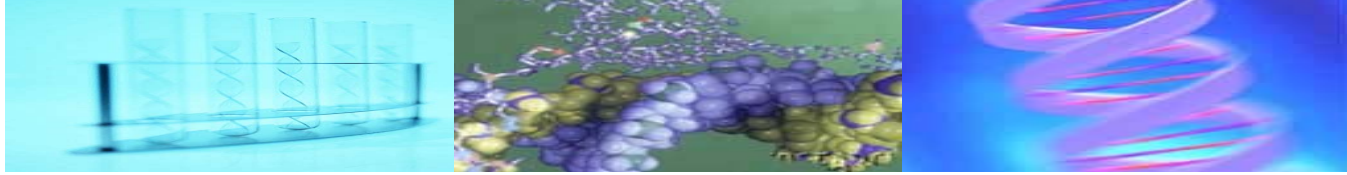
Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.



StaRT-PCR™

Use of Standardized Mixtures of Internal Standards in RT-PCR to Generate Validated Biomarkers And To Develop Standardized Transcript Abundance Reference Databases

James C. Willey
Professor of Medicine and Pathology
Medical University of Ohio
Consultant to Gene Express, Inc.

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

StaRT-PCR™ Meets FDA Guidelines
for Generating Valid Biomarkers

http://www.fda.gov/cdrh/oivd/guidance/1551.html#7

A.) Methodological Analytical Performance Characteristics	B.) *StaRT*-PCR™ Performance ▪ Validated in multiple independent studies ▪ Published in peer-review literature
1.) Reproducibility	Intra sample reproducibility CV 5-10% Day-to-day Reproducibility Monitoring <15%
2.) Lower Detection Threshold	< 10 molecules
3.) Quality Control	SMIS™ in each measurement controls for all known sources of variation
4.) Assay Specificity	Primers validated to ensure specificity of transcript
5.) Effective Assay Range	$>10^7$ ($<10^1$ to $>10^7$ molecules/10^6 β-actin molecules)
6.) Ability to identify small differences	As little as 20% differences due to high sensitivity and reproducibility

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

How *StaRT*-PCR™ Works

Willey et al, Methods in Molecular Biology, 2004;258:13-41

StaRT-PCR™ measures transcript abundance by utilizing Standardized Mixtures of Internal Standards (SMIS™) in every expression measurement

SMIS™ contains fixed molar ratios of gene-specific standards for each measurement.

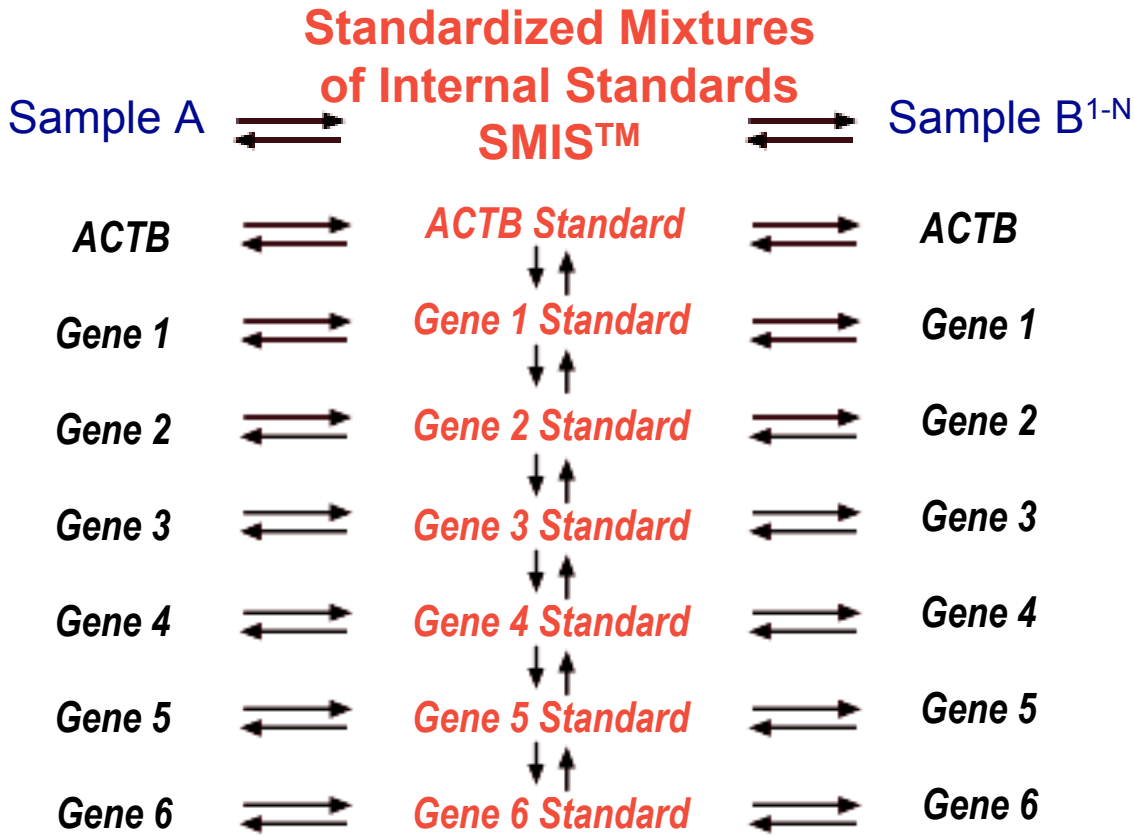
---

1. A shortened competitive template internal standard (10-20% shorter) is prepared for each gene and manufactured through a 29 step GLP process
   - 80-400 bp region bracketed by gene-specific primers
   - Primer sequences assessed for known SNP's
   - Primer-template annealing temperatures are the same for all genes
   - Each gene-specific IS amplifies with the same primer pairs as the NT gene in samples
   - Each gene-specific IS amplifies with same PCR amplification efficiency as NT in samples
2. Internal standards for 96 genes are combined into a single lot of SMIS™
3. The target gene internal standards are titrated over 7 logs relative to reference gene internal standards enabling standardized measurement across the full linear dynamic range of gene expression

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Key Proprietary Technology

Standardized Mixtures of Internal Standards (SMIS™)

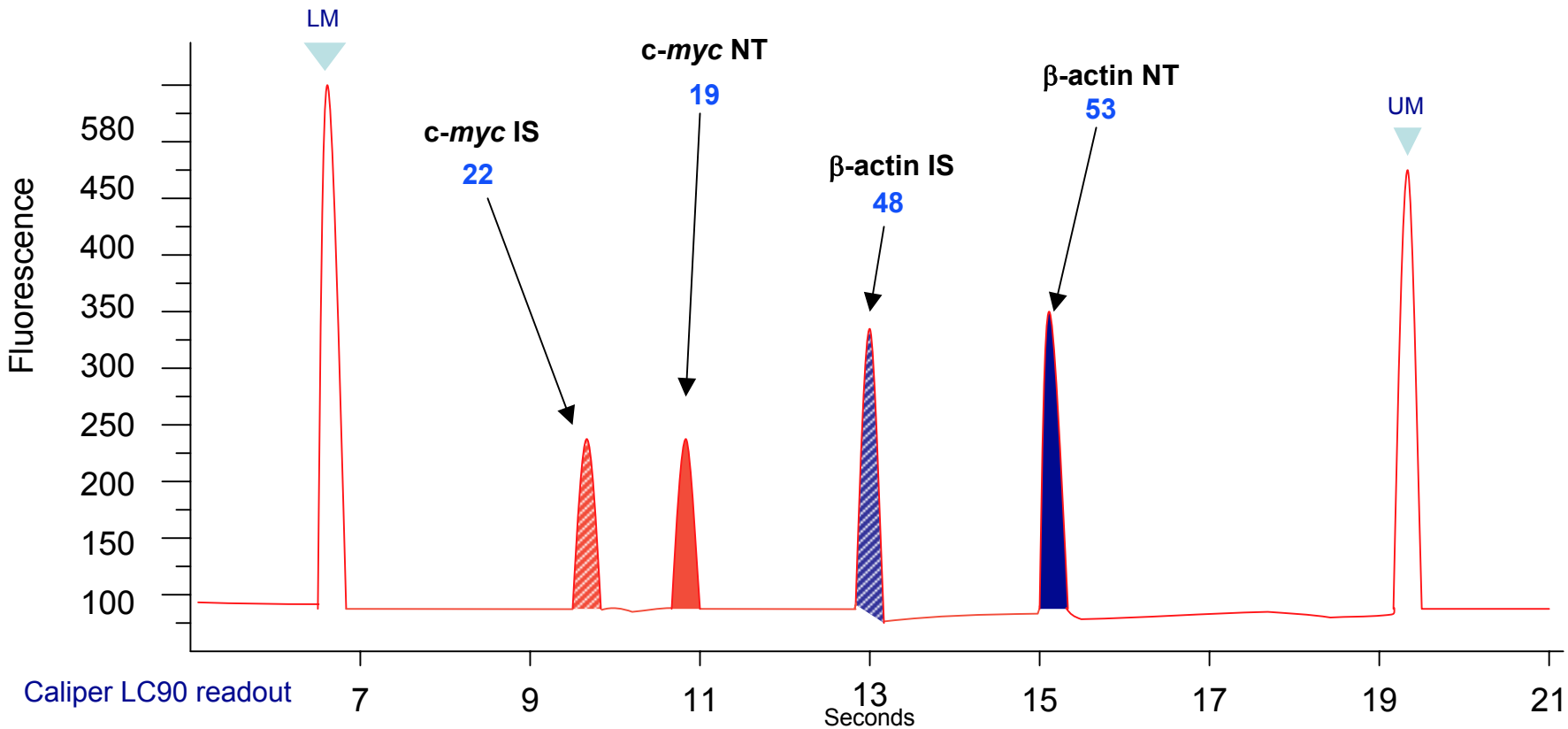


In each sample, each gene is compared to its respective internal standard within the SMIS™. All genes are subsequently directly comparable.

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Transcript Abundance Measurement

Numerical Data



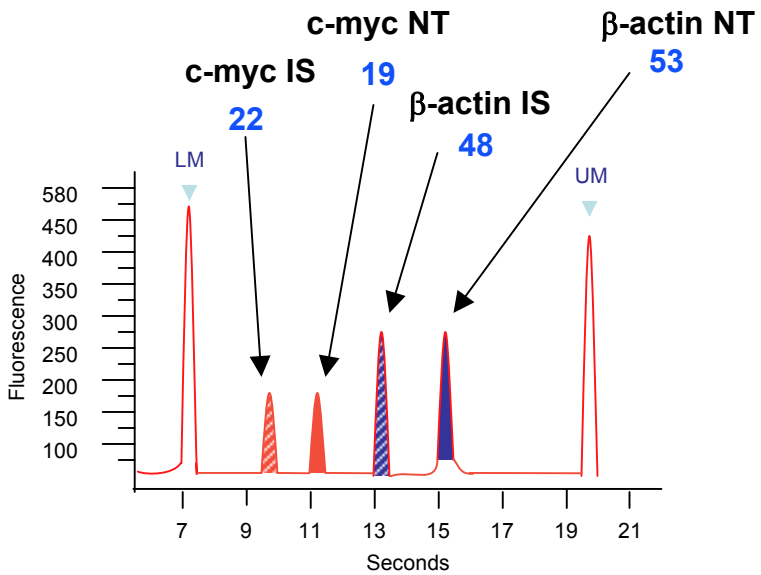
c-*myc*: 19/22 (NT/IS) X 60,000 c-myc IS molecules = 52,000 c-*myc* NT molecules

β-actin: 53/48 (NT/IS) X 600,000 β-actin IS molecules = 663,000 β-actin NT molecules

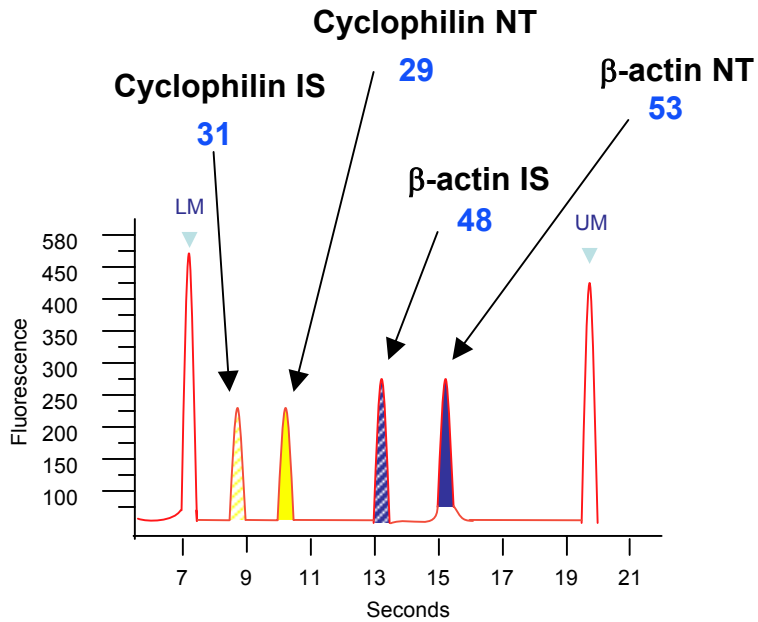
$$\left(\frac{52{,}000 \text{ c-}myc \text{ molecules}}{663{,}000\ \beta\text{-actin molecules}}\right) \times \frac{10^6 / 663{,}000\ \beta\text{-actin molecules}}{10^6 / 663{,}000\ \beta\text{-actin molecules}} = \frac{78{,}215 \text{ c-}myc \text{ molecules}}{10^6\ \beta\text{-actin molecules}}$$

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

SMIS™ Matrix
Enables Flexible Normalization



78,215 c-myc NT molecules
10^6 β-actin NT molecules



84,723 Cyclophilin NT molecules
10^6 β-actin NT molecules

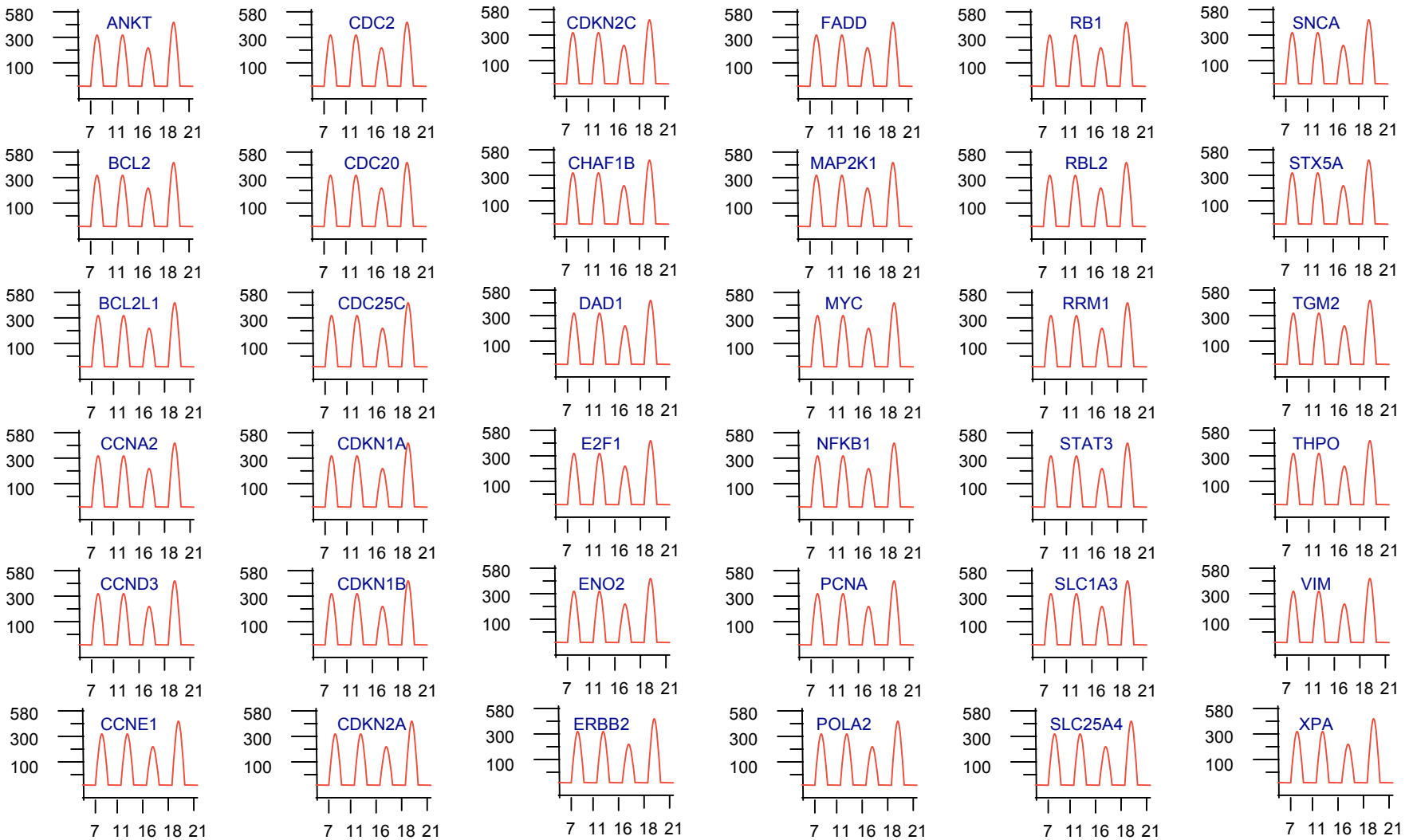
β-actin: 53/48 (NT/IS) X 600,000 IS molecules = 662,500 β-actin NT molecules

$$\frac{\text{78,215 c-myc molecules}}{10^6\ \beta\text{-actin molecules}} \times \frac{10^6\ \beta\text{-actin molecules}}{\text{84,723 Cyclophilin molecules}} = \frac{\text{923,184 c-myc molecules}}{10^6\ \text{Cyclophilin molecules}}$$

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Multi-Transcript Measurement

with *StaRT*-PCR™



Quantification of StaRT-PCR Products By Different Methods Yields Reproducible Results

- Blinded Multi-laboratory Study Using Different Electrophoresis Instruments Yields Reproducible Results (Crawford et al, Mol. Diagnosis, 2001)
- Electrophoresis of StaRT-PCR™ Products By Different Instruments yields reproducible results (Crawford et al, Agilent Application Note, 2001)
  - Slab gel electrophoresis
  - Agilent 2100
  - Perkin Elmer 310
- Analysis of StaRT-PCR™ Products By Microarray or Electrophoresis Yields Reproducible Results (Crawford et al, Proceedings of AACR, 2004)

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Non-Standardized Data

Can Be Problematic & Costly to "Fix"

Quantitative Real Time RT-PCR without SMIS™…

- Problem 1: Variation in Lower Detection Threshold for Each Gene
  - Because each cDNA sample has unknown copy number for each gene, the use of cDNA titration in real-time analysis as method to assess primer efficiency does not enable measurement of lower detection threshold (LDT) or transcript copy number
- Problem 2: Inter-sample variation in presence of PCR inhibitors is well-known (e.g. Heme)
  - PCR inhibitors are often gene-specific

The only way to control for sample-to-sample variation in gene-specific inhibitors is with use of SMIS™ in each measurement

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Reverse Transcription
Oligo dT Priming RT

- Primer Efficiency
  - Oligo dT and/or Sequence Specific Priming more efficient than random priming
- Linearity of RT with Oligo dT
  - Linear over wide range of expression
- MMLV and Superscript yield the same results
- Reproducibility of RT with oligo dT Priming
  - Multiple RT of same RNA yields highly reproducible results even with significantly degraded RNA samples (due to contribution of spontaneous priming?)
- Utility of RT with oligo dT Priming
  - Nearly all transcripts reverse transcribed in one RT reaction
- Limits
  - A small number of transcripts do not have poly A tail and may not be reverse transcribed as efficiently

Support for Inclusion of an Internal Standard in Each Gene Expression Measurement

Roche Molecular Biochemicals
Technical Note No. LC 12/2000

Roche

LightCycler

Absolute Quantification with External Standards and an Internal Control

Introduction

Requirements for an Internal Control

3. Guidelines for Designing an Internal Control

The Hybridization Probe format allows the LightCycler system to amplify and identify two target sequences in one capillary. This determination requires the use of two sets of Hybridization Probes that carry different labels (LC Red 640 and LC Red 705). The signals for the two amplicons are detected separately in fluorimeter channel 2 and channel 3, respectively.
One of these sequences can serve as an internal control that is co-amplified with the target sequence. Since factors that inhibit PCR will affect not only the target, but also the internal control fragment, the internal control:

- Can discriminate between a truly negative result and a false negative caused by inhibition of the PCR
- Serves as an indicator of PCR performance and therefore ensures accurate, reliable quantification

In this section you will find recommendations for designing a suitable internal control.

For best results, the internal control should meet all the criteria listed in the following table.

3. Guidelines for Designing an Internal Control

Introduction

The Hybridization Probe format allows the LightCycler system to amplify and identify two target sequences in one capillary. This determination requires the use of two sets of Hybridization Probes that carry different labels (LC Red 640 and LC Red 705). The signals for the two amplicons are detected separately in fluorimeter channel 2 and channel 3, respectively.
One of these sequences can serve as an internal control that is co-amplified with the target sequence. Since factors that inhibit PCR will affect not only the target, but also the internal control fragment, the internal control:

- Can discriminate between a truly negative result and a false negative caused by inhibition of the PCR
- Serves as an indicator of PCR performance and therefore ensures accurate, reliable quantification

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

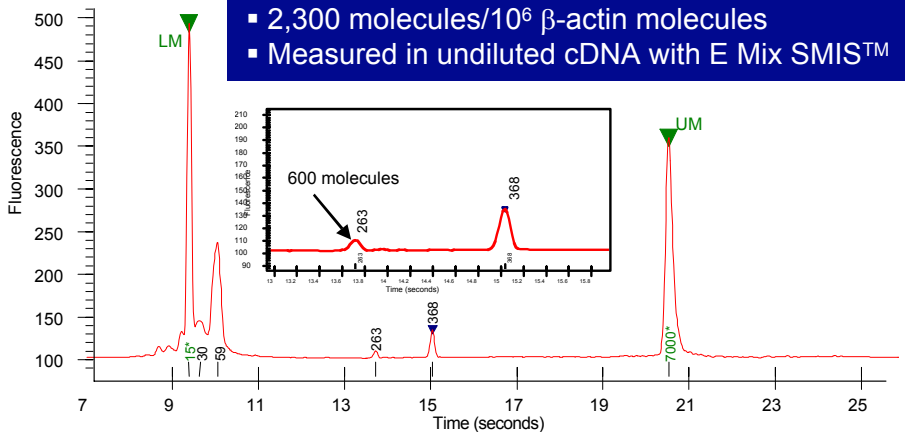
Client Project 1: Integrated Quality-Control

SMIS™ Controls for Gene-Specific Inhibition of PCR Amplification

- Gene-Specific PCR Inhibitor in Sample[339]
  - Has little affect on amplification of lowly expressed gene - GSTM3
  - Prevents amplification of highly expressed gene - ERBB2

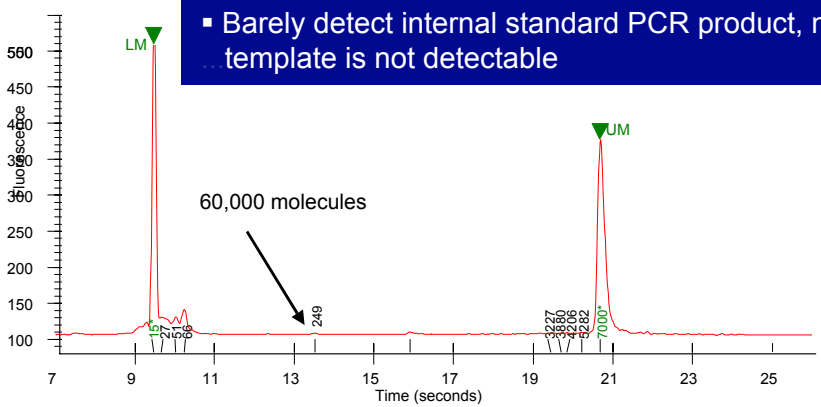
GSTM3

- 2,300 molecules/10^6 β-actin molecules
- Measured in undiluted cDNA with E Mix SMIS™



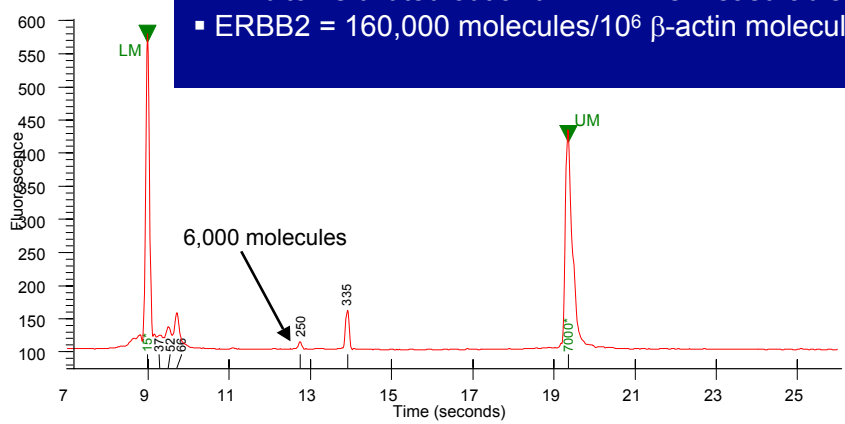
ERBB2

- Not measurable, due to PCR inhibitor
- Undiluted cDNA with C Mix
- Barely detect internal standard PCR product, native template is not detectable



ERBB2

- Dilute sample 10-fold and SMIS™ 10-fold
- Inhibitor is diluted out and ERBB2 is measurable
- ERBB2 = 160,000 molecules/10^6 β-actin molecules



Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Analyzing StaRT-PCR Data

StaRT-PCR Products May Be Analyzed on Any Platform

- CaliperLC90 Microfluidic Electrophoresis of StaRT-PCR™ Products
  - High throughput
  - Lower detection threshold of less than 10 PCR template molecules
  - Signal-to-analyte response 100%
  - Yields same results as on Agilent 2100. Studies with Agilent 5100 in progress
- MALDI-TOF Quantification of PCR Products
  - Excellent Lower detection threshold
  - Automated systems available (e.g. Sequenom)
- Microarray Analysis of StaRT-PCR™ Products
  - Design microarray with probe that recognizes both internal standard and native template StaRT-PCR product
  - Use 2-color fluor to distinguish
  - Validated in small studies to yield same result as electrophoretic analysis

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Client Project 2: StaRT-PCR™
Reproducibility Study

- **To determine day to day and run to run variability at SEM Center™:**
  - Three aliquots of the Stratagene Universal Human RNA were separately reverse transcribed
  - Nine genes selected at random and evaluated by *StaRT*-PCR™ in each samples in triplicate on three consecutive days.
  - Gene expression values were calculated and the variability assessed
  - For higher expressed genes CV <3% (FDA MAQC Study)

Samples 512, 513, 514 for days 1-3
Sorted on overall % CV

Gene (Standard Nomenclature)	Average Data Point CV %	Run to Run Average CV % (Days 1-3)	RT to RT Average CV % Samples (512-513)	Overall Average CV %
TMSB4X	3.3	4.9	8.5	8.6
BCL2L1	4.6	6.5	9.2	10.1
CALR	4.5	7.6	8.4	11.0
KRT18	2.9	10.9	7.5	11.5
CAT	8.6	5.4	11.8	13.5
RPS26	7.0	12.5	11.5	14.7
SSBP1	4.8	12.9	11.0	15.4
MGST1	7.9	23.4	7.2	22.1
ADH5	6.6	17.1	21.0	24.1
All Genes	**5.6**	**11.2**	**10.7**	**14.6**

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Client Project 3 Design

Dilution of Lung cDNA with Spleen cDNA

Samples

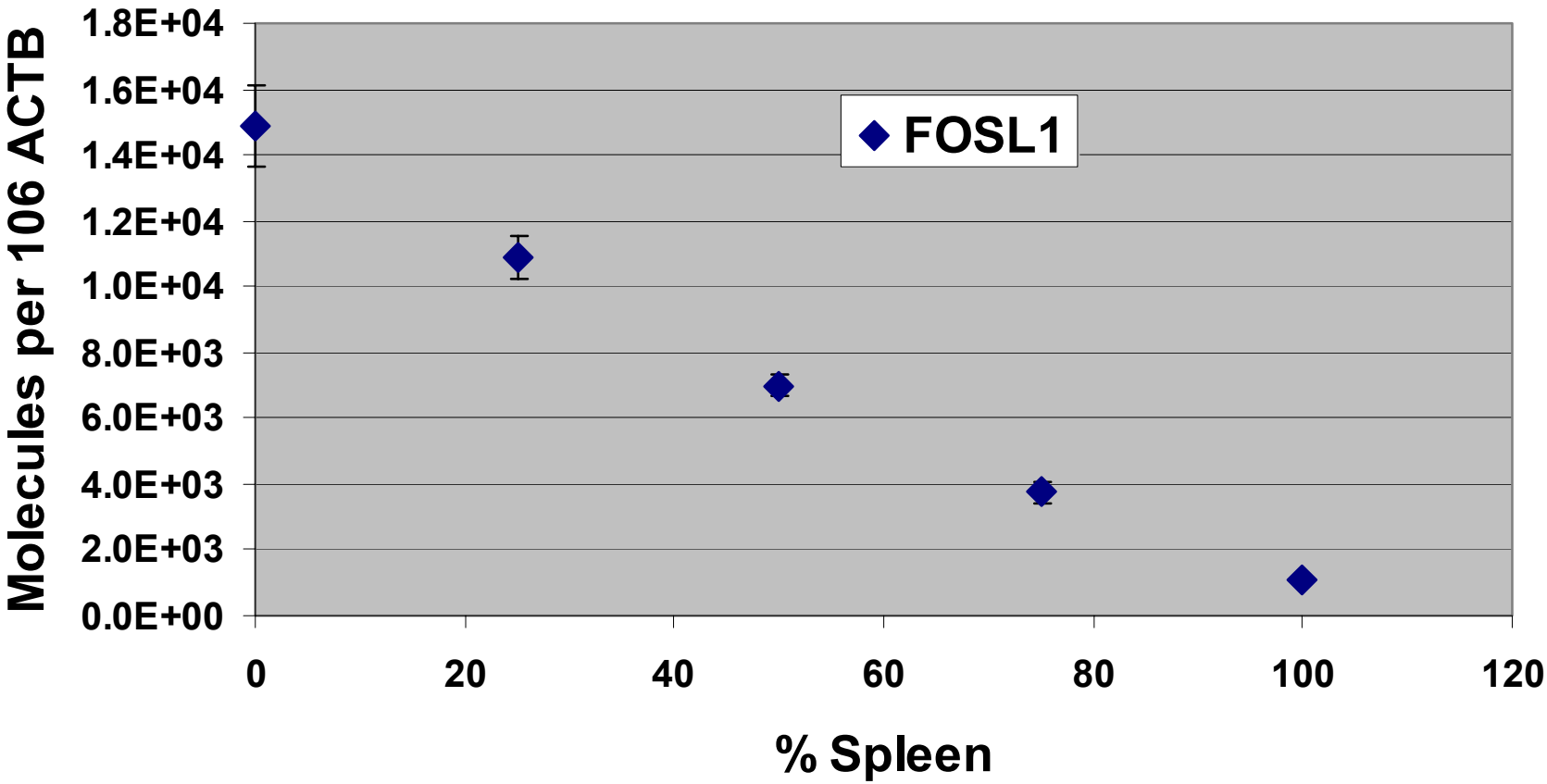
- #1 Lung
- #2 75% Lung, 25% Spleen
- #3 50% Lung, 50% Spleen
- #4 25% Lung, 75% Spleen
- #5 Spleen
- #6 Muscle

Purpose: Determine ability of StaRT-PCR to discriminate small differences in samples

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Client Project 3

Dilution of Lung cDNA with Spleen cDNA



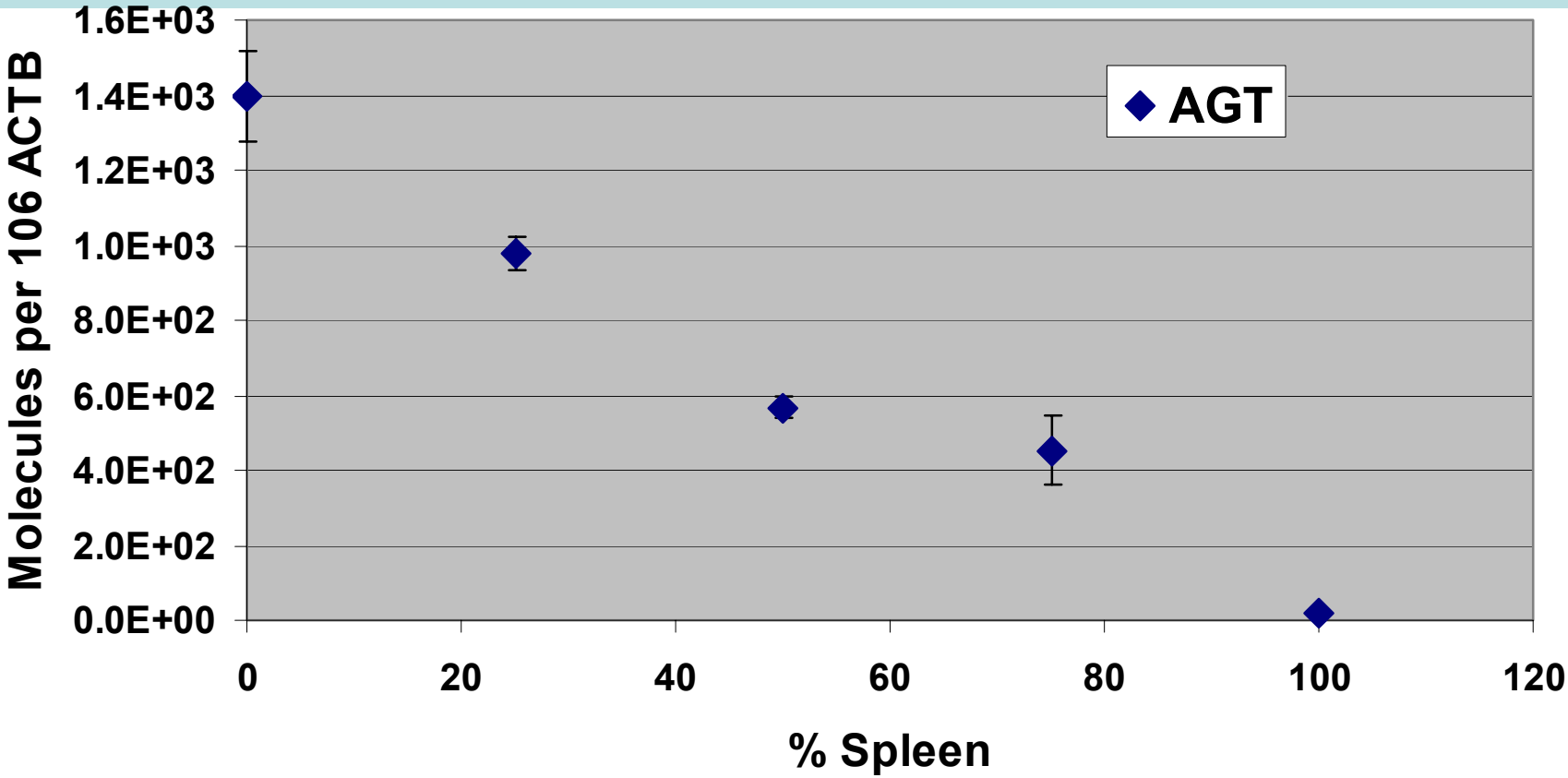
AFFY DATA

	Lung	P value	Spleen	P value
204420_at	232.7	0.002251	21.9	0.482865

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Client Project 3

Dilution of Lung cDNA with Spleen cDNA

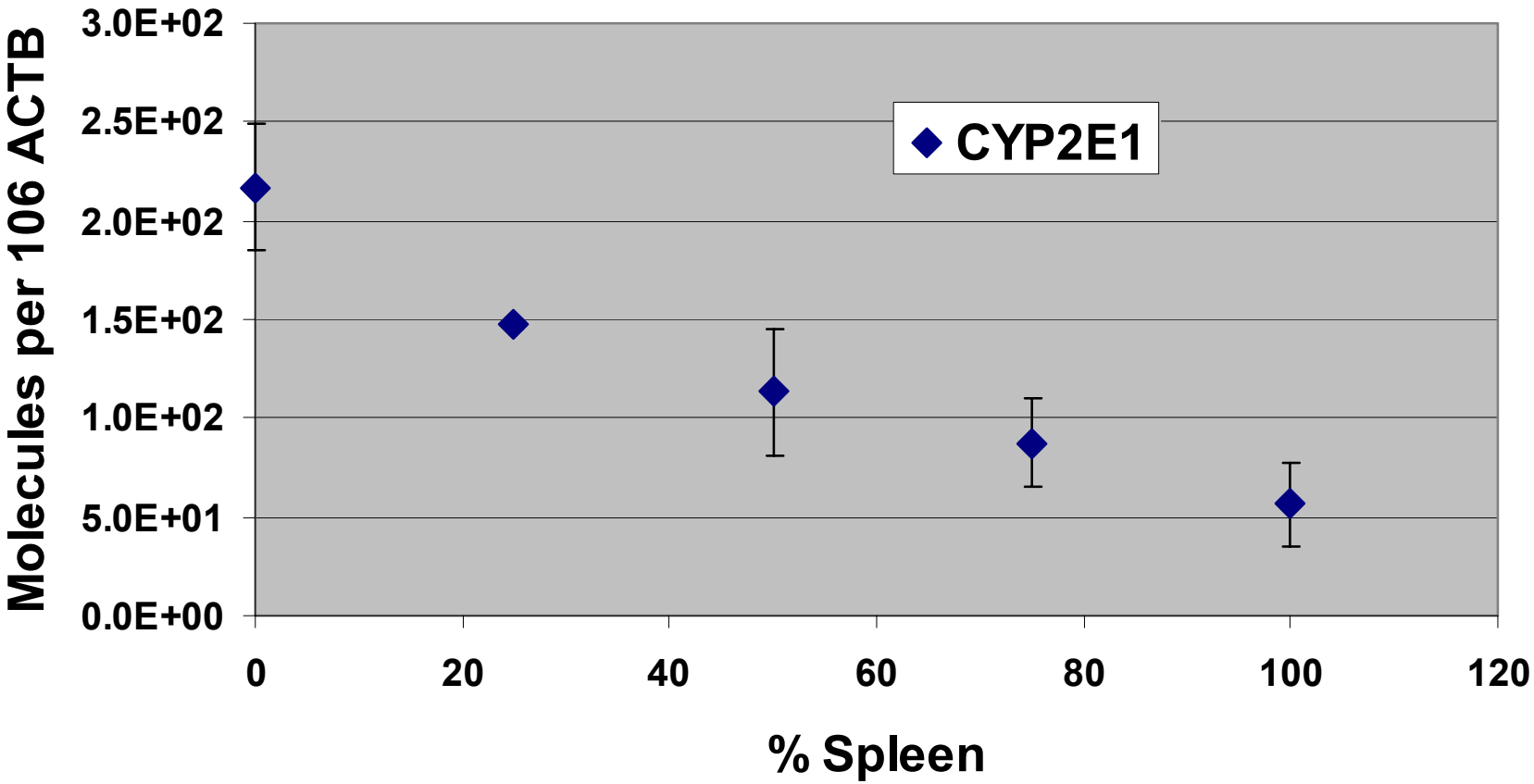


	AFFY Data			
	Lung	**P Value**	**Spleen**	**P Value**
202834_at	441.4	0.000704	31.9	0.274591

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Client Project 3

Dilution of Lung cDNA with Spleen cDNA



AFFY Data

	Lung	P value	Spleen	P value
209975_at	69.4	0.015789	38.5	0.222083

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

FDA MAQC Project Consistent with Client 2 and 3 Results

FDA MAQC Samples

- Sample A = 100% Stratagene Universal Human Reference RNA (SUHRR)
- Sample B = 100% Ambion Brain RNA Brain
- Sample C = (0.75 SUHRR + 0.25 Brain)
- Sample D = (0.25 SUHRR + 0.75 Brain)
- FDA MAQC Project Confirms Client 2 and 3 Project Results
  - High linearity: Median R2 > 0.95
  - High Precision: Average CV < 6%
  - Consistency with other Transcript abundance measurement methods
  - 100% Signal-to-Analyte response
  - Manuscripts in preparation

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Validation of a New Clinical Assessment Method: Collaboration with Pfizer

- Is it measuring what it is supposed to measure?
- Is the measurement sufficiently repeatable?
- Utility
  - Does a change truly reflect improvement/worsening?
  - Does the test distinguish the phenotypes of interest with sufficient sensitivity, specificity, and accuracy?
- Specific Questions:
  - Is whole blood a suitable sample for transcript abundance (TA) measurement by *StaRT*-PCR™?
  - Can sources of variation in TA measurement of the selected genes in blood samples be identified and quantified?

Pfizer

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Study Results – Subject Characteristics

- 15 healthy volunteers
- Gender
  - 8 males (53%)
  - 7 females (47%)
- Mean Age (range): 28 yrs (19 - 47)
- Race
  - 7 White 47%
  - 4 Hispanic/Latino 27%
  - 3 Black/African American 20%
  - 1 Asian 7%

Pfizer

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

StaRT-PCR™ Data for each Subject

- 18 replicate expression evaluations (transcript abundance)
  - 3 assay replicates x 3 samples x 2 visits
- for each of 19 genes (*name of genes blinded*)
- with 2 normalizer genes for standardization – ACTB, GAPD

~10,000 total TA data points

Pfizer

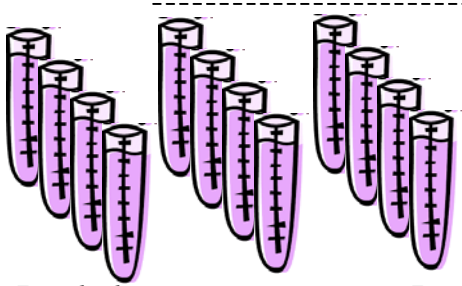
Study Design

15 healthy volunteers enrolled

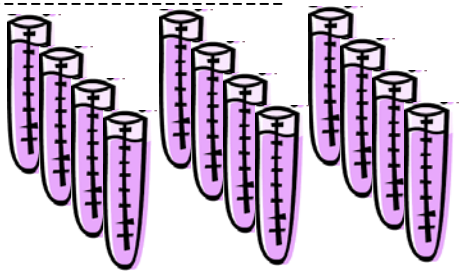
Visit 1 — *3-7 days* — Visit 2

Medical History
Physical Exam
ECG
Safety labs

Safety labs



Pooled Sample A *Pooled Sample B* *Pooled Sample C*



Pooled Sample A *Pooled Sample B* *Pooled Sample C*

***StaRT*-PCR™ Assay – Gene 1**

- Replicate 1
- Replicate 2
- Replicate 3

Sources of Variability

- within subject
  - between 3 replicate assays
  - between 3 pooled samples
  - between 2 visits
- between subjects

Cannot evaluate directly:

- SMIS™ mix
- technician
- equipment
- laboratory
- clinic

⬇

confounded with other factors or included as random error

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Statistical Representation

Observed TA = True TA
+ subject variation
+ visit variation
+ sample variation
+ replicate variation
+ random error (residual)

Biological Variation	Analytical Variation
✓	
✓	✓
✓	✓
	✓

Potential Components

- **different types of cells**
- **environmental effects**
- **temporal effects**

- ***StaRT*-PCR™ methodology**
- **RNA extraction**
- **RNA handling**
- **Reverse Transcription**

Pfizer

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Methodology for Evaluating Variability

- Estimate the variance components by finding the "best fit" of the preceding model to the actual study data
  - looks at the variability around each source:
    - replicate: how much do the individual assay replicates differ from mean of the replicates
    - sample: how much do the assay means for each of three samples differ from the sample mean
    - etc
  - calculate the percent of total variation for each component
- CV – coefficient of variation = standard deviation (SD) divided by mean expressed as a percentage

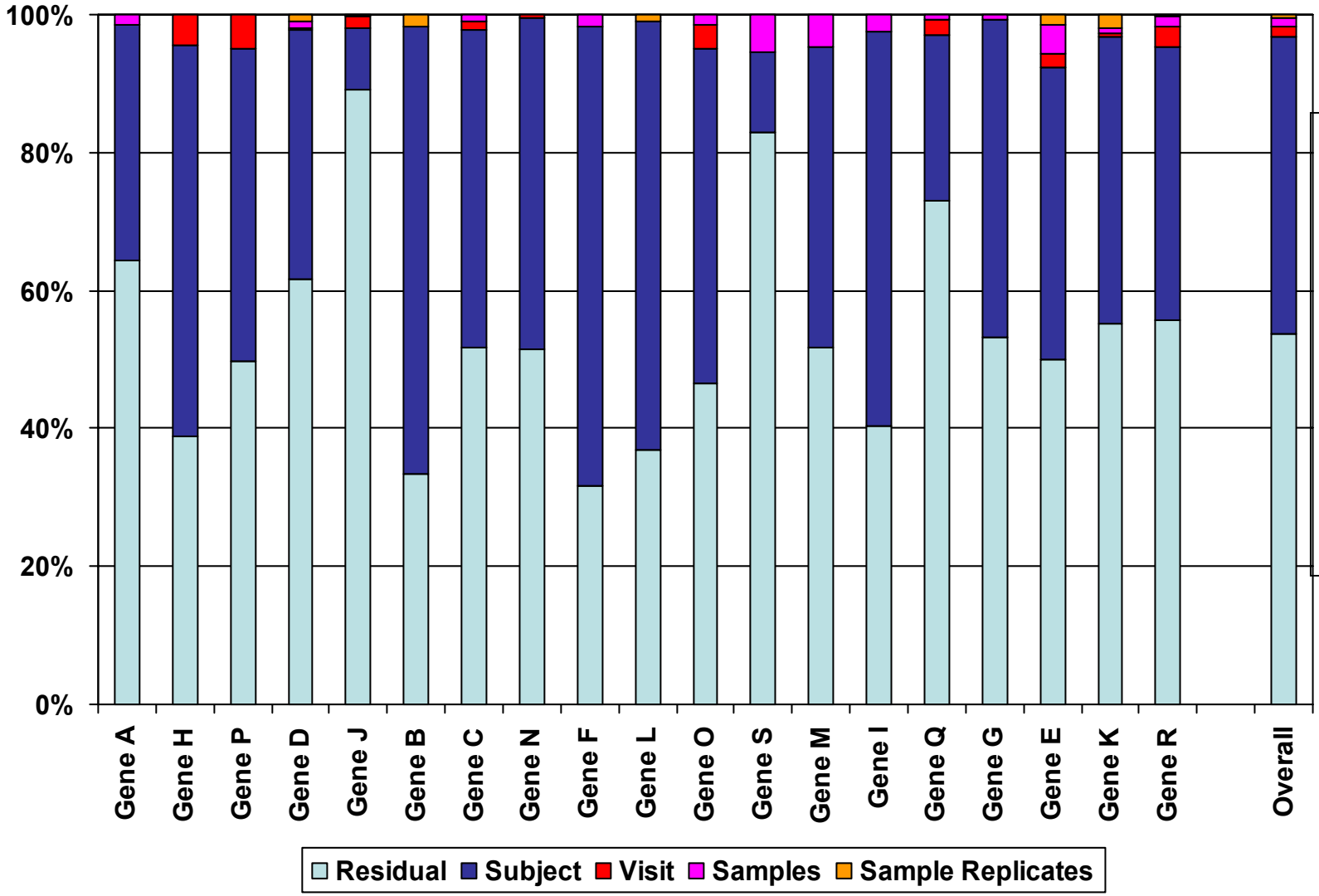
Pfizer

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Percentage of Total Variability
Attributed to each Component

Total %CV (includes all biological + analytical + random variation):

Gene A	Gene H	Gene P	Gene D	Gene J	Gene B	Gene C	Gene N	Gene F	Gene L	Gene O	Gene S	Gene M	Gene I	Gene Q	Gene G	Gene E	Gene K	Gene R	Overall
31	39	98	55	42	86	33	57	53	80	124	10	51	44	56	33	31	32	68	Ave: 54%



Overall Estimates

Assay replicates: negligible

Pooled Samples: 1% (across & within visits)

Visits: 1%

Subjects: 43%

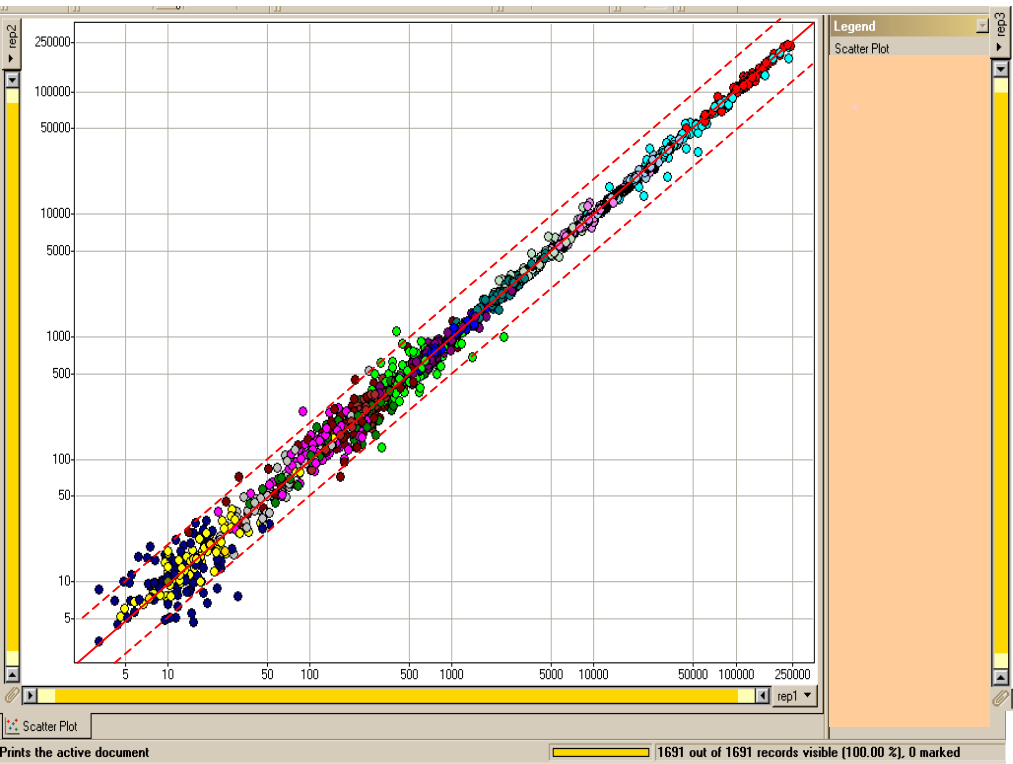
Residual: 54%

Pfizer

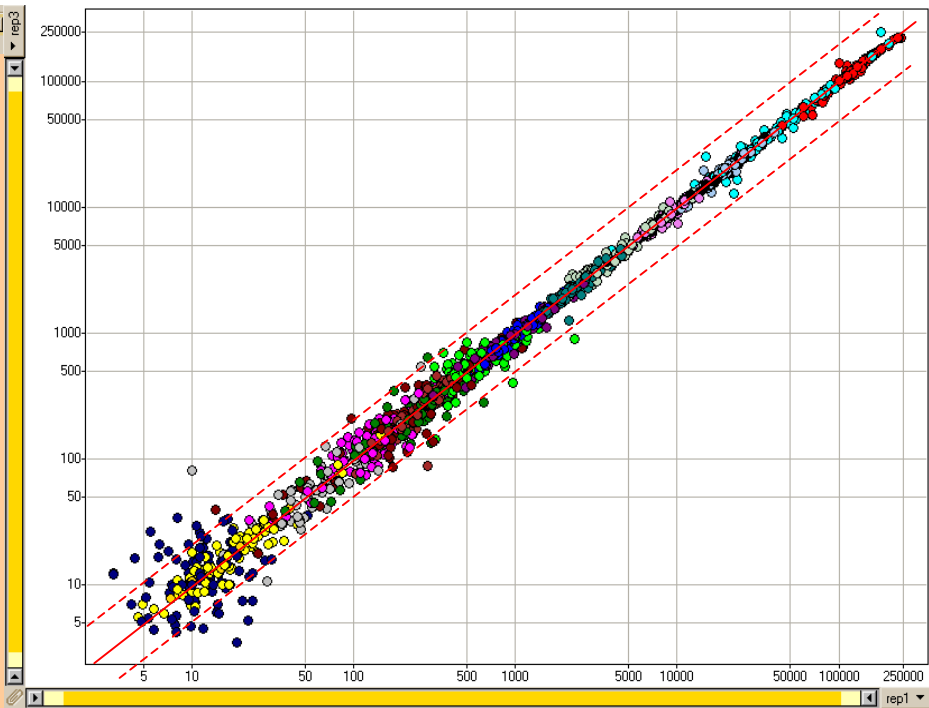
Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Comparison of *StaRT*-PCR™ Replicate Measurements of a Single Sample - ACTB

Replicate 1 vs. Replicate 2



Replicate 1 vs. Replicate 3

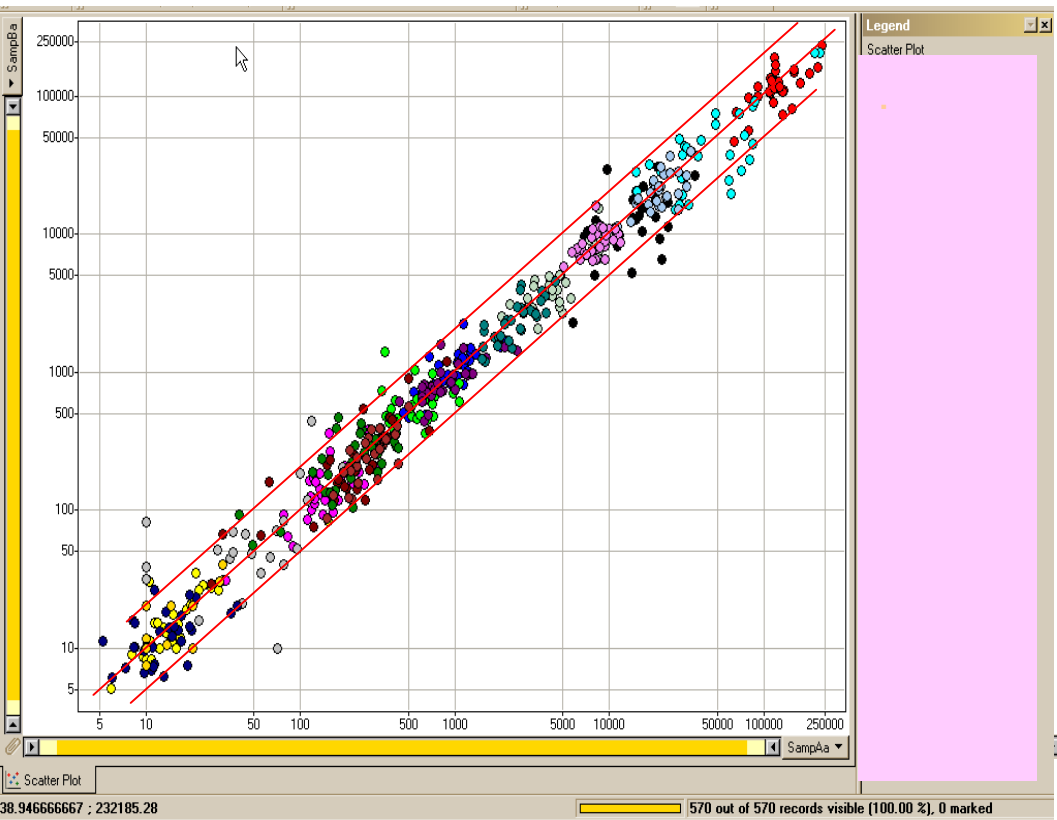


Pfizer

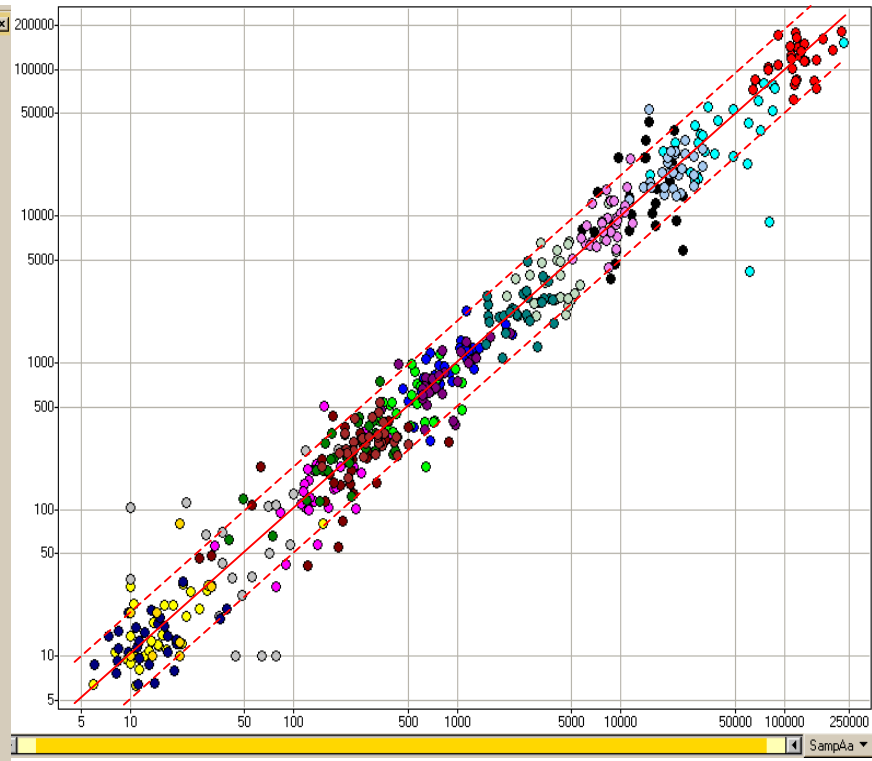
Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Comparison of Samples - ACTB

Sample A vs. Sample B



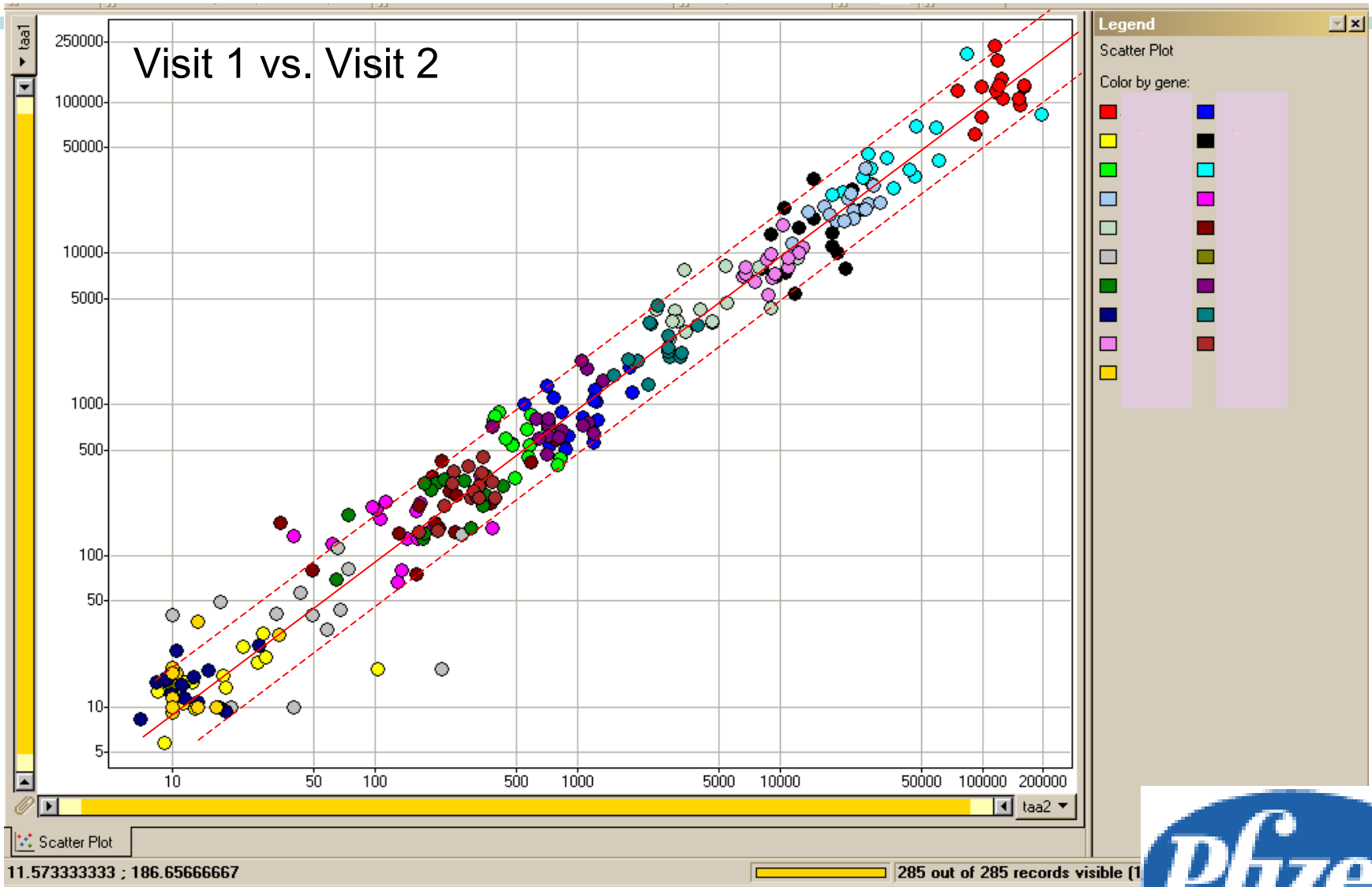
Sample A vs. Sample C



Pfizer

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Comparison of Visits



Pfizer

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Normal Ranges of Transcript Abundance (TA) Levels Across Subjects

Target gene cDNA molecules / 10^6 reference gene cDNA molecules

ACTB Gene	#Obs	Mean	SD	1SD Range Lower	1SD Range Upper	%Obs in Range	2SD Range Lower	2SD Range Upper	%Obs in Range	3SD Range Lower	3SD Range Upper	%Obs in Range
Gene A	30	123472	33871	89602	157343	76.7%	55731	191214	96.7%	21860	225085	96.7%
Gene H	30	1024	362	662	1386	73.3%	299	1749	90.0%	-63	2111	100.0%
Gene P	30	19	17	2	36	96.7%	0	54	96.7%	-33	71	96.7%
Gene D	30	14593	6402	8191	20994	63.3%	1789	27396	96.7%	-4613	33798	100.0%
Gene J	30	594	163	431	757	60.0%	268	920	100.0%	104	1083	100.0%
Gene B	30	51890	44651	7239	96542	93.3%	0	141193	93.3%	-82063	185844	93.3%
Gene C	30	21422	5695	15728	27117	73.3%	10033	32812	96.7%	4338	38507	100.0%
Gene N	30	163	80	83	243	76.7%	3	323	90.0%	-77	402	100.0%
Gene F	30	4998	2444	2554	7442	73.3%	109	9886	96.7%	-2335	12330	100.0%
Gene L	30	250	170	80	419	73.3%	0	589	93.3%	-259	758	96.7%
Gene O	30	56	60	-3	116	90.0%	0	176	93.3%	-123	235	96.7%
Gene S	30	<10	0									
Gene M	30	243	95	147	338	63.3%	52	433	100.0%	-44	529	100.0%
Gene I	30	879	361	518	1240	80.0%	157	1601	93.3%	-203	1962	100.0%
Gene Q	30	14	5	9	19	80.0%	4	24	90.0%	-1	29	100.0%
Gene G	30	2516	751	1765	3266	70.0%	1015	4017	96.7%	264	4767	100.0%
Gene E	30	8934	2203	6732	11137	80.0%	4529	13339	96.7%	2326	15542	100.0%
Gene K	30	287	75	212	362	73.3%	137	436	96.7%	62	511	100.0%
Gene R	30	14	7	7	21	90.0%	0	28	90.0%	-8	36	96.7%

Gene Express, Inc.
Setting the Standard in Genomics.™
Discovery. Molecular Diagnostics. Therapeutics.

Internal Standards
RNA Standards or cDNA Standards?

Why it is not necessary to use RNA Standards

- Most RNA degradation occurs prior to and/or during collection (i.e. biopsy)
  - RNA internal standards added after collection won't control for degradation of RNA that takes place prior to and/or during collection.
- RNA degradation during extraction, storage, or RT is minimal.
- There is no variation in representation of one transcript to another among multiple RTs of the same sample, even though the overall RT efficiency varies considerably (10-80%).
- One can assess for variation during RT by mixing samples with wide range in expression and determining whether the result is linear. We have observed no evidence for gene-specific RT inhibition.