Quantification Standards for 5' Nuclease Gene Expression Assays

Manohar Furtado Ph.D.
Applied Biosystems

NIST Workshop on Universal RNA Standards
March 28-29 2003

Applied Biosystems

5’ Nuclease Assay

1. Real time (Kinetic) PCR assay.
   - Closed tube format
   - Can be multiplexed

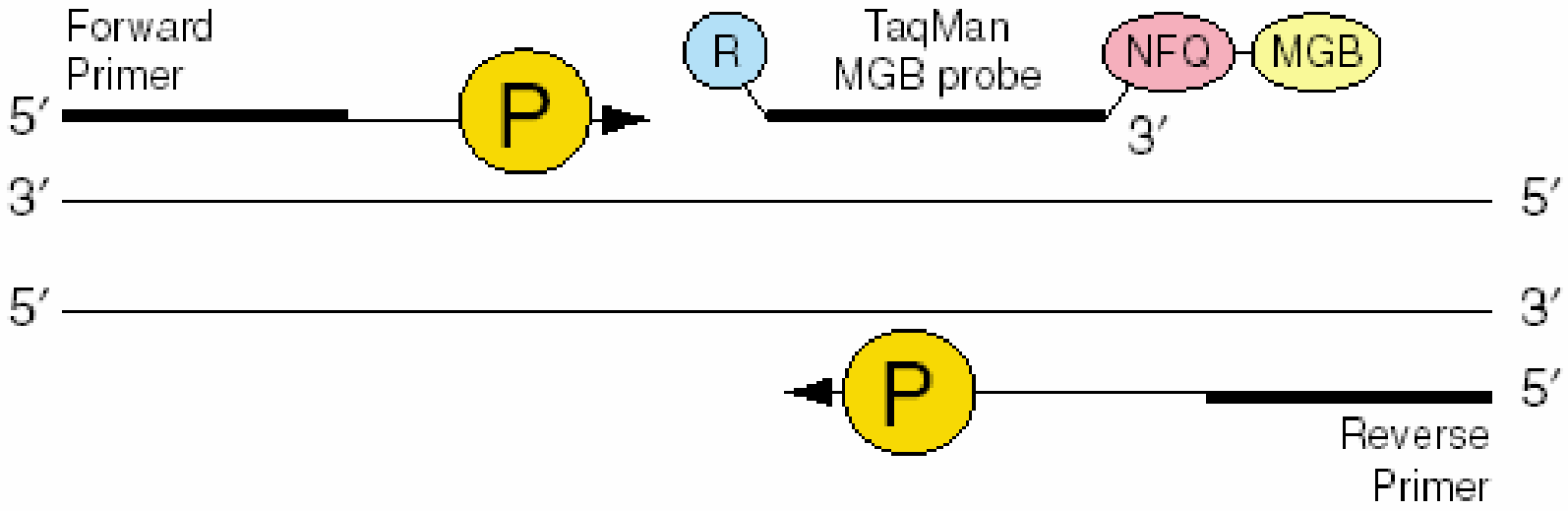
2. Contamination control is possible.

3. High degree of specificity from two primers and a probe.

4. Good sensitivity and broad dynamic range possible.

5. High efficiency (near 100%) associated with good design. Useful for HT relative quantification.

5’ Nuclease Assay



1. Reference for Expression Testing
2. Standards for Quantification.
3. Requirements for assay specificity.

Applied Biosystems

Expression Testing in Tissue Pools

Stratagene Universal Reference RNA (URR) Pool

Adenocarcinoma (mammary gland), Hepatoblastoma (liver), Adenocarcinoma (cervix),Embryonic carcinoma (testis), Glioblastoma (brain), Melanoma, Liposarcoma, Histocytic lymphoma (macrophage), Lymphoblastic leukemia, (T-cell), Plasmacytoma (B –cell)

Pool 1: testis, liver, kidney RNA (Clontech)

Pool 4: brain, lung & placenta RNA (Clontech)

Testing Gene Expression Assays-on-Demand™ Products

Assays for Human, Mouse & Rat genes

Human Gene Expression assays Ct < 35

First 2500 assays	URR pool	92%
	Pool 1	88%
	Pool 4	93%
	Combined	98%

Total of ~ 20,000 assays: URR pool ~ 70%

30% Tested with pool 1 & 4 ~ 23%

Assays-by-Design℠ Service: Any species / any target

Quantification Standards

Analytical quantification reference standards are a known / measured amount of target nucleic acid used to verify that the analytical process is functioning as expected. (e.g. WHO standards for HIV, HCV & HBV).

Quantification standards (QS) may also be used as internal or external quantitative calibrators in an assay. The signal from the calibrator is compared against signals from target for purposes of determining the amount of target. (from the manufacturer)

Based on NCCLS Guidelines

Quantification Standards

Non synthetic standards

Bacteria or virus grown in culture and titered

Cellular RNA (reference RNA pools)

DNA from well characterized cell lines.

(Quantified by culture, EM , biochemical methods).

Synthetic standards

Synthetic RNA and DNA templates.

DNA cloned in plasmids.

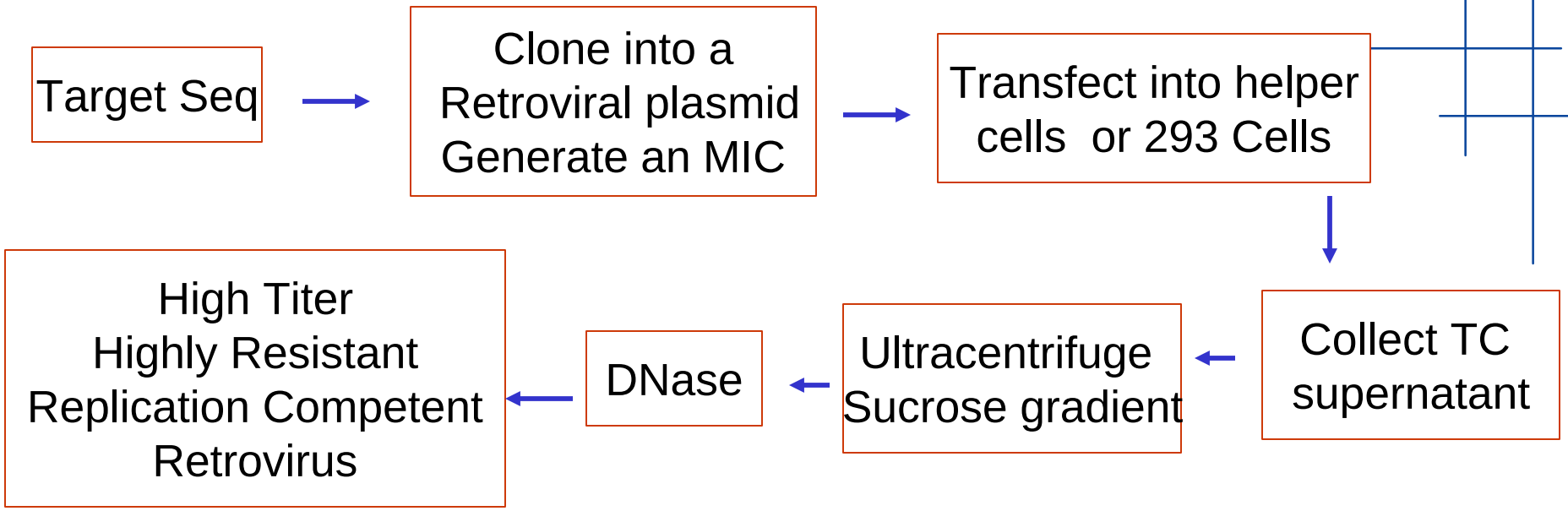
RNA generated by *in vitro* transcription.

Recombinant retrovirus with target sequence.

Recombinant phage with target sequence.

The phage and viral based standards can be spiked into the sample at the sample preparation stage and carried through the entire process.

Retrovirus Based Standards: HIV ViroSeq™ Kit



1. MICs and mutant virus with drug resistance mutations.
2. Sequenced 420 codons.
3. Reported 75 mutations on an Anti-retroviral Drug Resistance report.
4. Validated 32 mutations for detection at specific WT : Mutant ratios.
5. Each validated codon treated as an analyte.
6. Need for an Interpretive algorithm and specification of contraindicated drugs.

Applied Biosystems

Absolute Quantification: IQC Method

A quantitative calibrator is amplified along with target sequence in the same tube.

The IQC may be:

- The same size as the target and distinguished by altered probe binding sites.
- Contain a small insertion or deletion or a restriction enzyme cleavage site that helps differential detection.

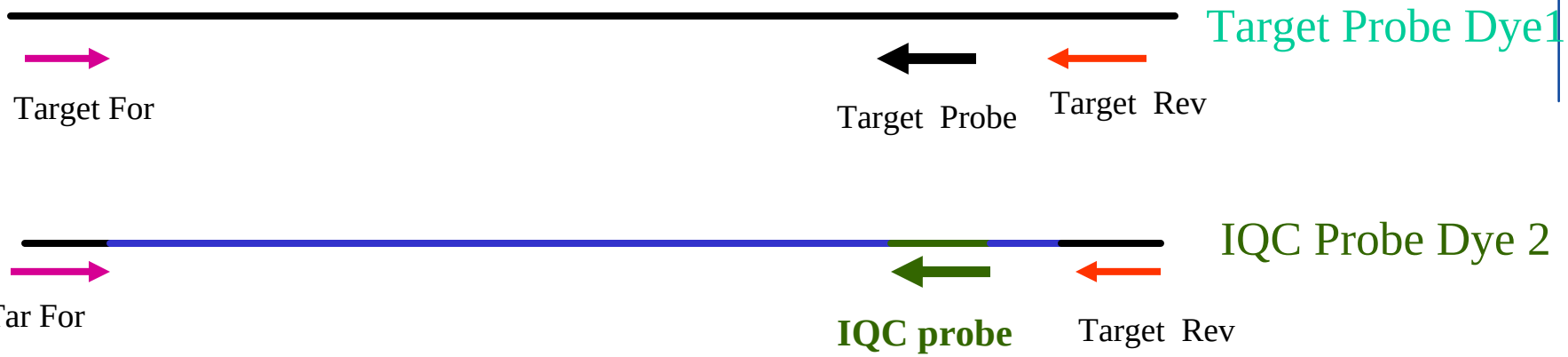
Allows detection of inhibitors.

Amplification efficiency of the target and IQC should be similar.

Reportable dynamic range within which the calibrator does not compete with the target should be defined.

Co-amplification of a Competing IQC Target Identical in Size to the Target

Efficiency Tar Amp ~ 100 % Efficiency IQC Amp ~ 100%



Introduction of 200- 500 copies/ml of IQC does not inhibit target amplification in the 50 to 5 x 10^6 copies/mL range. Needs extensive optimization

Co amplifying targets the same size as the Target amplicon.

1. Furtado M R;Murphy R; Wolinsky S M. Quantification HIV-1 tat mRNA as a marker for assessing the efficacy of antiretroviral therapy. Journal of Infectious Diseases. 167(1): 213-6, **1999.**

2. Mulder J, McKinney N, Christopherson C, Sninsky J, Greenfield L, Kwok S. Rapid and simple PCR assay for quantitation of human immunodeficiency virus type 1 RNA in plasma: application to acute retroviral infection. *J Clin Microbiol*. 1994;32:292-300. BASIS FOR THE ROCHE AMPLICOR ASSAY

Co amplifying targets smaller in size compared to the Target amplicon as IQCs

3. Furtado, MR, DS Callaway, JP Phair, KJ Kunstman, JL Stanton, CA Macken, AS Perelson and SM Wolinsky. Persistence of HIV-1 Transcription in Peripheral Blood Mononuclear Cells of Patients Receiving Potent Antiretroviral Therapy [see editorial review p1672]. New England Journal of Medicine. 340(21): 1614-1622, **1999.**

4. Wang AM, Doyle MV, Mark DF. Quantitation of mRNA by the polymerase chain reaction. *Proc Natl Acad Sci U S A*. 1989;86:9717-21 . CYTOKINES

Absolute Quantification: Standard Curve Method

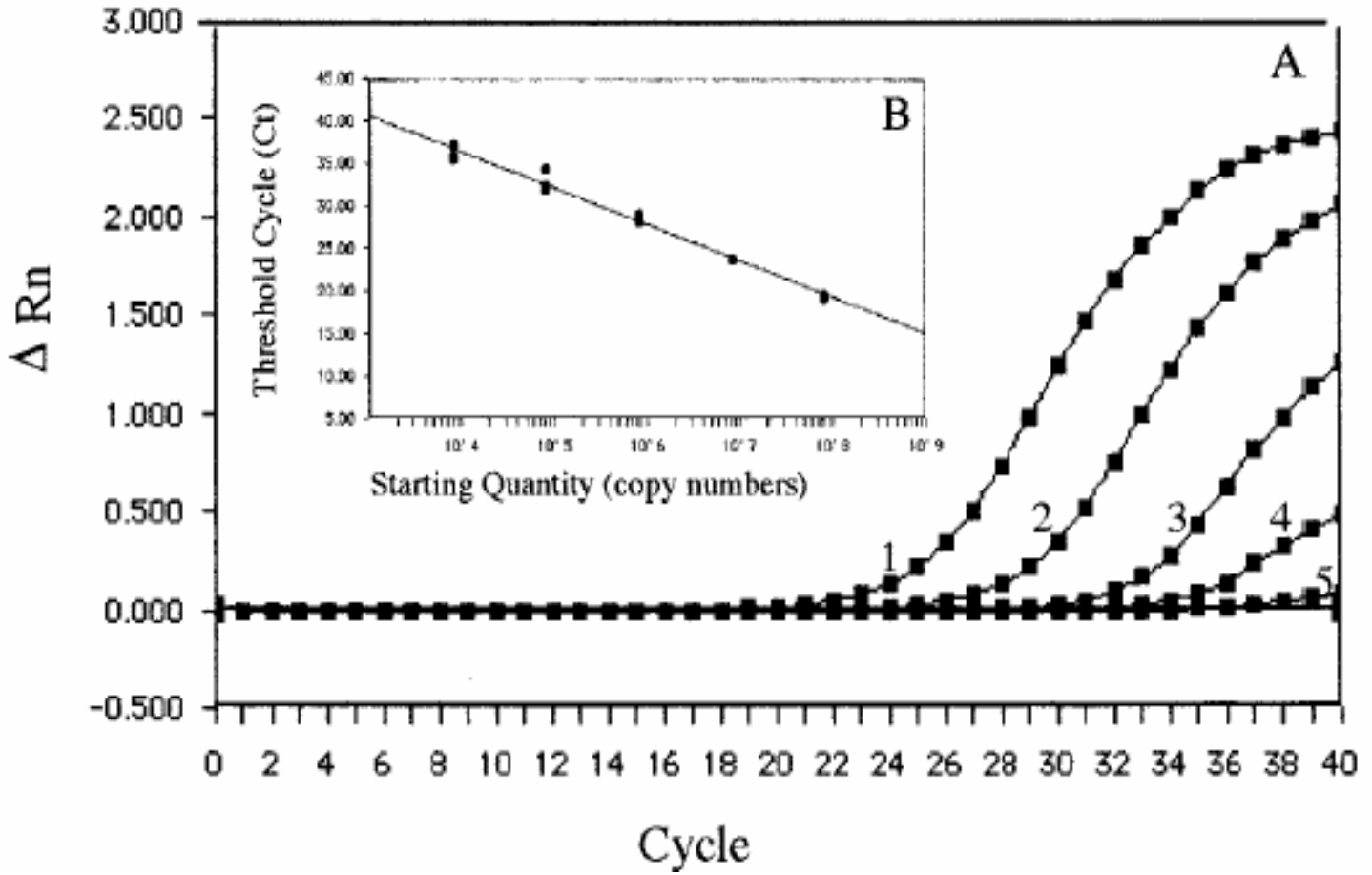
Requires the generation of a standard curve using known amounts of nucleic acid (analyte). Usually done by diluting a **quantitative standard (external calibrator).**

The external calibrator is not in the same well as the target.

Calibrators do not compete with the target. Targets identical in sequence can be used.

Assay conditions may vary from well to well. Minimal with newer instruments.

Use of Standard Curves



Wang and Morris Anal. Biochem. 269:198-201 (1999)

Applied Biosystems

Quantification Relative to an Endogenous Control

Cellular transcripts whose levels do not change during the course of the experiment can serve as endogenous controls / reference.

18S RNA	ADA	β-actin
GAPDH	Rb	
EF-1α	UbcH5B	MLN51

Warrington et al., *Physiol. Genomics*, **2**:143-147 (2000)
Hamalainen et al., *Anal Biochem*. **299**:63-70 (2001)

Is there an endogenous control ?

- User defined from experience with the system.
- Needs validation.
- Multiple controls.

Spike in a constant amount of an exogenous standard RNA to serve as an quantification standard into all samples.

Normalize data across all samples and identify an endogenous control that changes minimally across the study.

Relative Quantification

1. Using efficiency values in the PCR equation

$$\frac{(1+E_{ref})^{\{Ct_{ref2} - Ct_{ref1}\}}}{(1+E_{tar})^{\{Ct_{sam1} - Ct_{sam2}\}}} = \text{Fold Change}$$

Liu and Saint, Anal. Biochem.**302**:52-59; 2002

Efficiency values derived for target gene and reference can be used with the Ct values obtained to calculate fold change.

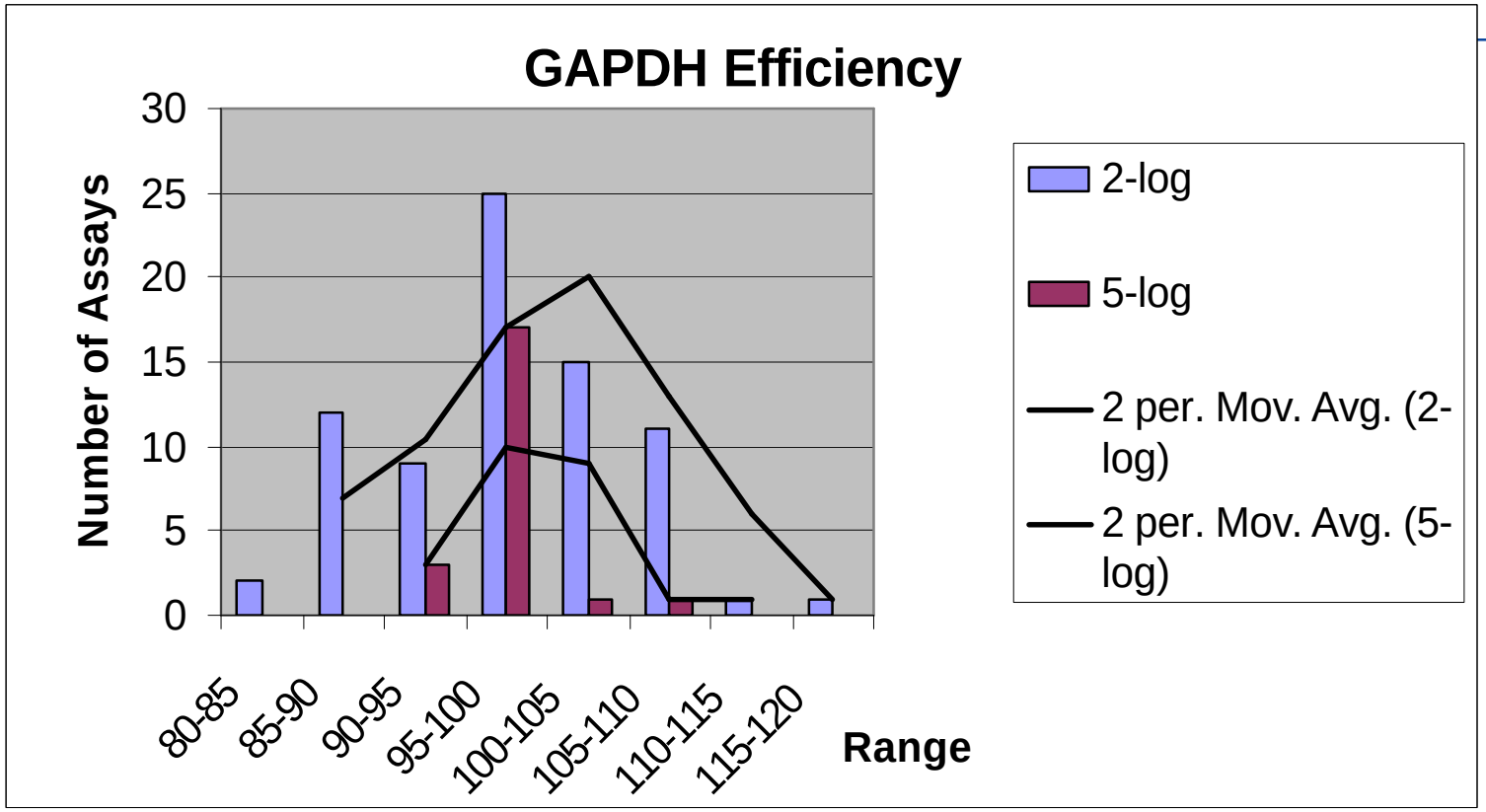
Accurate determination of assay efficiency is difficult. Requires the use of a broad dynamic range (5 to 6-logs) of target dilution and multiple determinations.

2. Using the delta delta Ct method Heid et al Genom Res, **6**:986-994; 1996.

If E ref = E tar =1 the above equation will change to $2^{-\Delta\Delta Ct}$

Most useful for higher throughput. No need for efficiency values.

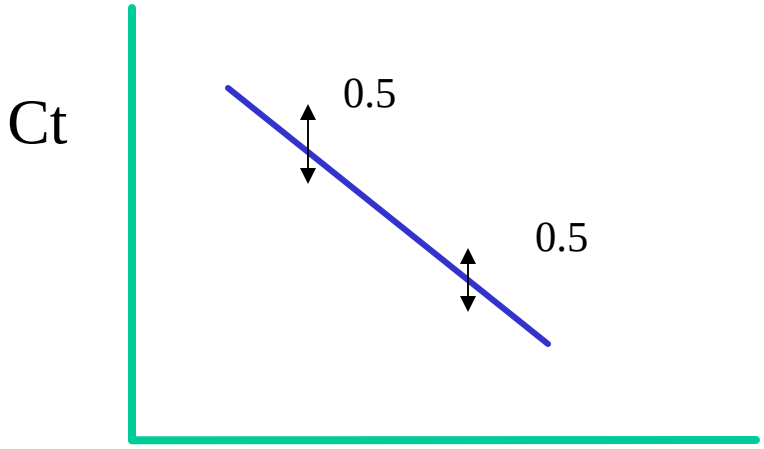
Measuring Efficiency Values



Range 2- log	82-112	97.6 Avg	N= 76
Range 5- log	92-105	98.4 Avg	N= 22

Effect of Ct Measurement Variations on Efficiency Results

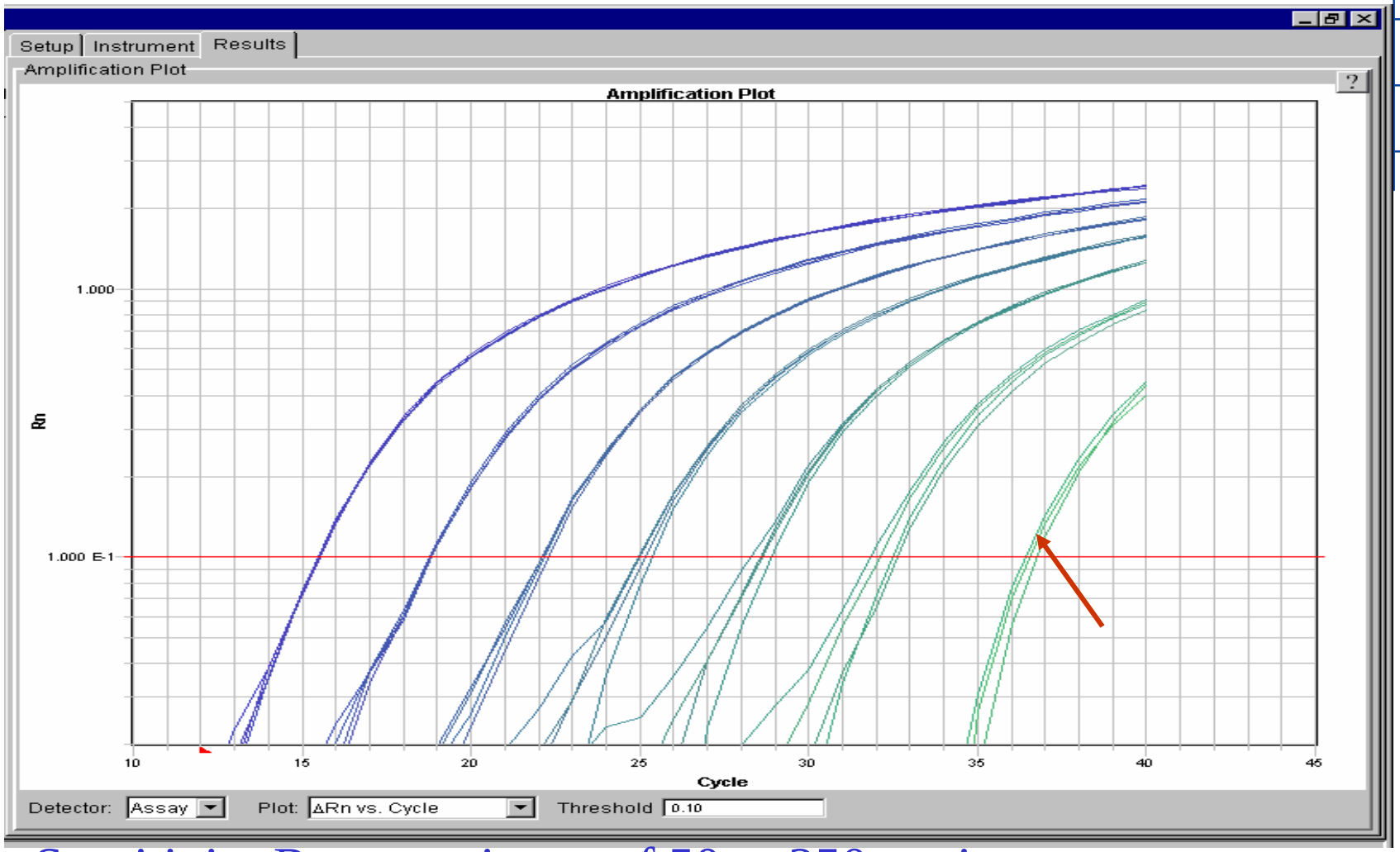
Log fold Dilution	Ct Variation	Measured Efficiency Range Exp 100
1.0	1.0	70 to 168%
2.0	1.0	82-125%
3.0	1.0	87-115%
4.0	1.0	89-110%
5.0	1.0	92-108%



$$\text{Efficiency} = 10^{(-1/\text{slope})} - 1$$

Applied Biosystems

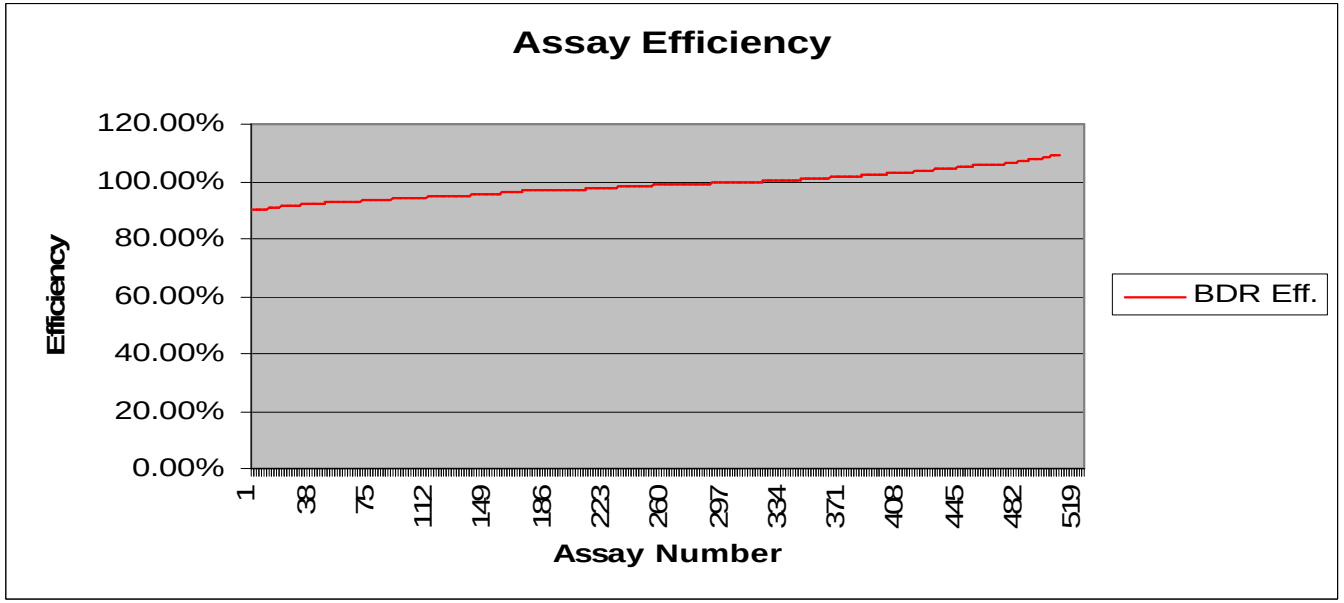
Dynamic Range and Sensitivity



Sensitivity Range estimate of 50 to 250 copies

Dynamic Range: 6-logs

Assay Efficiency:Total 508



Efficiency Range: 90-109% (100 ± 10%)

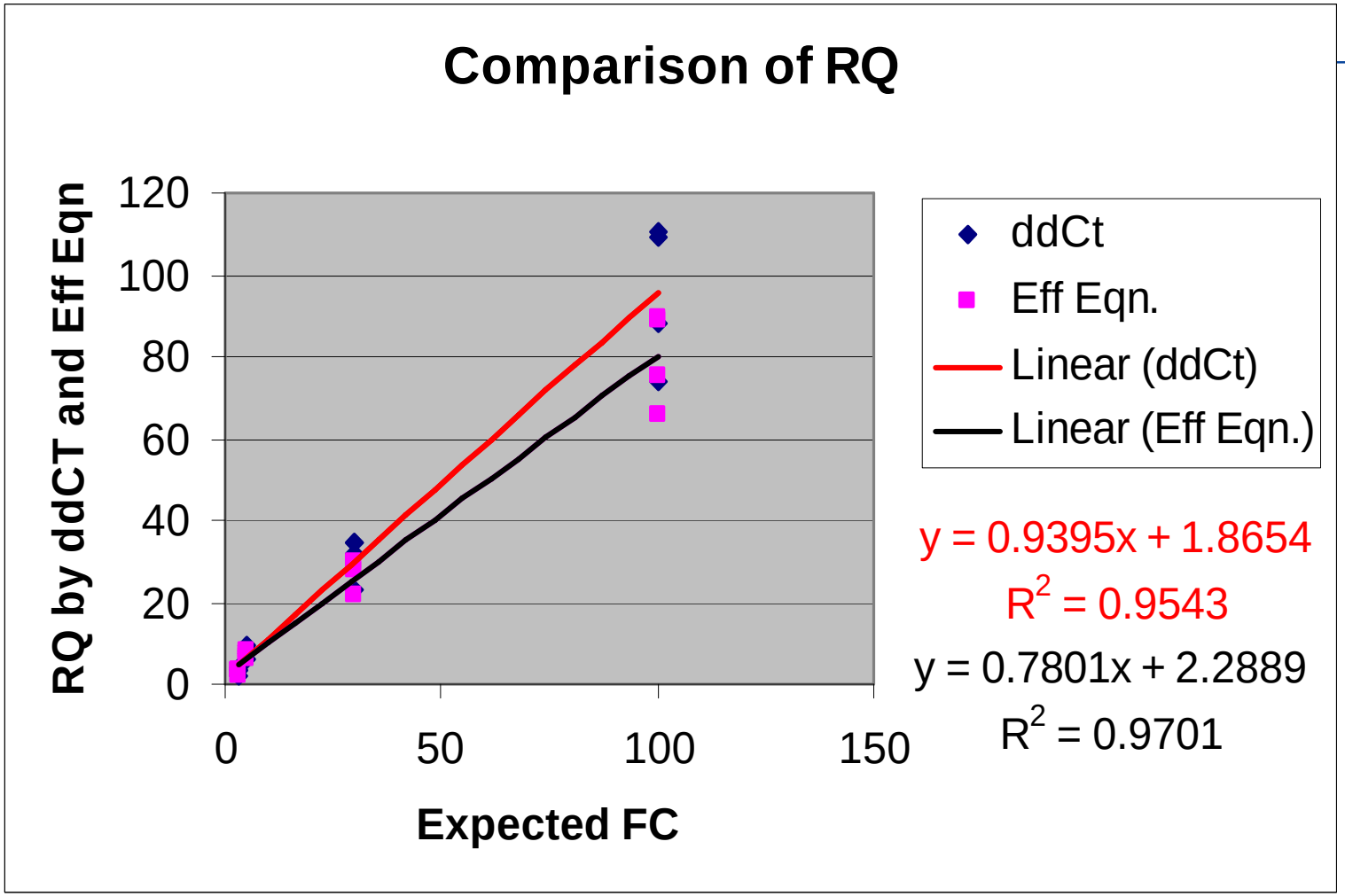
Over 700 assays tested.

Amplicon size range 56 -186.

No correlation to GC content.

No correlation to secondary structure.

Relative Quantification NHT @ 94%



GAPDH @ 97.5%. 3, 10 &100 fold dilution of total RNA

RQ Experimental Design; Reproducibility

1. Number of Assays Selected 15
   - Low Expressors 26-30 Ct 10,000 to 2000 5
   - Moderate Expressors 22-26 Ct 100,000 to 10,000 5
   - High Expressors 18-22 Ct 1,000,000 to 100,000 5
2. Dilutions/ Amounts: 100; 50; 30; 10 and 1 ng /well
3. Number of Operators 3
4. Number of Runs 3
5. Number of sites / instrument 1
6. Number of Replicates 3

RQ: Results Summary

Operator to Operator Variation: 2-Fold Change

	N	Range	AVG	SD	Median	95% CI	< 1.5	Accuracy (1.33)	Accuracy (1.50)
Exp_FC 2 and All OPER	387	0.97 to 6.03	2.23	0.5	2.18	1.23 to 3.23	6 or 1.5%	97.40%	96.80%
Exp_FC 2 and OPER 1	120	1.59 to 4.71	2.29	0.53	2.15	1.23 to 3.35	0	100%	100%
Exp_FC 2 and OPER 2	133	0.97 to 6.03	2.14	0.62	2.03	0.90 to 3.38	6 or 4.5%	92.50%	90.90%
Exp_FC 2 and OPER3	134	1.76 to 3.14	2.29	0.29	2.23	1.71 to 2.87	0	100%	100%

Conclusions : Quantification

1. Internal calibrators or standard curves are necessary for accurate absolute quantification.

2. For relative quantification it is useful to have an endogenous control that is minimally altered during the study / experiment.

3. Quantification standards that can be added into the sample (or during NA extraction) and carried through the entire process are useful for normalization across large studies and identification of appropriate endogenous controls.

Conclusions .. continued

4. Efficiency measurements have a high degree of variability. Using these values in RQ measurements gives less robust results.

5. The ΔΔCt method produced a better correlation to expected results as compared to measurements using efficiency values in PCR equations.

6. Good assay design strategies can ensure near 100% efficiency for real time PCR assays.

7. Relative quantification using the ΔΔCt method generated reliable results.

Assay Specificity

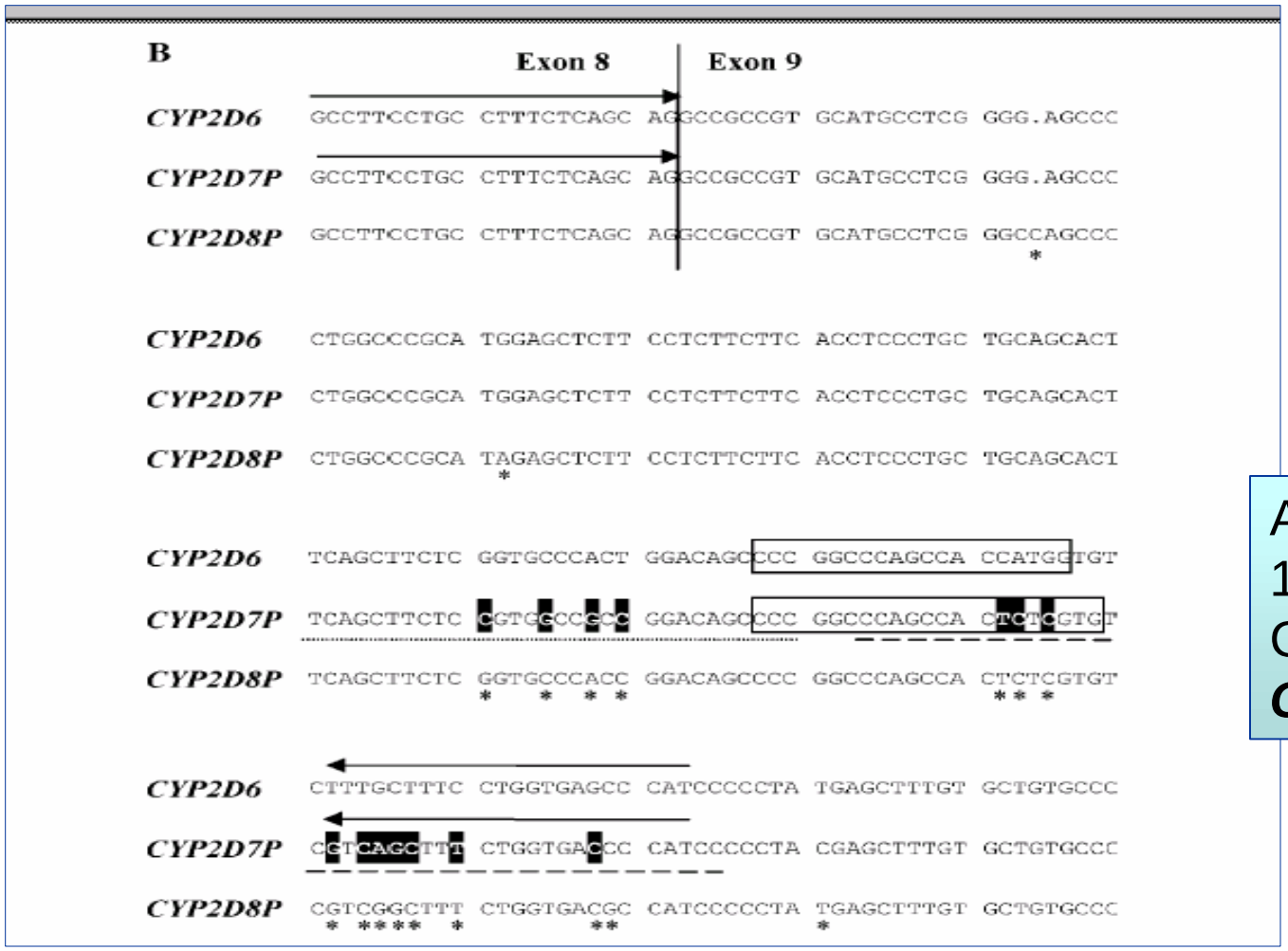
Duplicated Loci in the Human Genome

Science 297:1003 Aug 2002

INTERPRO (entry number)	Description	Number		Enrichment*
		Duplicated	Unique	
003006	Immunoglobulin and major histocompatibility complex	38	280	4.0
001400	Somatotropin hormone family	17	1	31.8
001254	Serine proteases, trypsin family	11	75	4.3
001909	KRAB box	10	87	3.5
001128	Cytochrome P450 enzyme	8	41	5.5
002999	Tudor domain	6	21	7.5
001870	Domain in various γ-carboxylases	5	35	4.2
003877	SPla and the ryanodine receptor (SPRY)	5	42	3.6
001664	Intermediate filament proteins	5	42	3.6
000566	Lipocalin-related protein and Bos/Can/Equ allergen	5	21	6.5
000359	Cystine-knot domain	5	17	7.7
001039	Major histocompatibility complex protein, class I	5	9	12.0
001811	Small cytokines, interleukin 8–like	4	40	3.1
000436	Sushi domain/SCR repeat/CCP module	4	39	3.1
001545	Glycoprotein hormone β chain	4	2	22.5
001271	Mammalian defensin	4	2	22.5
000340	Dual-specificity protein phosphatase	3	39	2.4
003575	Small GTPase, Ras subfamily	3	24	3.7
004045	Glutathione S-transferase NH_2 terminus	3	18	4.8
000863	Sulfotransferase	3	16	5.3
001079	Galectins (previously S-lectins)	3	10	7.8
000971	Globin	3	8	9.2
000461	Glycoside hydrolase family 13	3	3	16.8
000353	Class II histocompatibility antigen, β chain, β_1	3	2	20.2

*Enrichment was calculated as the fraction of duplicated domains for an INTERPRO number over the average fraction for all INTERPRO domains detected in the genome (647 duplicated/21,147 total). Table S7 provides a complete list of all INTERPRO domains examined by this analysis.

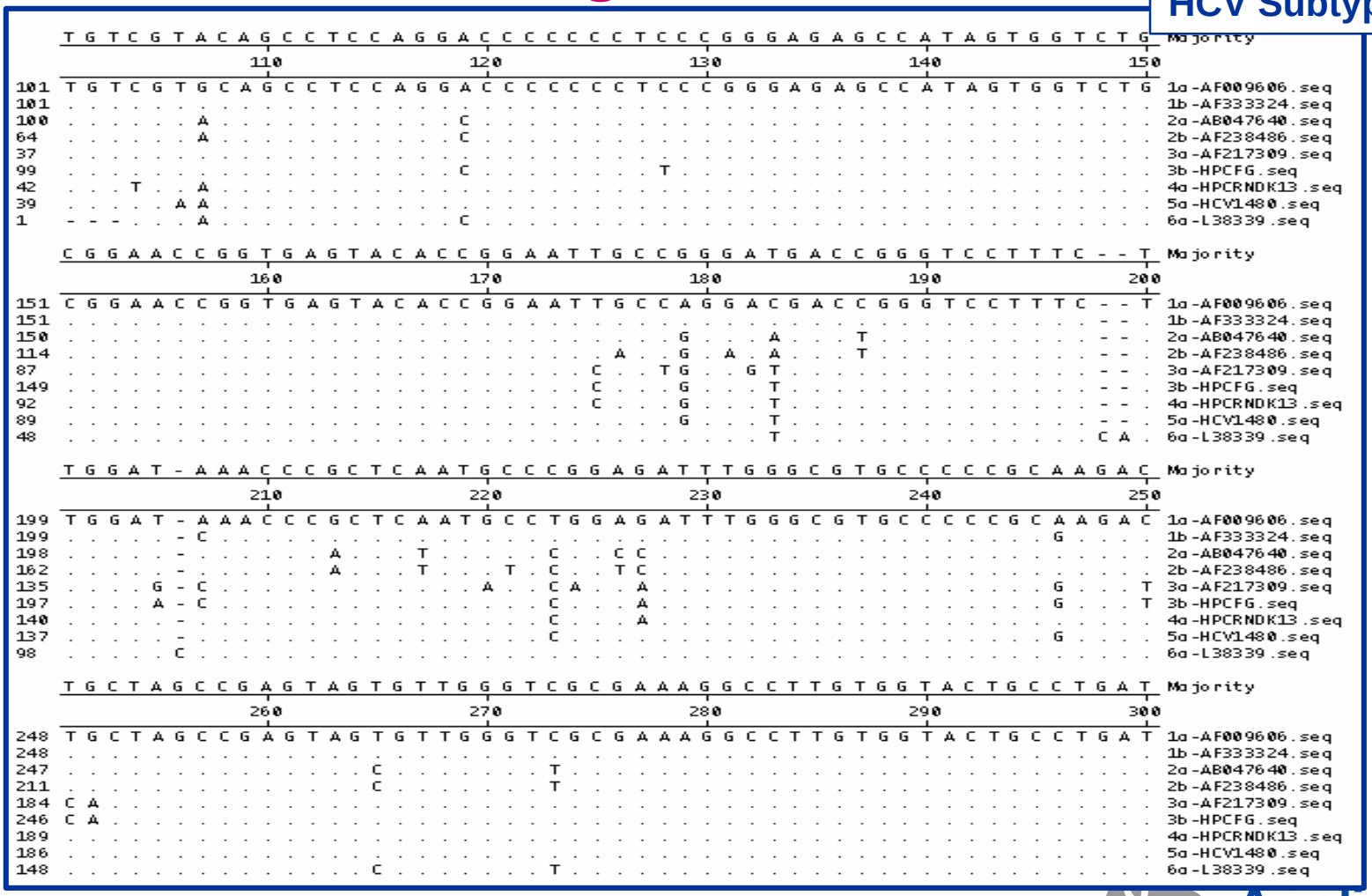
Selectivity by Primer Mismatches



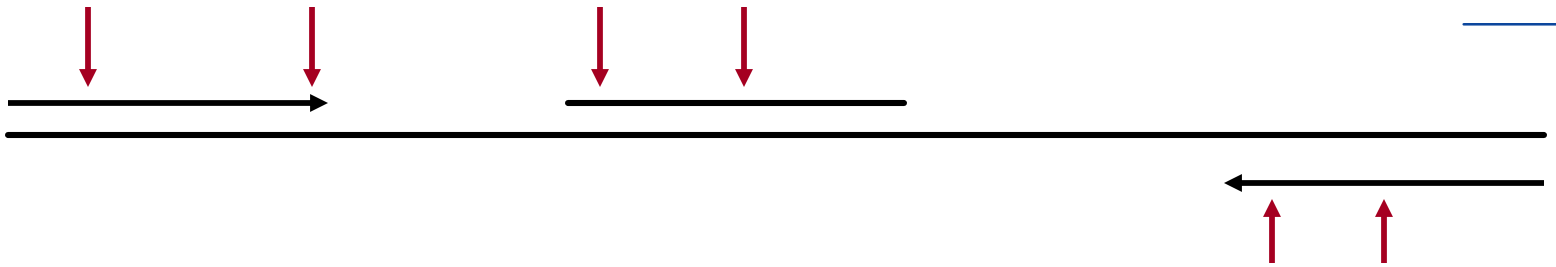
Anal Biochem. 300: 121-131. 2002 Cytochrome P450 ***CYP2D6***

Selectivity by Probe Mismatch: HCV Alignment

HCV Subtypes



Effect of a Mismatch



The Ct value is dependent on both amplification & probe cleavage.

Effect of mismatches in primers & probe on Ct.

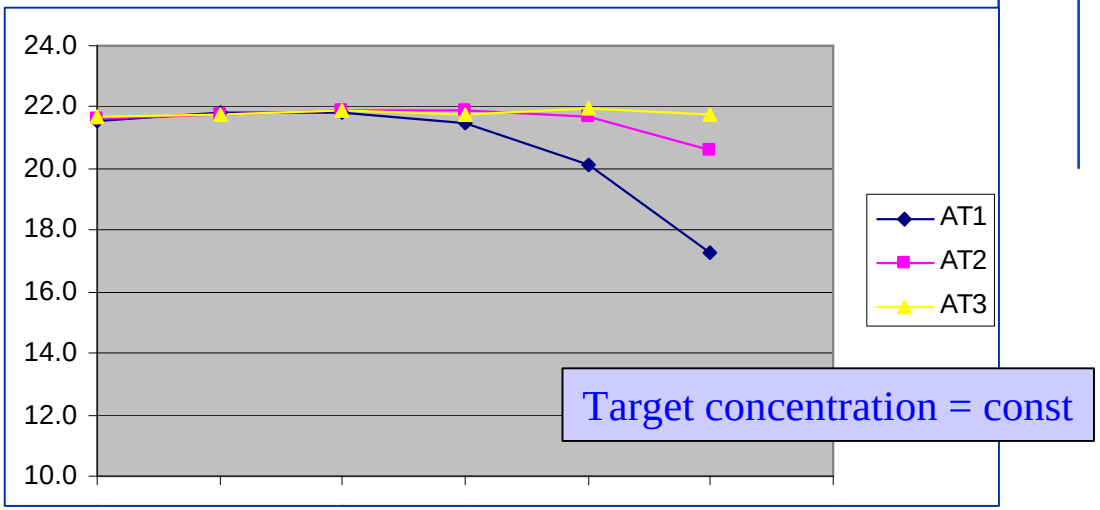
- Position specific effects.
- Mismatch specific effects.
- Combinatorial effects of multiple mismatches

Relative contributions from primers & probes.

In silico QC against the known transcriptome.

Determining Level of Assay Specificity

Ct diff AT1 vs T = 9



Mismatches

Probe

Target T

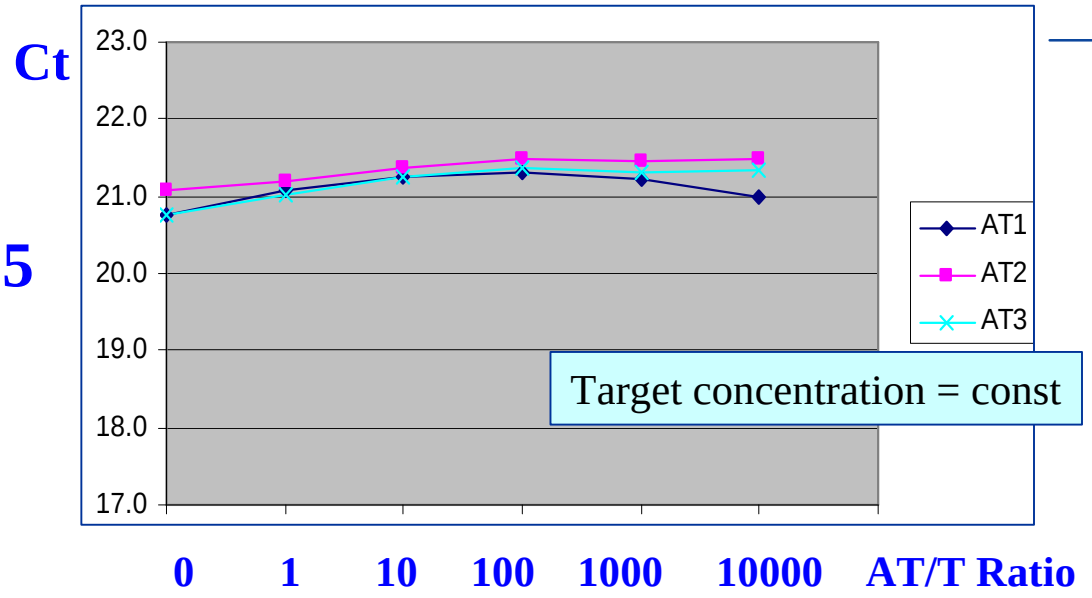
Related gene AT

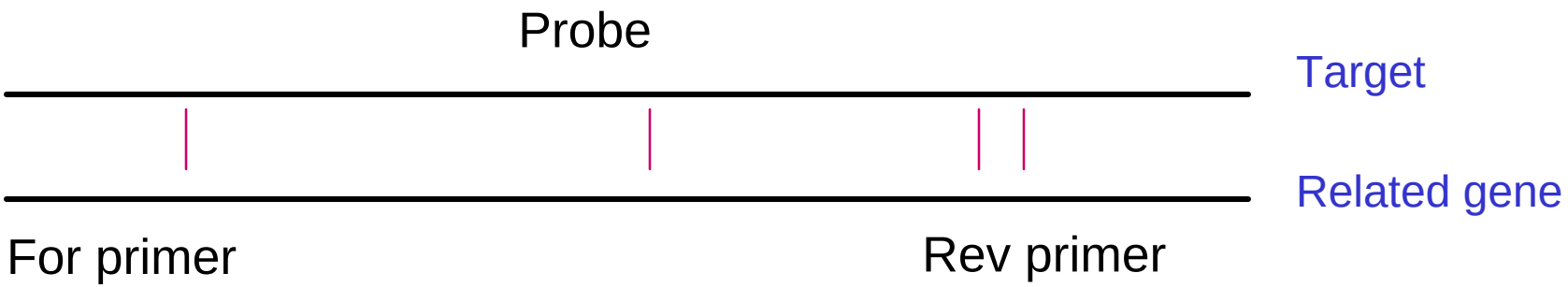
For primer

Rev primer

Determining Level of Assay Specificity

Ct diff AT1 vs T= 15





Transcript Abundance

Classes	Copies/cell	No. transcript/cell	Abundance
Low	1--15	11,000	< 0.004%
Intermediate	200-400	500	< 0.1%
High	10,000	< 10	3%

Based on Abundance & distribution of SAGE tags

Gene Expression: Levels of Induction

Expression levels of the majority of genes in a cell are altered by 2 to 10-fold *in vivo*.

Data from IFN regulated genes; Brian Williams; Cleveland Clinic

Fold Change	IFN -a	IFN -b	IFN -g
>1000	None	None	None
100-1000	None	2	None
50-100	1	1	None
25-50	3	3	2
10 to 25	8	4	4
2 to 10	143	178	194
<2-fold	1185	1152	1140
Total	1340	1340	1340

Conclusions

1. Real time PCR assays that target transcripts emanating from closely related genes should have high selectivity.

2. Suppression of co-amplification can be achieved by placing primers at positions where there are mismatches.

3. Mismatches within probe binding regions will also enhance selectivity.

4. The level of selectivity should be defined by the user based on the study / experiment and is related to transcript abundance and levels of induction / suppression of the related genes.

Acknowledgements

Dennis Gilbert Karl Guegler Ken Livak
Gene Spier

Olga Petrauskene Kathleen Shelton
Karen Poulter
Rixun Fang
Michael Malicdem
Shashi Amur
Ada Wong
Sonny Young
Eileen Westphale

Labeling

For Research Use Only. Not for use in diagnostic procedures.

The PCR process and 5' nuclease process are covered by patents owned by Roche Molecular Systems, Inc. and F. Hoffmann-La Roche Ltd.