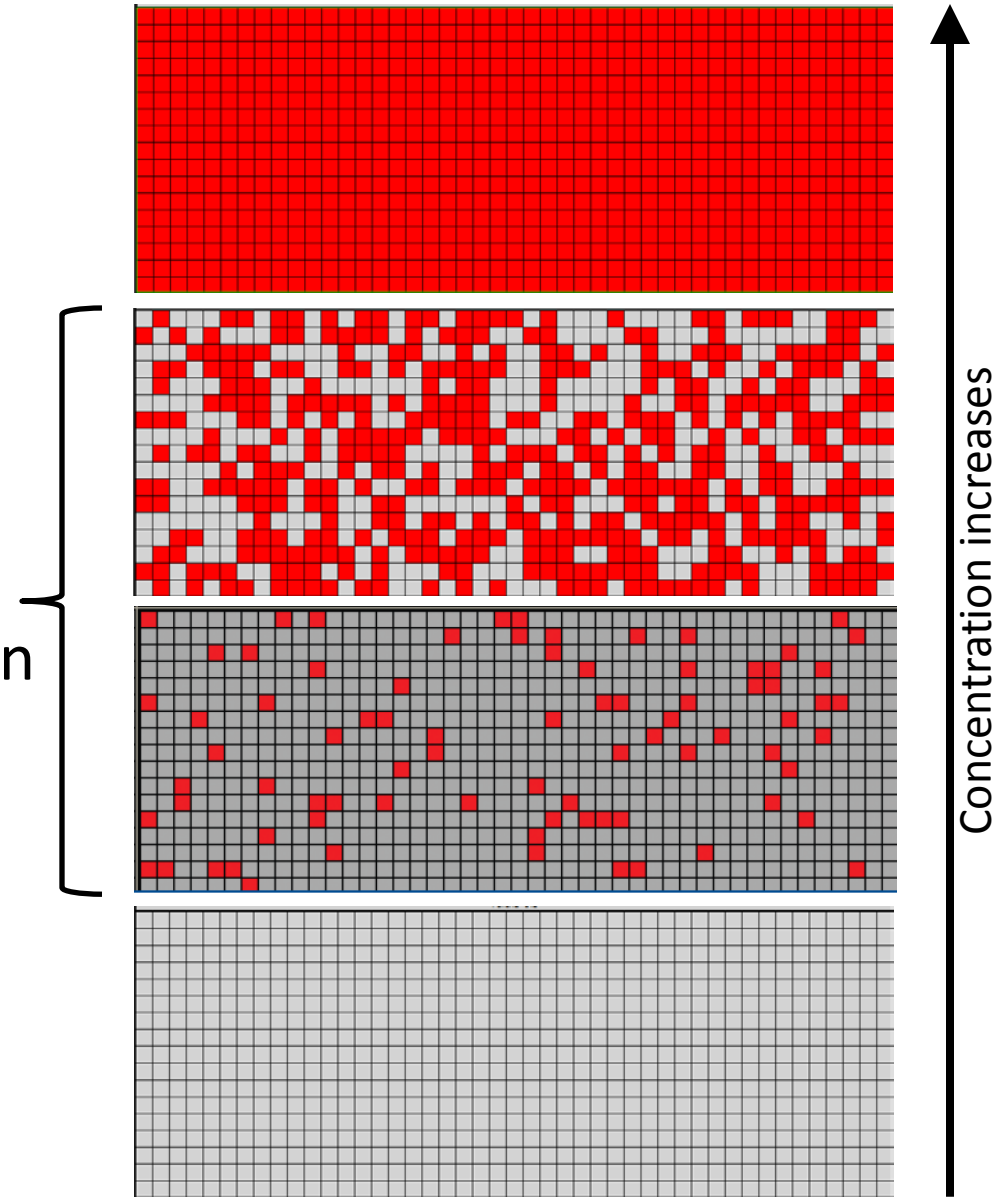
100 molecules

100 wells

37 wells with 0 molecules
63 wells with ≥1 molecules
Poisson calculated 99 molecules
95%CI 77 to 129 molecules

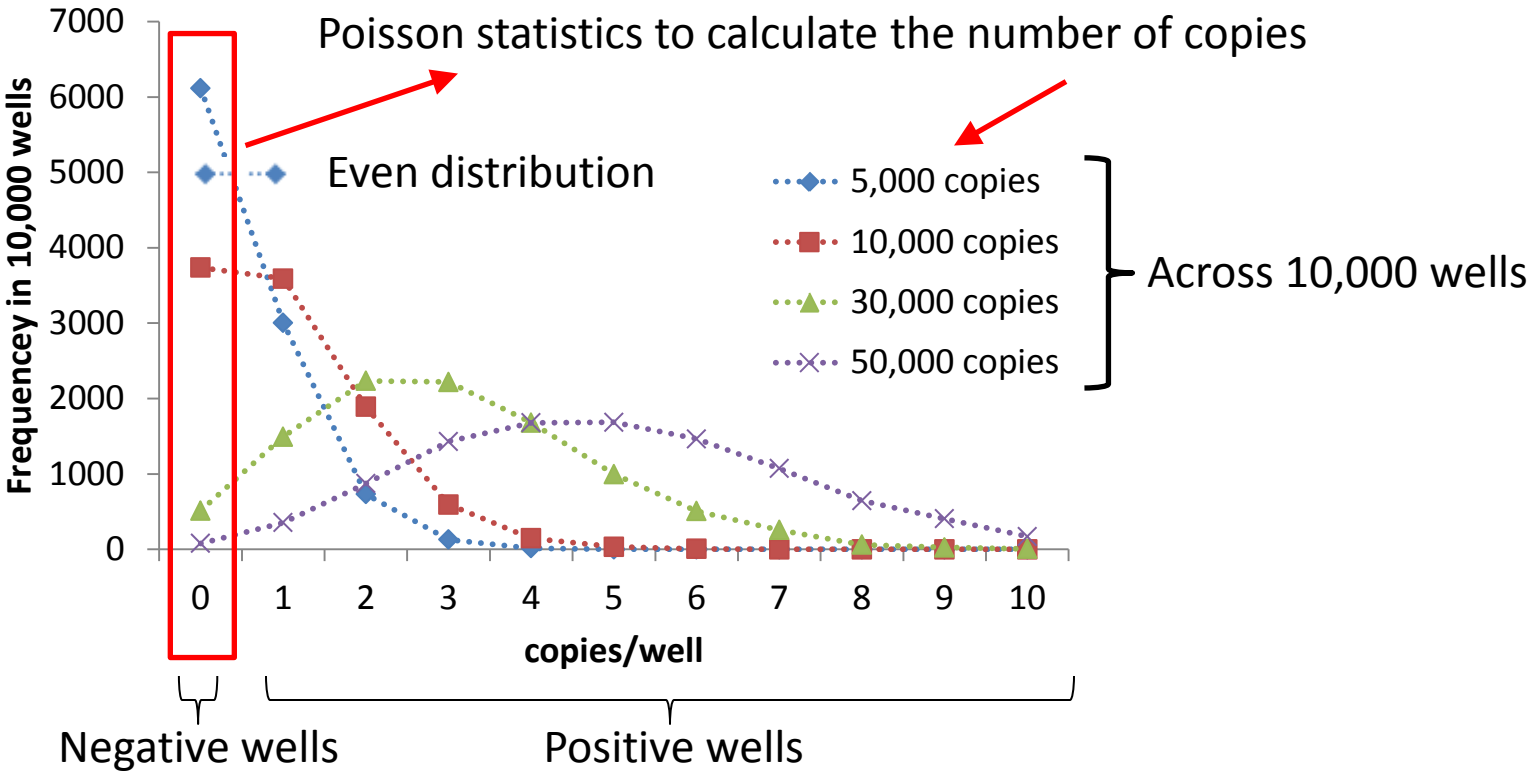
Range of Concentrations

- Saturated
  Every well has at least one copy
- Binary detection
  Calculate concentration
- No amplification
  < 1 copy/total volume



Poisson Statistics

- Poisson Distributions look like a normal distribution crashing into zero
  - When negative values are impossible



Poisson Statistics

- Look it up from a table of values

Concentration = μ/volume of one PCR reaction

Estimate the average number of copies per PCR reaction (μ)

Proportion of negative wells

TABLE A.2
Poisson probabilities

	μ									
k	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0	0.6065	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067

	μ									
k	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
0	0.0041	0.0025	0.0015	0.0009	0.0006	0.0003	0.0002	0.0001	0.0001	0.0000
1	0.0225	0.0149	0.0098	0.0064	0.0041	0.0027	0.0017	0.0011	0.0007	0.0005
2	0.0618	0.0446	0.0318	0.0223	0.0156	0.0107	0.0074	0.0050	0.0034	0.0023
3	0.1133	0.0892	0.0688	0.0521	0.0389	0.0286	0.0208	0.0150	0.0107	0.0076
4	0.1558	0.1339	0.1118	0.0912	0.0729	0.0573	0.0443	0.0337	0.0254	0.0189
5	0.1714	0.1606	0.1454	0.1277	0.1094	0.0916	0.0752	0.0607	0.0483	0.0378
6	0.1571	0.1606	0.1575	0.1490	0.1367	0.1221	0.1066	0.0911	0.0764	0.0631
7	0.1234	0.1377	0.1462	0.1490	0.1465	0.1396	0.1294	0.1171	0.1037	0.0901
8	0.0849	0.1033	0.1188	0.1304	0.1373	0.1396	0.1375	0.1318	0.1232	0.1126
9	0.0519	0.0688	0.0858	0.1014	0.1144	0.1241	0.1299	0.1318	0.1300	0.1251
10	0.0285	0.0413	0.0558	0.0710	0.0858	0.0993	0.1104	0.1186	0.1235	0.1251
11	0.0143	0.0225	0.0330	0.0452	0.0585	0.0722	0.0853	0.0970	0.1067	0.1137
12	0.0065	0.0113	0.0179	0.0263	0.0366	0.0481	0.0604	0.0728	0.0844	0.0948
13	0.0028	0.0052	0.0089	0.0142	0.0211	0.0296	0.0395	0.0504	0.0617	0.0729

1 negative in 10,000

Pagano & Gauvreau "Principles of Biostatistics" 2nd ed. Appendix A pA-6

Poisson Statistics

- $N\lambda = N \ln(N/N-x)$
  - $N\lambda$ = number of copies across all wells
  - N = number of wells
  - x = number of positive wells
  - λ = average number of copies in a well
- Concentration (copies/µL) =

  Number of copies / total volume x dilution factor

$$\text{Concentration (c/µL)} = \frac{\text{Number of wells} \times \ln\left(\frac{\text{Number of wells}}{\text{Number of negative wells}}\right)}{\text{Volume of all PCR reactions}}$$

Heyries et al. Nat Methods. 2011 Jul 3;8(8):649-51.

Uncertainty Calculations

- If number of PCR reactions (N) is large enough, calculate the uncertainty for a binomial

$$\hat{p} \pm z_c \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}$$

$N\lambda = N \ln(N/N-x)$

- Z_c is 1.96 for 95 % confidence interval
- $\hat{p}$ is the proportion of PCR reactions that amplified target; equal to x/N
- Solve for x_{hi} and x_{lo} & plug into Poisson equation

Dube et al. PLoS One. 2008 Aug 6;3(8):e2876.

Absolute Quantitation

- PCR amplify → 1000 reactions
- Count positive wells → 594 reactions amplified
- Poisson stats → 900 copies
- Divide by total volume → 20 µL
- Correct for dilutions → 10 fold dilution
- Concentration → 450 c/µL = (900/20)x10
- Uncertainty is based on binomial statistics → 95 % CI: 415 to 489 c/µL

Uncertainty Calculations

Define parameters:

- N = 1000 reactions
- X = 594 reactions amplified
- Z_c is 1.96 (for 95% CI)
- 10-fold dilution; 20 µL rxn

$$\hat{p} \pm z_c \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}$$

$$0.594 \pm 1.96 \sqrt{\frac{0.594 \text{ X } 0.406}{1000}}$$

0.594±0.0304 = (0.564, 0.624)

X_{lo} = 564 X_{hi} = 624

Copies = Nλ = N ln(N/N-x)

Nλ = 1000 ln(1000/406) = 900

900 c/20 µL x 10 = 450 c/µL

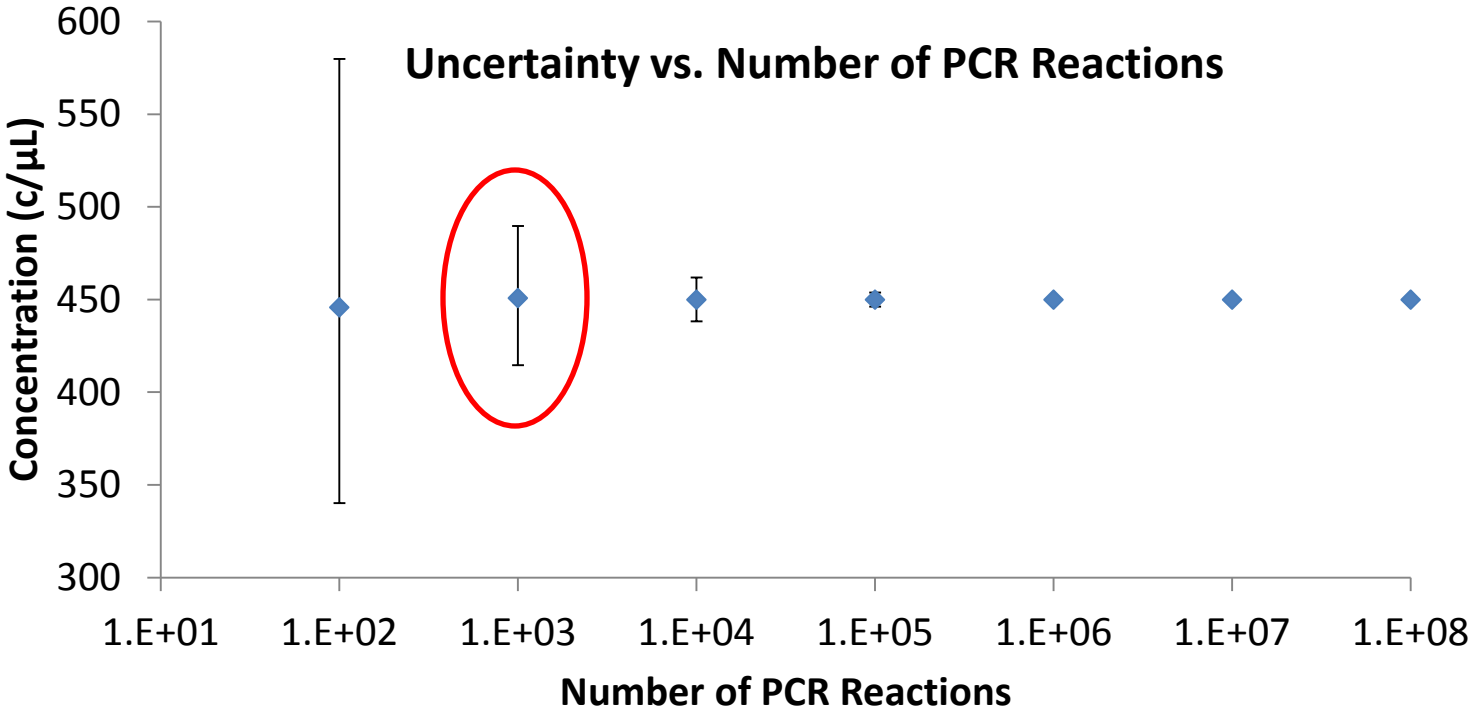
$N\lambda_{hi}$ = 1000 ln(1000/376) = 978

978 c/20 µL x 10 = 489 c/µL

$N\lambda_{lo}$ = 1000 ln(1000/436) = 830

830 c/20 µL x 10 = 415 c/µL

Poisson Uncertainty Depends on Number of PCR Reactions



Number of PCR reactions	n=100	n=1000	n=10k	n=100k	n=1m	n=10m	n=100m
Uncertainty	53.75%	16.67%	5.26%	1.66%	0.53%	0.17%	0.05%

Poisson Statistics

- Monte Carlo simulations of Poisson distribution
- uCOUNT(SM) Digital PCR: University of Utah
- Analysis of dPCR already run or
- Prediction of number of positive reactions based on number of copies
- Preset and user generated number of wells

http://www.dna.utah.edu/ucount/uc.html

Poisson Statistics: Conclusions

1. What constitutes a “large” number of PCR reactions depends on the uncertainty required
   - More PCR reactions will allow discrimination of smaller changes (1:2 vs. 11:12)
2. Poisson Statistics can be used because DNA molecules are randomly distribute
3. Arithmetic division is not an appropriate calculation to determine number of copies

Agenda

- Overview of NIST
- Quantitative PCR versus Digital PCR
- Digital PCR Applications
- Poisson Statistics
- Standard Reference Material 2366
  - Cytomegalovirus for DNA Measurements
- Technology Types
- Measurement Issues

Standard Reference Material 2366

- The first viral DNA SRM produced at NIST is SRM 2366 Cytomegalovirus for DNA Measurements
- Released in 2011
- Certified for concentration using dPCR
- Sequence verified for select regions

SRM 2366

Cytomegalovirus (CMV) for DNA Measurements

- DNA in solution; BAC containing entire genome of the Towne strain
- Components are dilutions of DNA
- Certified for concentration using dPCR

Table 1. Certified Values of the Number of Amplifiable CMV Genome Copies

Component	Value (copies per microliter)	Standard Uncertainty (copies per microliter)	Relative Uncertainty (%)	Expanded Uncertainty 95 % Confidence Interval (copies per microliter)
A	420	56	13.3	301 to 523
B	1702	130	7.6	1446 to 1959
C	19641	365	1.8	18924 to 20359

https://www-s.nist.gov/srmors/view_detail.cfm?srm=2366

SRM 2366

- Select regions sequenced as information value
  - Reference GenBank accession AY315197.2 Towne
  - SRM matched reference except at one base

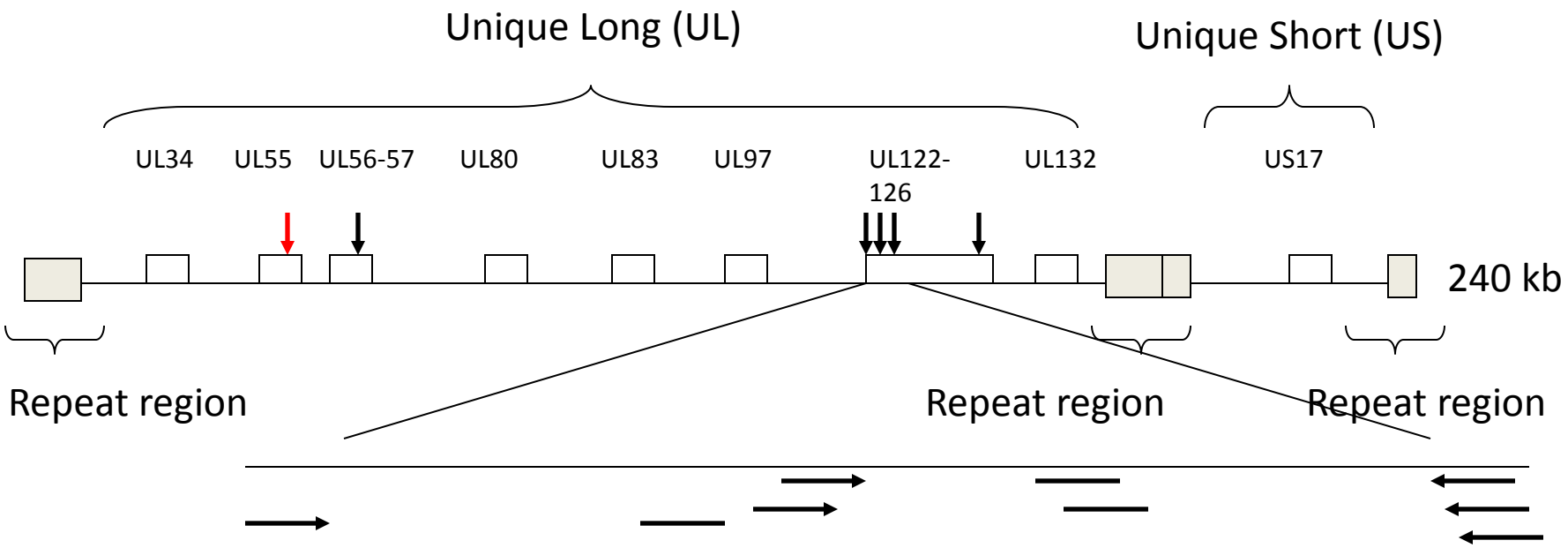
Table 2. Exact Nucleotide Ranges That Were Sequenced for CMV BAC.

Reference Sequences	Nucleotide Range			# Bases Sequenced
UL34	43202	to	44971	1770
UL54[(a)]	77695	to	79992	2298
UL55 to 56	80848	to	82731	1884
UL80	114401	to	116793	2393
UL83	118890	to	119937	1048
UL97	140784	to	142090	1307
UL122 to 126	170525	to	173182	2658
UL132	176380	to	177192	813
US17	198929	to	199312	384

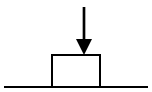
[(a)] The sequence exception at position 78651 (UL54) is described as a Y (C/T) at NIST and a C in GenBank.

https://www-s.nist.gov/srmors/view_detail.cfm?srm=2366

CMV Genome



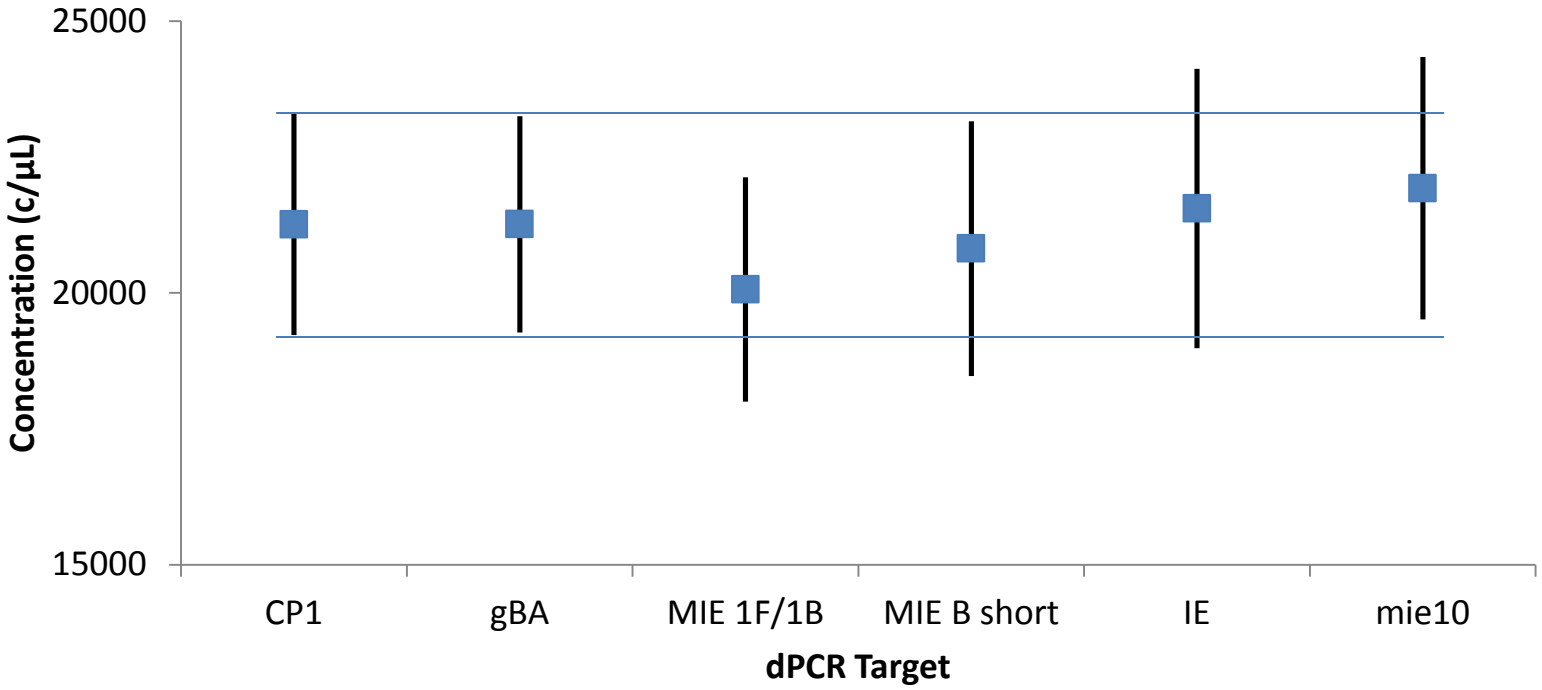
dPCR assay	Size of PCR product	Location
CP1	72 bp	UL55
gBA	254 bp	UL56
IE	127 bp	UL122
MIEB short	138 bp	UL122
MIE 1F/1B	427 bp	UL122
mie10	86 bp	UL125



A dPCR assay site
Regions sequence (17 kb total)

6 dPCR Targets

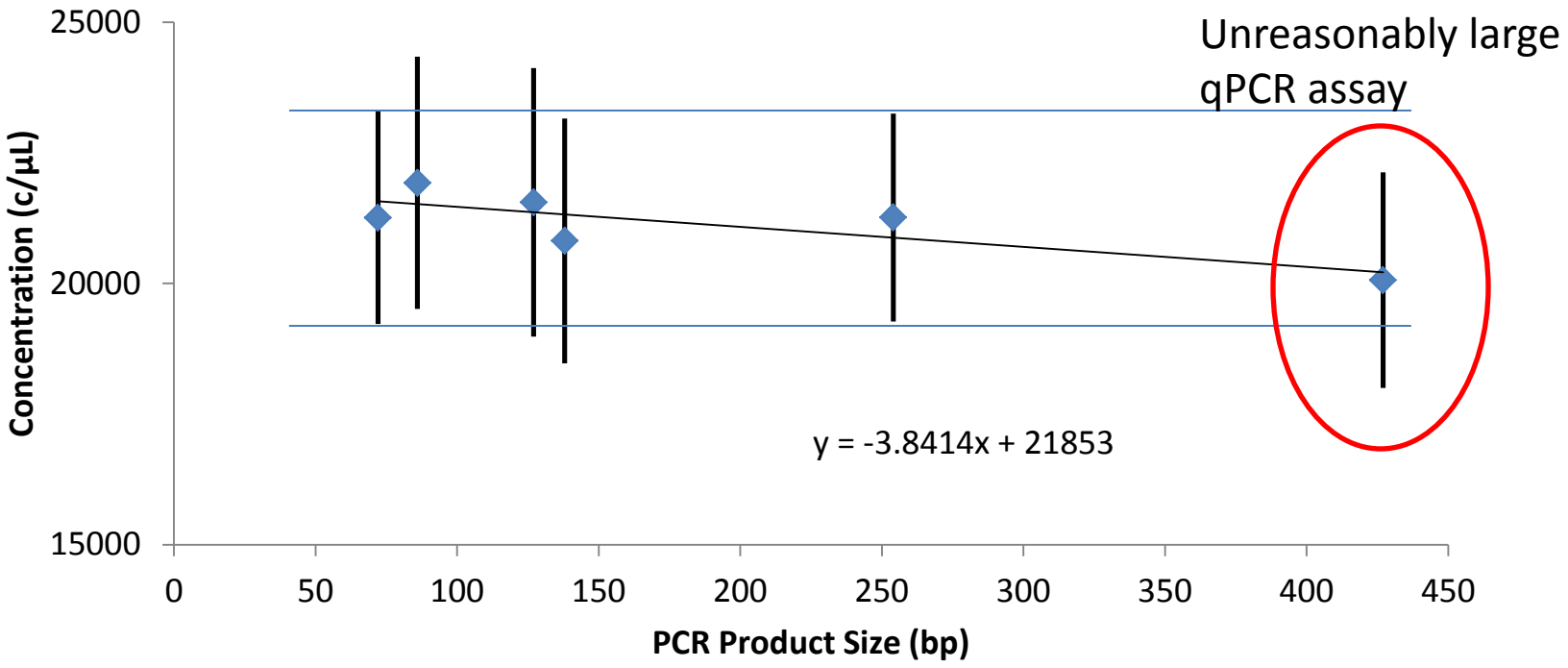
- ANOVA results p-value = 0.43 (no difference)



Error bars ±1 SD; n = 12 panels from Fluidigm 12.765 digital arrays

6 dPCR Targets

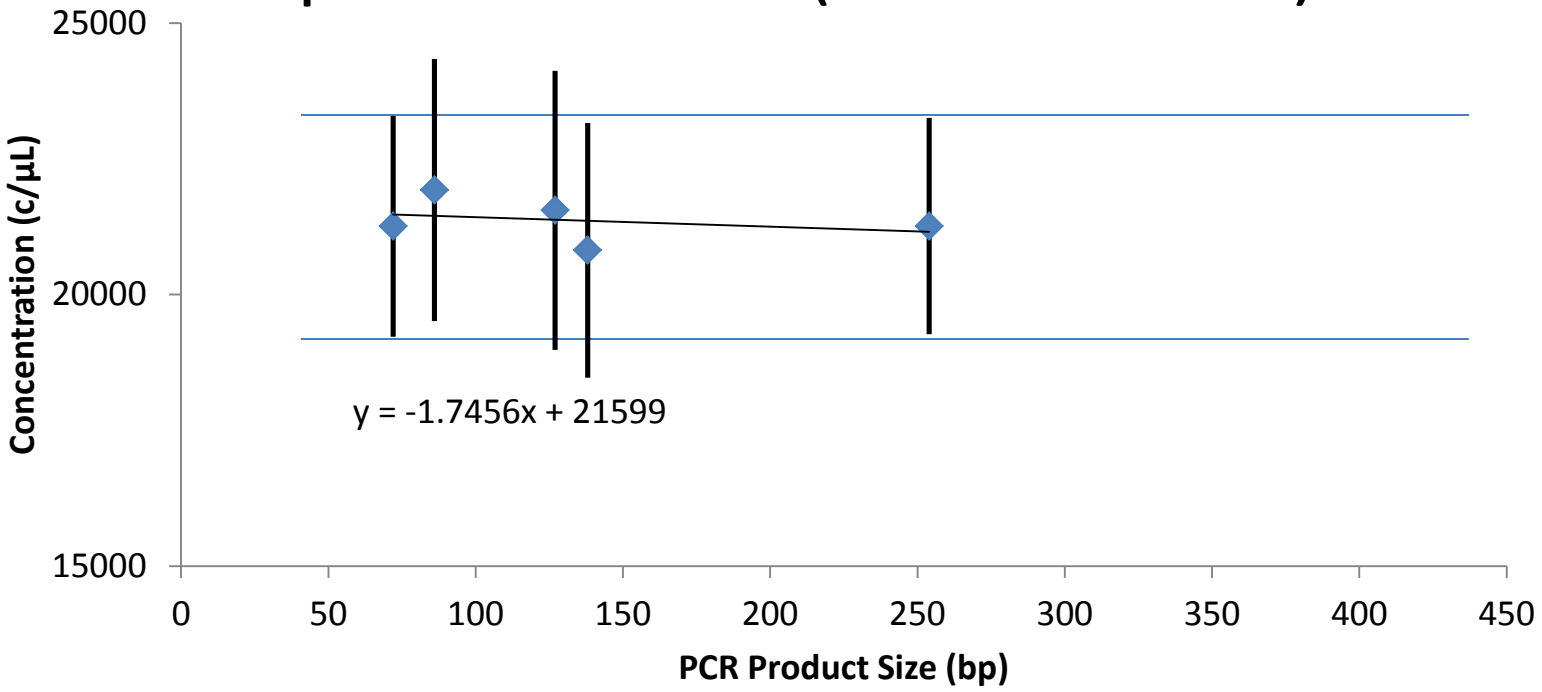
- Correlation p-value = 0.07 (no difference)



Error bars ±1 SD; n = 12 panels from Fluidigm 12.765 digital arrays

~~6~~ 5 dPCR Targets

- Correlation p-value = 0.70 (no difference)
- ANOVA p-value = 0.81 (no difference)



Error bars ±1 SD; n = 12 panels from Fluidigm 12.765 digital arrays

5 dPCR Targets

- Size of PCR product did not appear to correlate with concentration
  - When PCR product size was between 72 bp and 254 bp
- Insufficient evidence to declare the 427 bp PCR product gave significantly different results
- Certification was done using CP1 assay
  - Smallest PCR product

SRM 2366 Measurements

- 99 panels Fluidigm 12.765 digital array
- 75,735 individual PCR reactions (99 x 765)

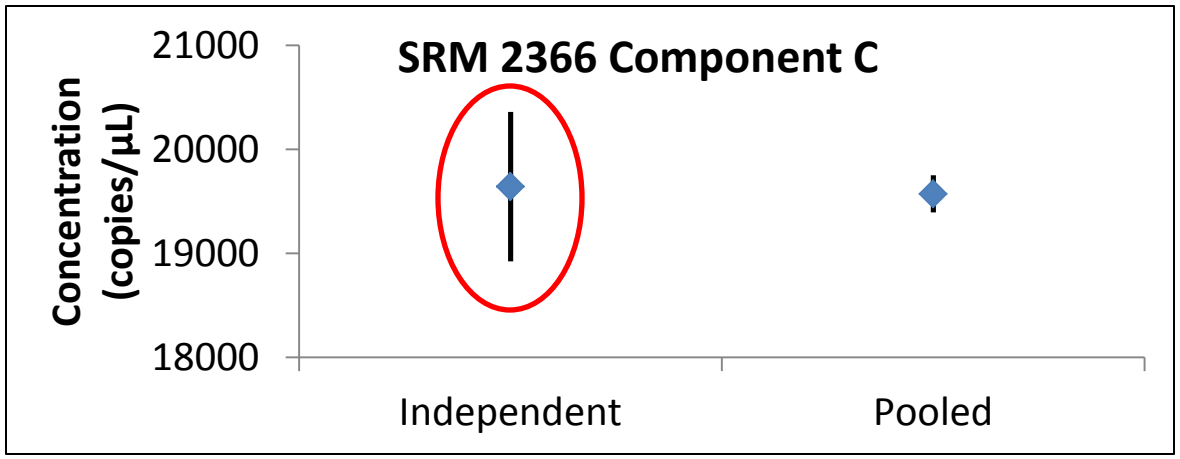
Number of positive reactions

499	550	530	492	511	524	517	510	542
503	553	524	519	517	506	490	514	549
540	553	515	511	517	521	520	533	548
514	522	525	528	506	489	540	546	547
522	518	521	506	515	505	512	524	533
540	523	501	534	506	518	526	541	535
504	506	537	499	499	522	522	558	552
536	516	536	522	507	508	526	556	537
510	521	523	484	498	529	516	527	532
509	529	529	498	525	521	526	515	532
517	553	526	506	503	515	517	547	557

Two Strategies For Combining Data

- Treat each panel (n=99) as independent measurement
- Capture:
  - Run to run variability
- Larger 95 % CI
- Better for conservative measure of uncertainty

- Pool 75735 individual PCR reactions, then calculate Poisson number of copies
- Smaller 95 % CI
- Better for detecting small differences between two samples



SRM 2366: Conclusions

- No statistically significant difference in dPCR assays
  - Used CP1 assay to certify with dPCR
  - Smallest PCR product
- dPCR measurements were treated as independent in order to calculate a more conservative (i.e. larger) uncertainty

Agenda

- Overview of NIST
- Quantitative PCR versus Digital PCR
- Digital PCR Applications
- Poisson Statistics
- Standard Reference Material 2366
  - Cytomegalovirus for DNA Measurements
- Technology Types
- Measurement Issues

Technology Types

- Types of dPCR systems
- Advantages of each
- Examples of advantages
- Comparison of specifications

Technology Types

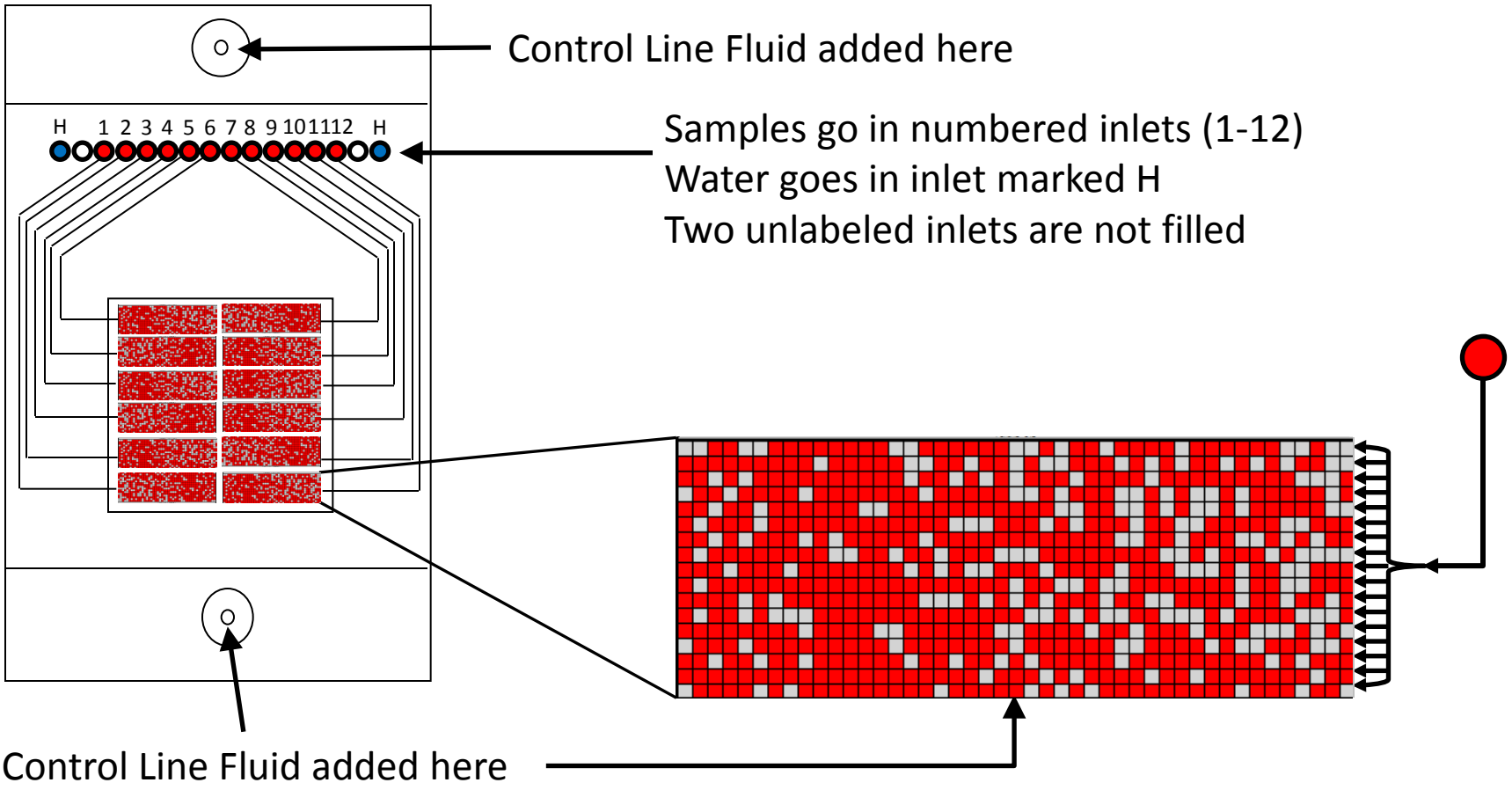
- Pre-manufactured microfluidic chambers
  - wells are pre-machined and static in space
    - Fluidigm – BioMark™ & EP1™
    - Life Technologies – QuantStudio® 12K Flex
- Emulsion based chambers
  - oil in water emulsion with reactions of same size
    - BioRad – QX100™
    - RainDance – RainDrop™ System

*This is not an exclusive list of manufacturers. This is only a list of known manufacturers by the author at the time this presentation was created.

Pre-manufactured microfluidic chambers

- Microfluidic technologies used to aliquot sample into massive number of PCR reactions
- Geometry of the well dictates volume of chamber – can not fit 20 nL in 10 nL space
- Chambers are fixed in space – therefore images can be taken after each cycle, just like real-time qPCR
  - Troubleshooting
  - Assay optimization
  - Multi-purpose instruments (not just dPCR)

Fluidigm 12.765 Digital Array



See a video here: http://www.youtube.com/watch?v=UwzDc6wcGZg&feature=relmfu

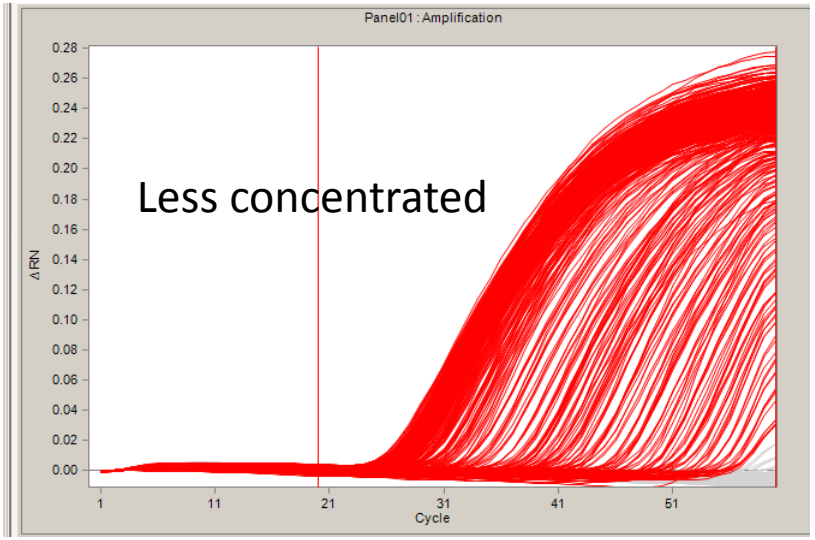
Troubleshooting

- Intact and linear plasmid diluted the same amount are run on dPCR

Intact plasmid DNA

Positive chambers = 429

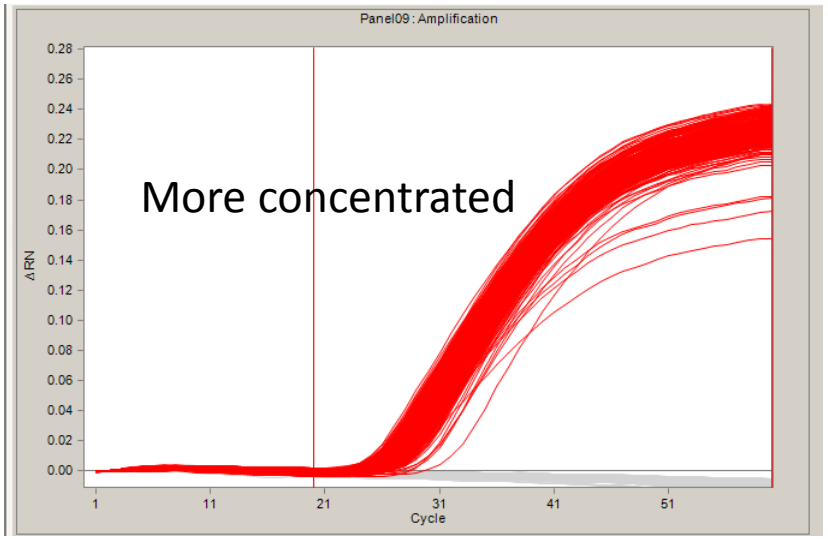
Conc. 1.2 x 10^6 copies/µL



Linear plasmid DNA

Positive chambers = 534

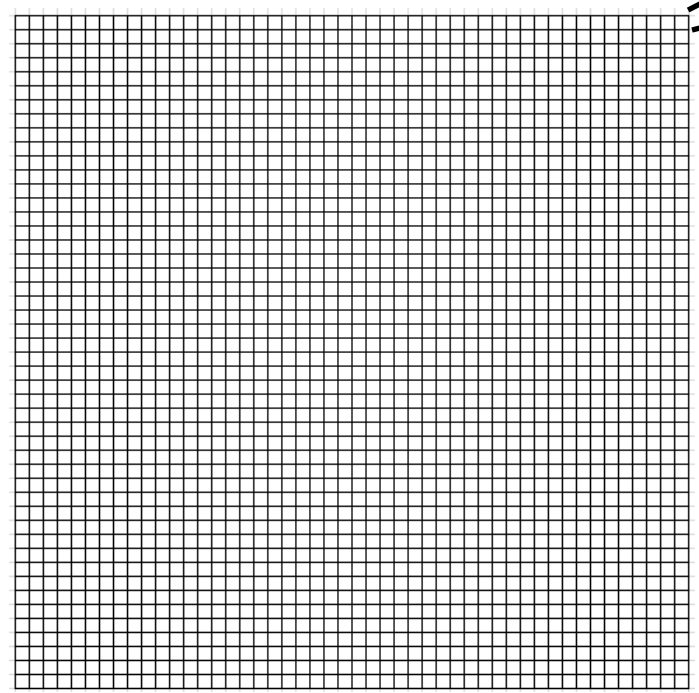
Conc. 2.0 x 10^6 copies/µL



Would more chambers be positive if more cycles were run?
Is this a true difference in concentration?

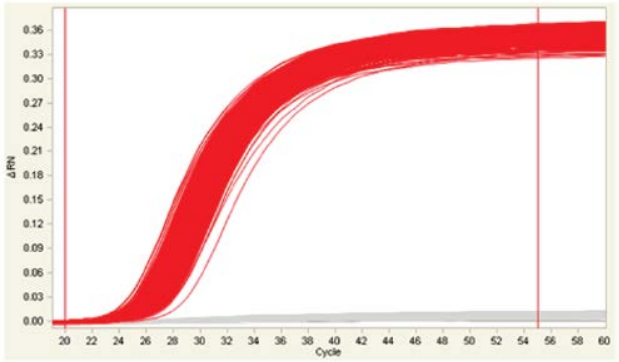
Multi-Use Instrument

48 samples



48 qPCR assays

2304 qPCR curves



http://www.youtube.com/watch?v=s9HUhuCbbhU

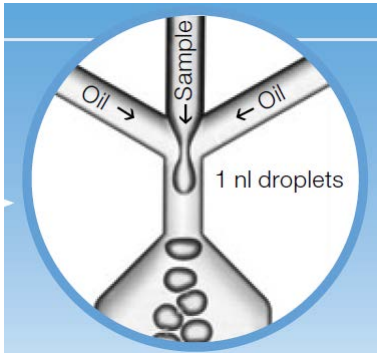
Emulsion based chambers

- Emulsion generator produces water-in-oil emulsion chambers of equal size
- Geometry of the generator and stability of emulsion determines chamber size
- Chambers are not fixed in space – end-point detection only
  - Cheaper to create emulsion chambers
  - More reactions – better statistics
  - All dPCR technologies use end-point data for Poisson calculations

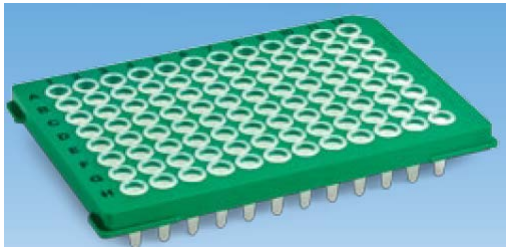
BioRad QX-100

1) Create a PCR mastermix as if for qPCR

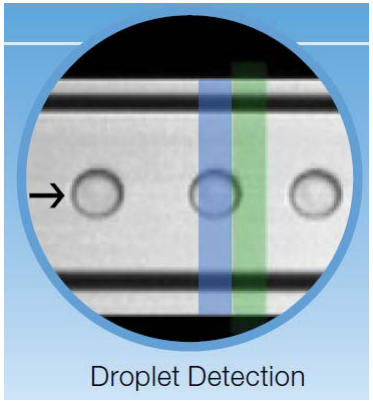
2) Generate 20,000 emulsion reactions



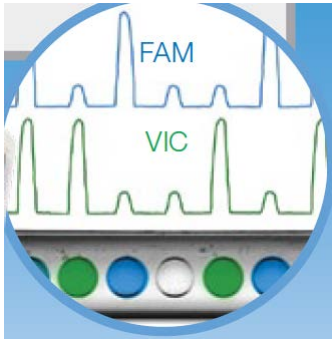
3) PCR in standard thermal cycler



4) Measure fluorescence of individual PCR reactions



5) Analyze reactions for presence or absence of amplification & perform stats



http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6237.pdf

Emulsion based chambers

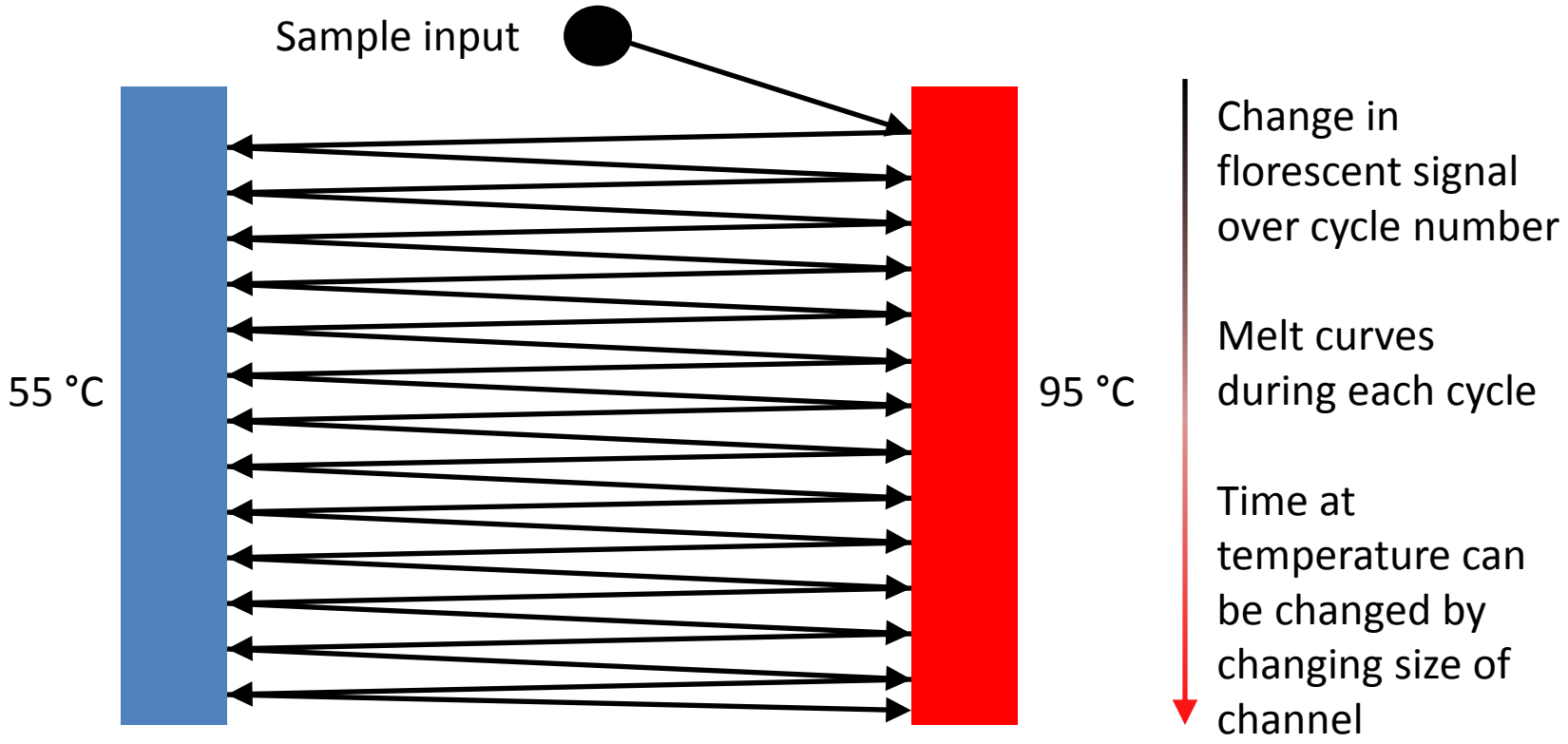
Troubleshooting is not impossible

- Validate different sample types (plasmid, genomic, viral, etc) for number of cycles required
- Restriction digestion has been shown to improve qPCR and dPCR efficiency*
- Additional confidence can be gained by using multiple target genes across the genome

*Bhat et al. Anal Bioanal Chem. 2009 May;394(2):457-67.

Update: continuous detection possible

- "University of Utah looks to commercialize continuous flow microfluidic PCR"



http://www.genomeweb.com//node/1125071?hq_e=el&hq_m=1345992&hq_l=3&hq_v=88d133a09c

Comparison of Specifications

System	Fluidigm	Life Technologies	BioRad	RainDance Technologies
Type	Pre-MFR'ed	Pre-MFR'ed	Emulsion	Emulsion
Reactions per sample	765 or 770	64 to 9,216	20,000	10 million
Number of samples	12 or 48	1 to 144	8 to 96	8 to 96
Total number of reactions	9,180 or 36,960	9216	160,000 to 1.9 million	80 million to 960 million
Detection	Real-Time or End point	Real-Time	End Point	End Point
Reaction volume	6 nL or 0.85 nL	33 nL	1 nL	1 pL
Total volume	4.6 µL or 4µL	304 µL	20 µL	25 µL
Cost per run	$320[%]	$150[&#]	$40*[1]	$240*[2]
Cost per 10,000 reactions	$348 or $86	$488	$2.50	$0.03

[%]based on quote for 12.765 digital arrays 2012
[&]list price Sept 2012
[#]based on one OpenArray (n=3072 reactions)

*cost for 8 samples
[1]based on quote 2012
[2]based on pre-commercial market research

Technology Types: Conclusions

Lab requirements may influence system requirements:

- Uncertainty
- Throughput
- Cost
- Space
- Sample types

Emulsion	Pre-manufactured
Cheaper to Create → More reactions → More Power to Discriminate Small Differences	More expensive per reaction → Fewer Reactions → Less Power to Discriminate Small Differences
Poisson calculations only use end-point data	Real-time data collection gives more data for troubleshooting and optimizing
Single-purpose instrument	Multi-purpose instrument (space saver)

Agenda

- Overview of NIST
- Quantitative PCR versus Digital PCR
- Digital PCR Applications
- Poisson Statistics
- Standard Reference Material 2366
  - Cytomegalovirus for DNA Measurements
- Technology Types
- Measurement Issues

Measurement Issues

As with any technology there are potential measurement issues with dPCR

1. Caveats of dPCR setup
2. Assumptions of dPCR
   - What can happen if assumptions are not met

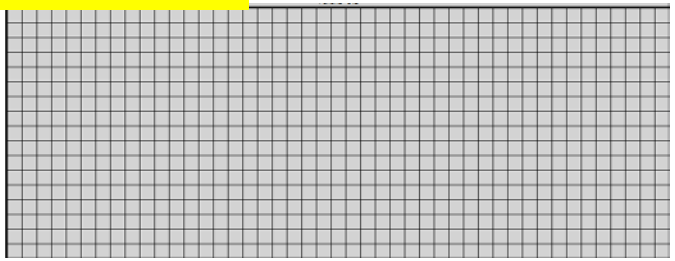
dPCR

1) Create a PCR mastermix as if for qPCR

Manufacturer specific reagents required:
E.g.:
- Fluidigm 20X loading reagent
- BioRad ddPCR master mix

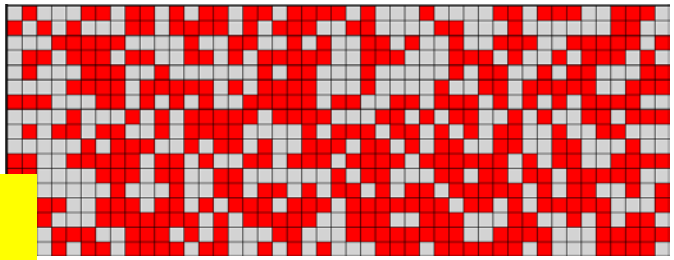
2) Aliquot across 100s or 1000s of wells

Number of wells is system specific and may not be user changeable



3) Thermal cycle as if for qPCR & count wells with amplification

Only end-point data is used for Poisson calculations
Real-time data may help with optimization and troubleshooting



4) Use Poisson statistics to determine concentration of starting material

All assumptions must be met for Poisson calculations to be **valid**

Measurement Issues

Assumptions:

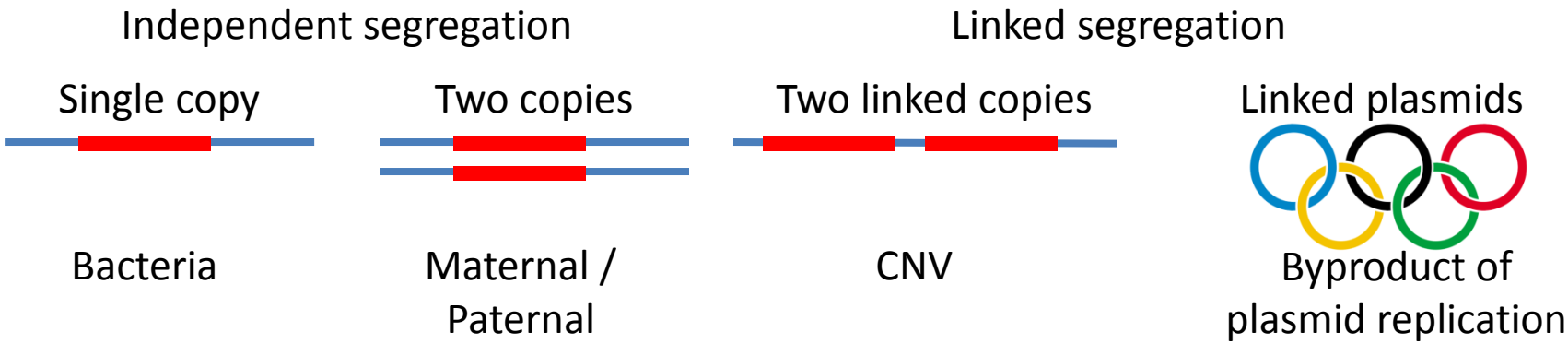
1. Large number of PCR reactions
2. Random distribution
3. Independent segregation of molecules
4. Every copy gives a signal
5. Every molecule is dsDNA

Assumptions 1 & 2

1. Large number of wells
   - "Large" depends on the uncertainty required
     - PCR reactions required to discriminate a CNV 10:11 is different that CNV 2:3
2. Random distribution
   - Use Poisson statistics
   - Testable with Ripley's K function
     - Not a everyday use statistic

Assumption 3
Independent Segregation

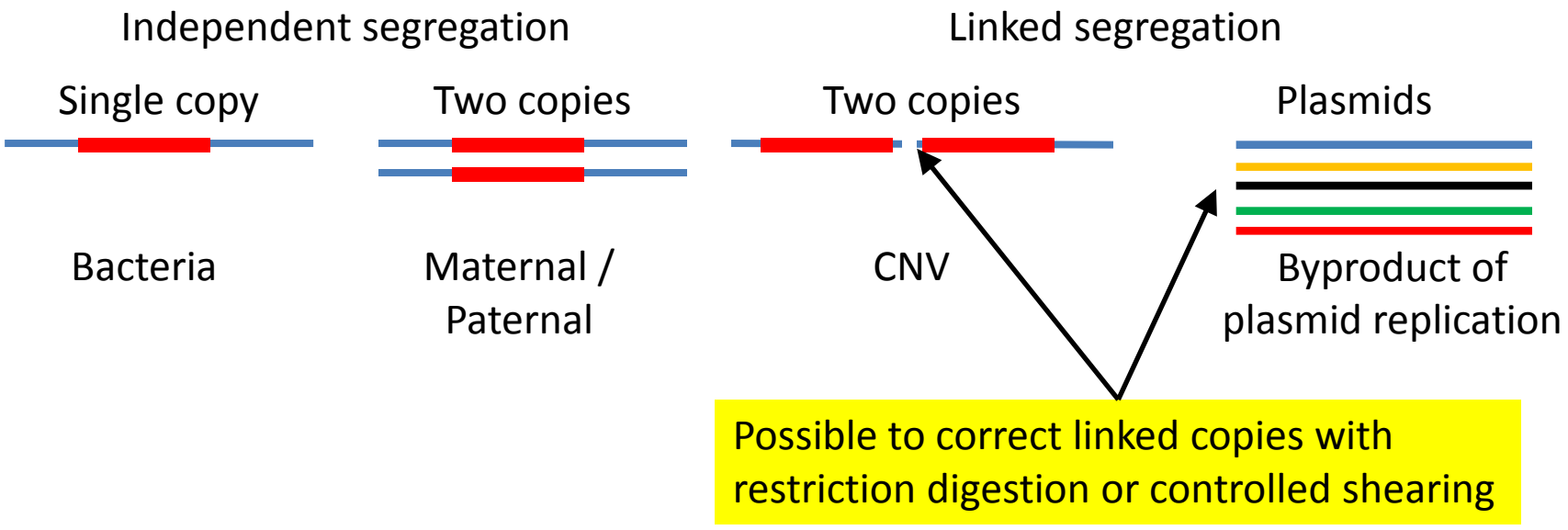
- If molecules are concatemers or physically bound to one another one “amplification-forming-unit” will consist of multiple copies
  - i.e. two linked copies will be counted as one copy



Possible to correct linked copies with restriction digestion or controlled shearing

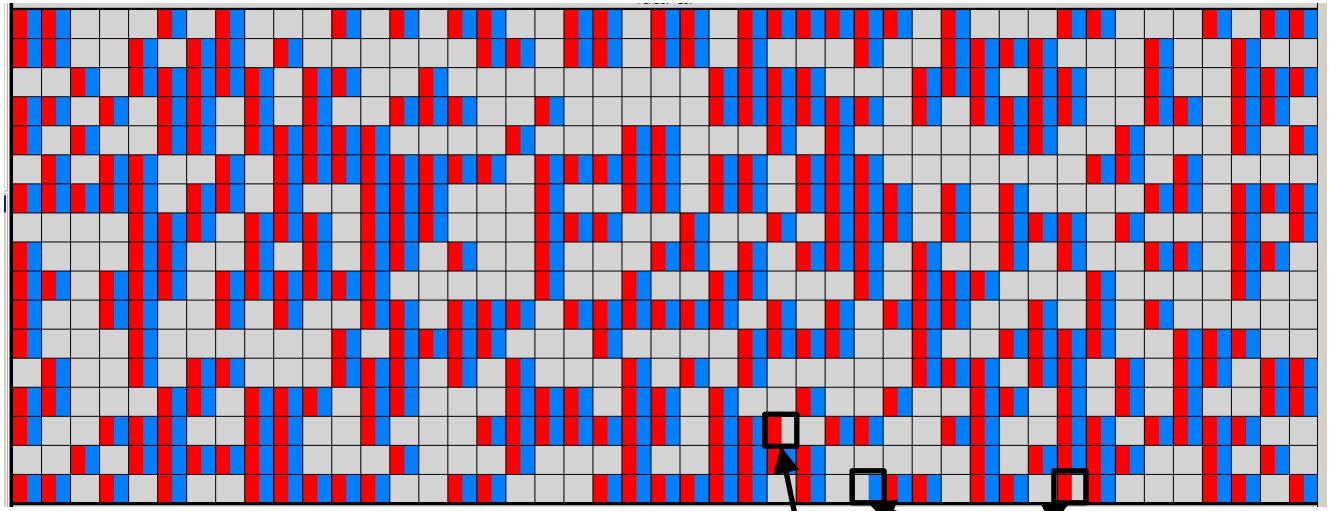
Assumption 3
Independent Segregation

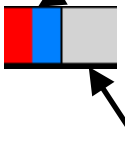
- If molecules are concatemers or physically bound to one another one “amplification-forming-unit” will consist of multiple copies
  - i.e. two linked copies will be counted as one copy



Assumption 4
Every Copy Gives a Signal

- Two linked genes 
- Duplex dPCR: each PCR reaction should have both or neither





Both genes

Neither

Linear plasmid ~0.7 % only one target was detected
Supercoiled plasmid ~3 %

Red Gene only

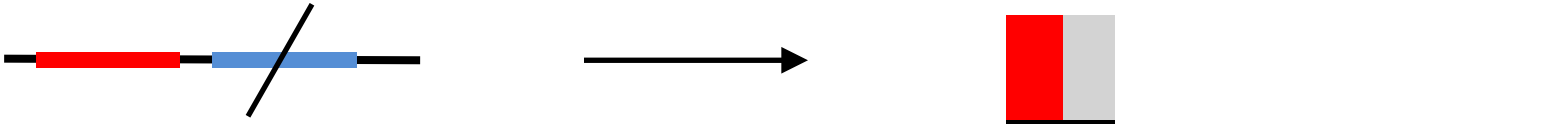
Blue Gene only

Data generated by Ross Haynes (NIST)
IRMM CRM-AD413 certified to have 1:1 ratio of HMG and MON810 Maize genes

Assumption 4
Every Copy Gives a Signal

Plasmid type	Only one target detected	Possible cause(s)
Linear	~0.7 %	dPCR bias Fragmentation
Supercoiled	~3 %	dPCR bias Inaccessibility of target due to supercoiling

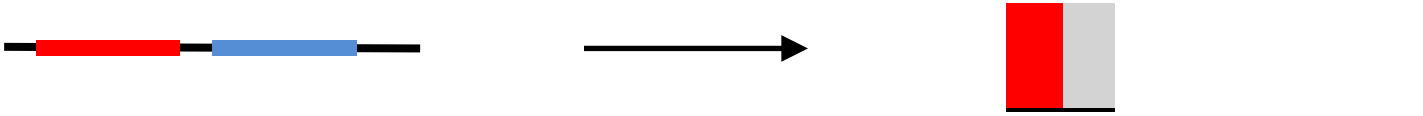
Fragmentation



If blue gene is cut

Then only the red gene will be detected

dPCR bias



Both genes present

But only the red gene is detected

Assumption 4
Every Copy Gives a Signal

Plasmid type	Only one target detected	Possible cause(s)
Linear	~0.7 %	dPCR bias Fragmentation
Supercoiled	~3 %	dPCR bias Inaccessibility of target due to supercoiling

Supercoiling

Reagents can not “see” blue gene

Then only the red gene will be detected

dPCR bias

Both genes present

But only the red gene is detected

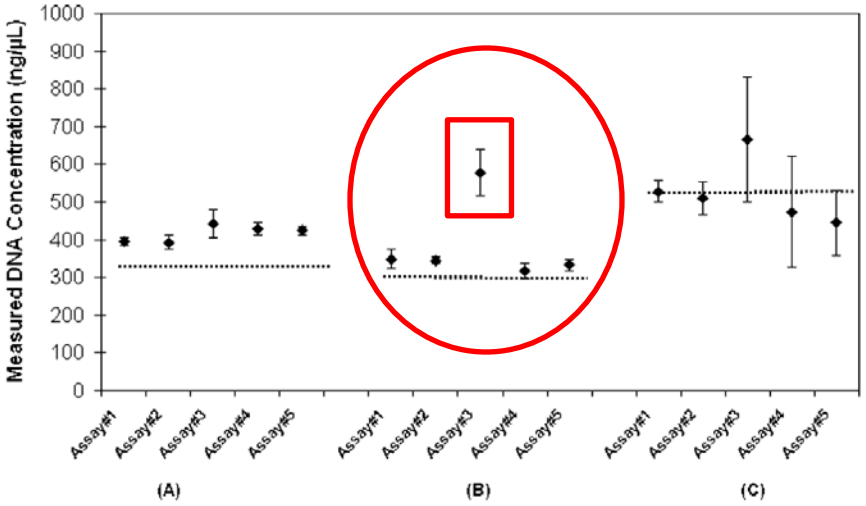
http://bitesizebio.com/articles/is-supercoiled-dna-derailing-your-cloning/

Assumption 4
Every Copy Gives a Signal

- Extraction method may leave PCR inhibitors
  - Also a problem for qPCR
  - Some direct extraction methods use alkaline solution to free DNA from cell
    - DNAzol® Direct: 10-fold dilution of sample into master mix required to avoid PCR inhibition

Assumption 5
Every Molecule is dsDNA

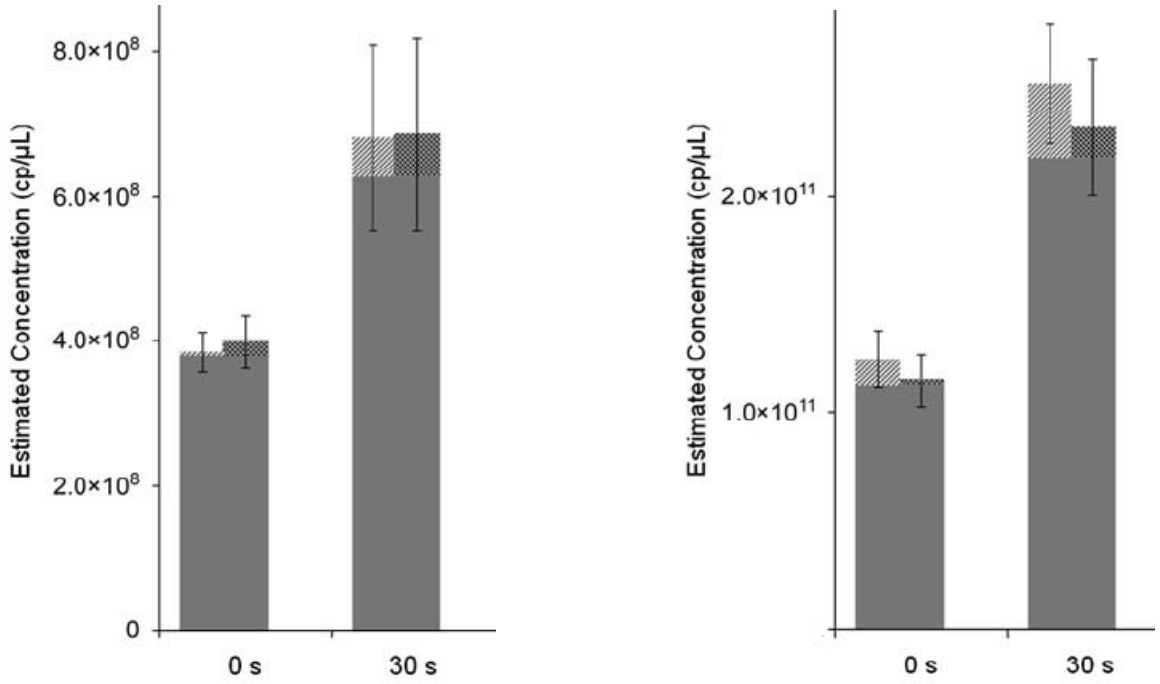
- NMI Australia:
  - Five (5) targets across genome used to quantify amount of DNA; one assay gave a 2-fold increase in concentration
  - Traced to low local GC-content; ssDNA



Bhat et al. Anal. Chem. 2010, 82, 7185-7192

Assumption 5
Every Molecule is dsDNA

- NMI Australia:
  - dsDNA vs. ssDNA gave 2-fold difference in concentration (95 °C for 30 sec and snap cooled)



Bhat et al. Analyst. 2011 Feb 21;136(4):724-32.

Assumption 5
Every Molecule is dsDNA

- Extraction method may leave ssDNA
  - Some direct extraction methods use alkaline solution or heat inactivating enzymes
    - DNAzol® Direct – alkaline solution
    - ZyGEM forensicGEM™ – heat inactivation of protease
- Options:
  - Validate method that produces 100% ssDNA then apply a 2-fold factor to concentration calculations
  - Validate method that produces 100% dsDNA

Holden et al. J Agric Food Chem. 2009 Aug 26;57(16):7221-6.

Measurement Issues: Conclusions

- Optimization may still be required when changing qPCR assays over to dPCR
- Poisson statistics require assumptions to be met
  - "Large" number of reactions
    - Depends on your needs or applications
  - Random distribution of molecules
    - Allows use of Poisson statistics
  - Molecules must segregate independent
    - Otherwise one amplification-forming-unit may contain many copies
  - Possible to underestimate concentration if not every target amplifies
  - Possible to overestimate concentration if DNA has significant portion of ssDNA
    - Each molecules of ssDNA has two amplification-forming-units for every one dsDNA molecule

Overall Conclusions

- NIST produces Standard Reference Materials
- dPCR is a calibrant free method
- Applications:
  - Absolute quantitation
  - Relative quantitation (DNA or RNA)
  - Rare allele detection
  - Investigation of individual alleles; haploid typing
- Poisson Statistics:
  - “Large” depends on uncertainty required
  - More PCR reactions give power to differentiate smaller differences in concentration
  - Used because molecules distribute randomly

Overall Conclusions

- SRM 2366: Cytomegalovirus for DNA Measurements
  - 5 dPCR targets statistically not different
  - Treated data as independent measurements to calculate a more conservative (i.e. larger) confidence interval
- Technology Types:
  - Laboratory needs will dictate “best” system
  - Emulsion-based: more PCR reactions → more power to discriminate small differences in concentration
  - Pre-manufactured chambers: real-time data helps optimize and troubleshoot
    - Multipurpose instrument (e.g. SNP detection, 96-well qPCR)

Overall Conclusions

- Measurement Issues:
  - Extraction issues:
    - ssDNA left by heat or alkaline solution
    - PCR inhibitors
  - Sample type:
    - Difference in PCR efficiency → bias in dPCR measurement
    - E.g. supercoiled versus linear plasmid
  - Linked copies:
    - Measured per amplification-forming-unit
    - Fixed by restriction digestion or controlled shearing
  - Number of copies:
    - ssDNA gives two amplification-forming-units per dsDNA

Questions

Ross Haynes
Ross.Haynes@nist.gov

This presentation will be available online at
~~http://www.nist.gov/mml/biochemical/genetics/clinical_dna.cfm~~
http://www.nist.gov/mml/bmd/genetics/clinical_dna.cfm
Or Google “CDIR NIST”

Select References (1)

- Ruano et al. Proc Natl Acad Sci U S A. 1990 Aug;87(16):6296-300.
  - Single molecule PCR for separating maternal and paternal chromosomes before sequencing – heredity
- Monckton et al. Genomics. 1991 Oct;11(2):465-7.
  - Single molecule PCR for separating maternal and paternal chromosomes before sequencing – human ID
- Vogelstein et al. Proc Natl Acad Sci U S A. 1999 Aug;96(16):9236-41.
  - Investigation of rare mutant KRAS alleles
  - First paper to suggest dPCR could be used for quantitating DNA
- Dube et al. PLoS One. 2008 Aug 6;3(8):e2876.
  - Uncertainty calculations for digital PCR

Select References (2)

- Bhat et al. Anal Bioanal Chem. 2009 May;394(2):457-67.
  - Ripley's K to measure random distribution of molecules
  - Restriction digestion to increase dPCR efficiency
- Holden et al. J Agric Food Chem. 2009 Aug 26;57(16):7221-6.
  - Discusses how ssDNA can affect DNA quantitation
- Bhat et al. Anal. Chem. 2010, 82, 7185-7192
  - ssDNA can over estimate concentration of dPCR
- Bhat et al. Analyst. 2011 Feb 21;136(4):724-32.
  - Heating of DNA can effect dPCR, extended heating can damage DNA

Select References (3)

- Heyries et al. Nat Methods. 2011 Jul 3;8(8):649-51.
  - Statistics of digital PCR in supplementary section
- Pagno & Gauvreau “Principles of Biostatistics” 2nd ed. ISBN-10: 0534229026
  - Statistics textbook
  - Copy number variation using digital PCR
- https://www-s.nist.gov/srmors/view_detail.cfm?srm=2366
  - SRM 2366 webpage
- uCOUNT(SM) University of Utah http://www.dna.utah.edu/ucount/uc.html
  - Monte Carlo simulations of Poisson statistics

Select References (4)

- http://www.youtube.com/watch?v=s9HUhuCbbhU
  - Loading of Fluidigm 48.48 array
- http://www.youtube.com/watch?v=UwzDc6wcGZg&feature=relmfu
  - Loading of Fluidigm 12.765 digital array
- http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6237.pdf
  - Product sheet for Bio-Rad QX100
- http://www.genomeweb.com//node/1125071?hq_e=el&hq_m=1345992&hq_l=3&hq_v=88d133a09c
  - Article on continuous flow microfluidic PCR, which could give emulsion PCR the advantages of real-time data collection