

StaRT-PCR[™] Method Validation

An IRB-Approved Study to
Determine the Biological and Analytical Variation
in Transcript Abundance (TA) Levels of Selected Genes
in Normal Human Whole Blood Samples
Measured by *StaRT-PCR*[™] in Gene Express, Inc.
Standardized Expression Measurement (SEM) Center[™]

Drug Information Association 41th Annual Meeting
Washington, DC
June 26-30, 2005

Shelley Ann des Etages, Ph.D.

Principal Scientist, Genomic Technology
Pfizer Global Research and Development
Pfizer Inc., Groton, CT

Sandra Rojas-Caro, Ph.D.

Pfizer Global Research and Development
Pfizer Inc., Ann Arbor, MI



Validation of a New Clinical Assessment Method

- 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? ←

Experimental Design –

Normal & Healthy Subjects

■ Initial Screening Questionnaire

■ Inclusion Criteria

- Age 18-55 years, inclusive
- Healthy males/females - No clinically relevant abnormalities
- Body Mass 18-30 kg/m² and Total Body Weight >50 kg

■ Exclusion Criteria

- Evidence or history of clinically significant medical problem(s)
- Sensitivity to heparin or heparin-induced thrombocytopenia
- Use of tobacco
- Positive result of Breathalyzer test
- Alcohol or substance abuse

■ Adverse Events

■ Concomitant Medications

■ Serum Chemistry Panel – (23) comprehensive metabolic profile, cholesterol & iron content, and hepatitis B & C testing

■ Hematology Analysis – complete blood count and differential & c-reactive protein assessment

■ Immunology – HIV- 1+2 Antibody

■ Urinalysis – (8) standard assays

■ Urine Drug Screen – (10) drug abuse panel

■ Anthropometrics – height & weight

■ Physical Examination – eyes, neck, lungs, heart, & abdomen

■ Vital Signs - heart rate & blood pressure

■ Electrocardiogram – 12 lead

■ Breathalyzer Alcohol Test

■ Carbon Monoxide Test

Experimental Design – *Sample Collection*

■ **Peripheral Blood Collected for Safety Labs**

- Whole blood collected by site's standard venipuncture procedure
- Two tubes drawn for serum chemistry & hematology measurements
- These were the first tubes drawn from each subject

■ **Peripheral Blood Collected for Gene Expression Analysis**

- Whole blood collected per subject per visit was ~ 30 ml
- Collected as 2.5 ml aliquots in 12 PAXgene™ Blood RNA Tubes
- Each PAXgene™ Blood RNA Tube was inverted 10X immediately
- Each PAXgene™ Blood RNA Tube was allowed to set at 18-25°C for 2 hours according to carefully defined SOPs
- Immediately following RT incubation, tubes were stored upright at -80°C

■ **PAXgene™ Blood Tubes shipped on dry ice to Ambion**

Experimental Design – *Sample Extraction Processing*

■ Total RNA Extraction performed by Ambion

- RNA Isolated by a proprietary modification of the PAXgene™ protocol
- RNA extracted from individual tubes of blood
- RNA samples combined into groups of four tubes
 - First four tubes (#1 - #4) were combined to become Sample A
 - Second four tubes (#5 - #8) were combined to become Sample B
 - Last four tubes (#9 - #12) were combined to become Sample C

■ RNA Quality Control Assessment performed by Ambion

- DNase-treated and evaluated for residual DNA contamination levels by using Ambion's SuperTaq™ Real-Time and a TBP TaqMan® probe
- RNA quality assessed spectrophotometrically A_{260}/A_{280} and by the Agilent 2100 Bioanalyzer™ 28S and 18S ribosomal RNA ratio
- Evaluated for nonspecific endonuclease/nickase activity and exonuclease activity
- All RNAs undergo accelerated stability testing in which RNA integrity is checked before and after a 14-18 hr incubation (37°C) - Agilent 2100 Bioanalyzer

■ RNA samples shipped on dry ice to Gene Express, Inc.

Experimental Design – *Sample Processing & StaRT-PCR™ Analysis*

SEM Center™ Analysis by Gene Express, Inc.

■ Reverse Transcription of Samples

- RNA reverse transcribed using MMLV with oligo-dT priming

■ **StaRT-PCR™ performed at the Standardized Expression Measurement (SEM) Center™**

- Analysis with SMIS™ to produce VMTA™ data

- Standardized and quantitative TA values

- Target gene (NT)
cDNA molecules /10⁶ reference gene cDNA molecules

■ **Statistical Analysis performed by Innovative Analytics of Virtually-Multiplexed Transcript Abundance™ (VMTA™) Data**

Completeness of StaRT-PCR™ Data

- 14 (of 15) subjects provided complete data/samples
- 1 subject missing sample C from first visit
- For 13 (of 19) genes all TA levels were greater than 10 molecules/ 10^6 β -actin molecules
- Genes with some TA levels less than 10 molecules/ 10^6 β -actin molecules (*name of genes blinded*):

Gene N	1%	(1 replicate out of 267)
Gene Q	4%	(10 replicates out of 267)
Gene P	16%	(42 replicates out of 267)
Gene O	37%	(98 replicates out of 267)
Gene R	92%	(246 replicates out of 267)
Gene S	100%	(267 replicates out of 267)

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%

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



Study Design

Sources of Variability

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

15 healthy volunteers enrolled



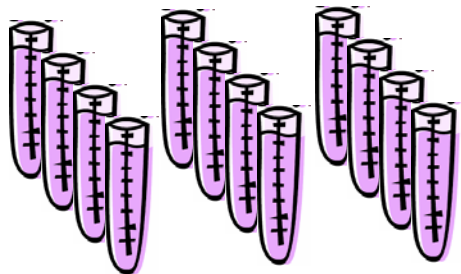
Visit 1

3-7 days

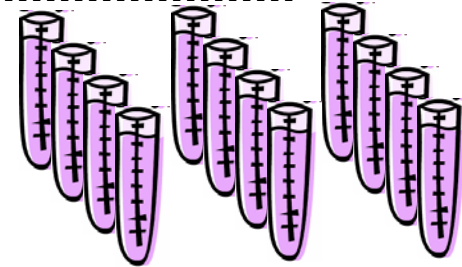
Visit 2

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

Cannot evaluate directly:

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



confounded with other factors or included as random error

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

Methodology for Evaluating Variability

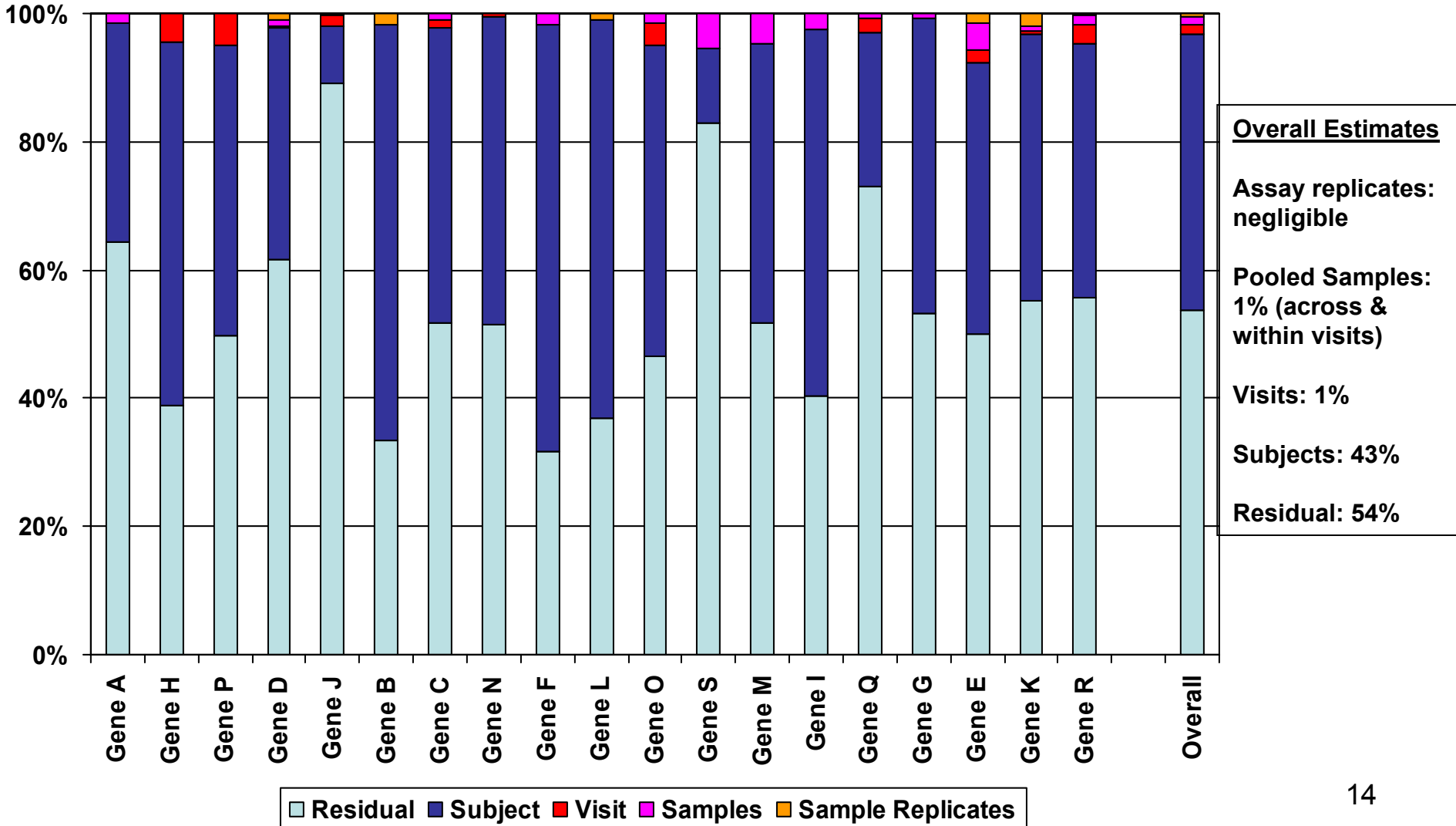
- Estimate the variance components by finding the “best fit” of the preceding model to the actual study data
 - looks at the variability around each source:
 - replicate: how much do the individual assay replicates differ from mean of the replicates
 - sample: how much do the assay means for each of 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

Percentage of Total TAM Variability Attributed to each Component - ACTB

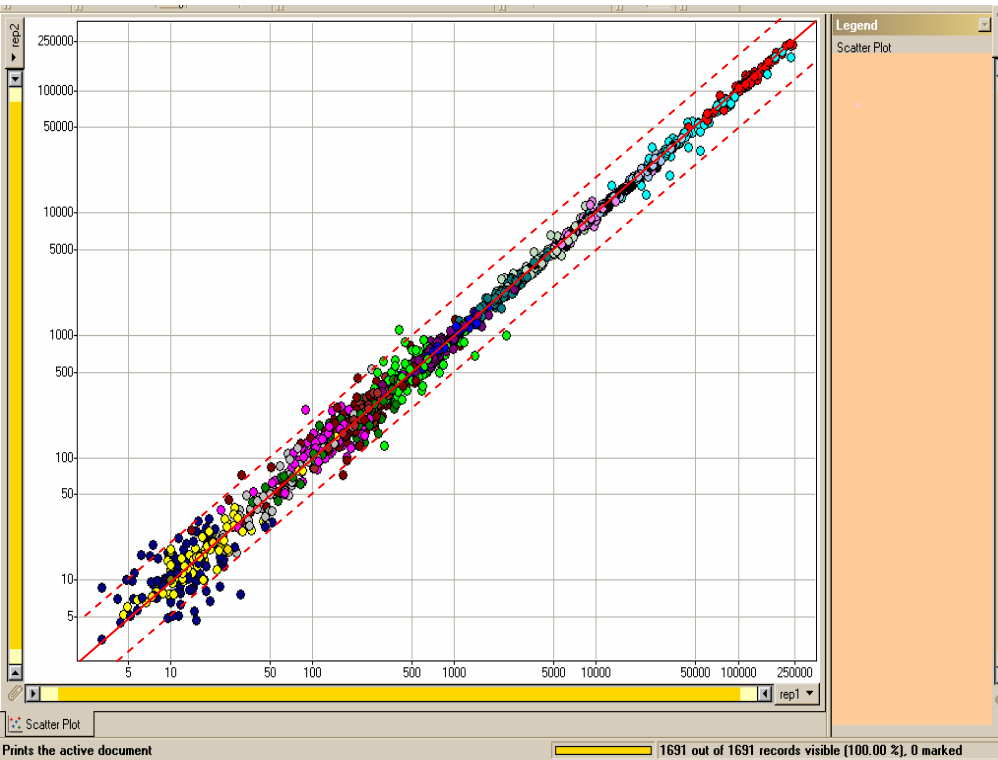
Total %CV (includes all biological + analytical + random variation):

31 39 98 55 42 86 33 57 53 80 124 10 51 44 56 33 31 32 68 Ave: 54%

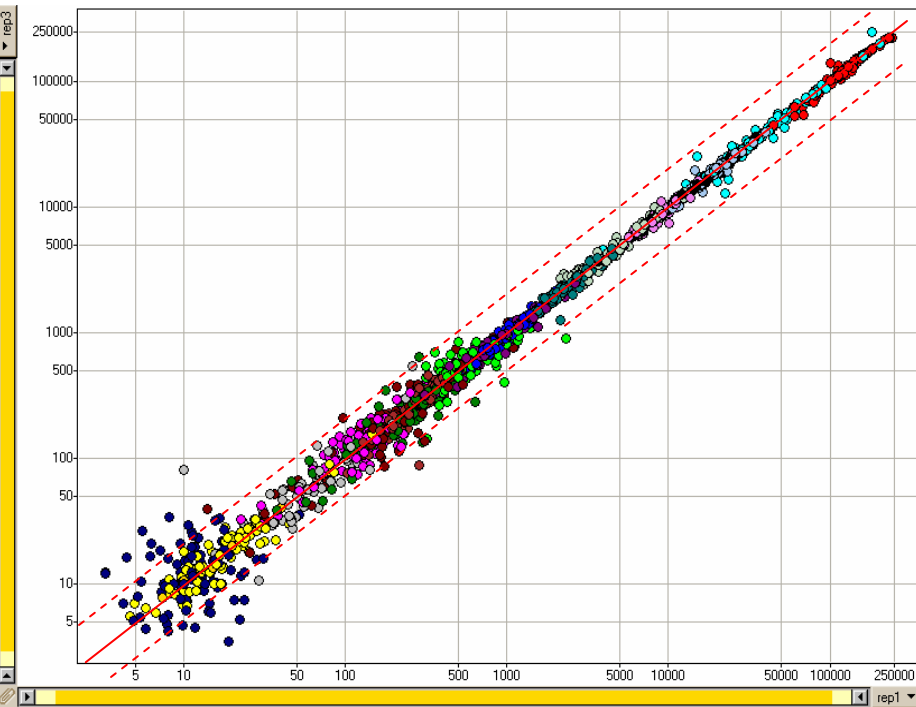


Comparison of *StaRT-PCR*TM TAM Replicates - ACTB

TAM Replicate 1 vs. TAM Replicate 2



TAM Replicate 1 vs. TAM Replicate 3



Average %Coefficient of Variation (%CV) for *StaRT*-PCR™ Assay Replicates

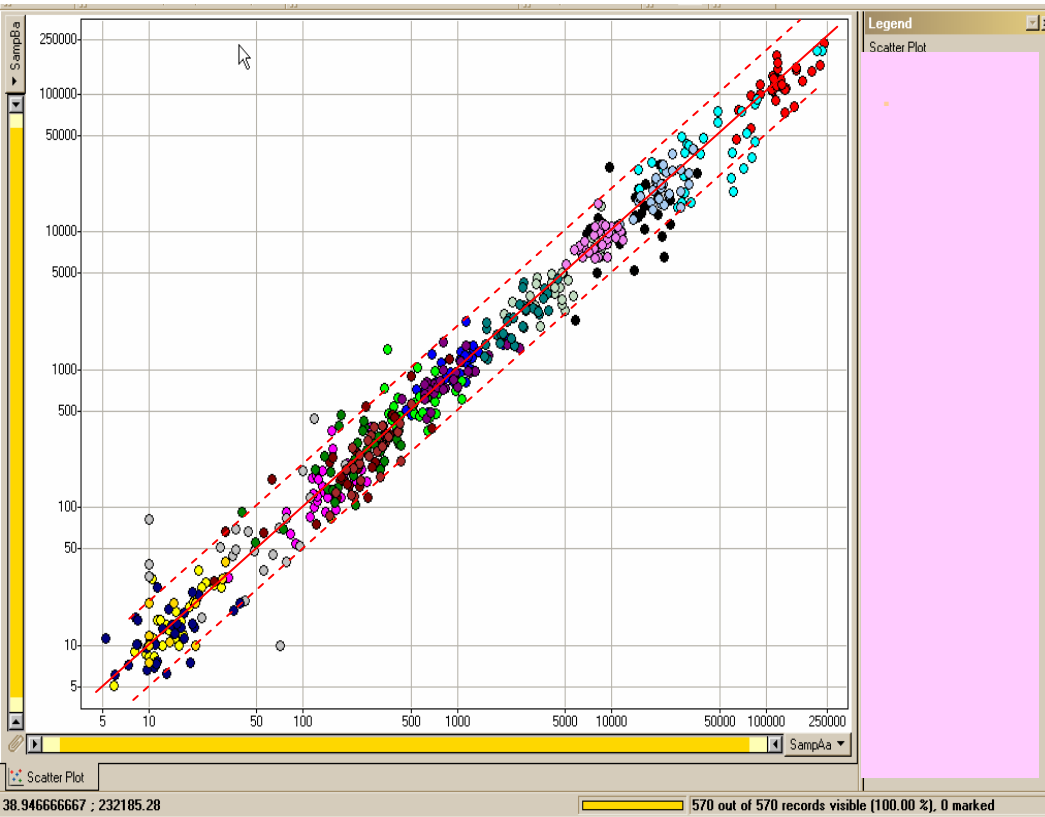
Target gene cDNA molecules / 10⁶ reference gene cDNA molecules

Gene	#Obs*	Average <i>StaRT</i> -PCR TAM - ACTB				Average <i>StaRT</i> -PCR TAM - GAPD			
		Replicate 1	Replicate 2	Replicate 3	Average %CV	Replicate 1	Replicate 2	Replicate 3	Average %CV
Gene A	89	122211	123217	121265	3.2%	2941315	2950933	2898741	3.2%
Gene H	89	1033	1020	1017	5.0%	25317	25109	24944	5.0%
Gene P	89	19	19	19	12.6%	459	466	462	12.8%
Gene D	89	14483	14341	14397	2.5%	326836	323157	324362	2.5%
Gene J	89	605	601	567	15.5%	14399	14463	13747	15.5%
Gene B	89	50420	48943	50979	6.3%	1285131	1256811	1287399	6.3%
Gene C	89	21298	21320	21134	3.1%	539110	537507	533946	3.1%
Gene N	89	161	164	161	16.7%	3948	4082	4058	16.7%
Gene F	89	4929	5007	4948	5.7%	123014	124164	122890	5.7%
Gene L	89	244	255	244	15.1%	5940	6147	5878	15.1%
Gene O	89	53	57	57	11.0%	1303	1385	1391	11.0%
Gene S	89	<10	<10	<10	0.0%	<263	<263	<263	0.0%
Gene M	89	246	238	242	13.3%	5983	5905	5916	13.3%
Gene I	89	882	873	846	6.9%	21464	21467	20677	6.9%
Gene Q	89	14	13	14	34.3%	340	330	364	34.1%
Gene G	89	2503	2519	2459	4.7%	61265	61723	60313	4.7%
Gene E	89	8985	8933	8870	3.6%	226491	225600	222083	3.6%
Gene K	89	288	290	287	7.5%	7233	7264	7131	7.5%
Gene R	89	14	14	14	1.3%	332	329	326	1.3%
Overall	1691	12021	11991	11975	8.9%	294218	293006	291310	8.9%

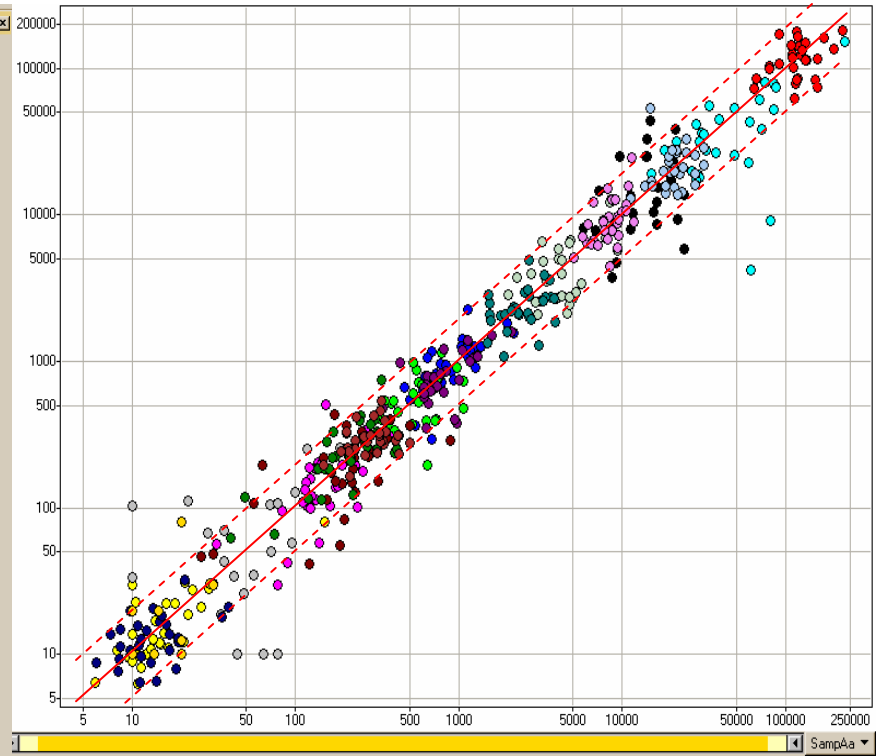
* number of triple replicates across subjects, visits, samples

Comparison of TAM Samples - ACTB

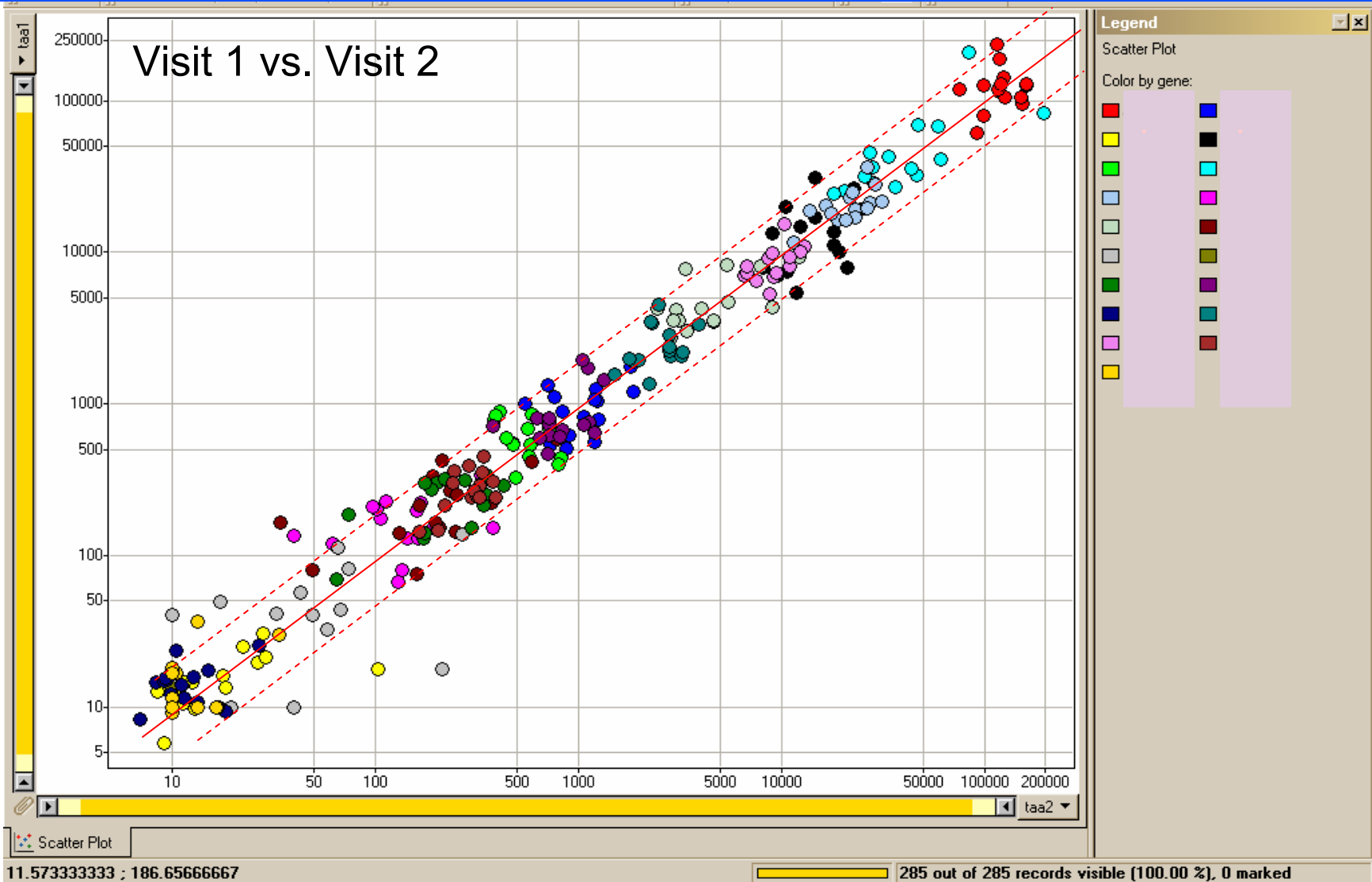
TAM Sample A vs. TAM Sample B



TAM Sample A vs. TAM Sample C



Comparison of TAM Visits - ACTB



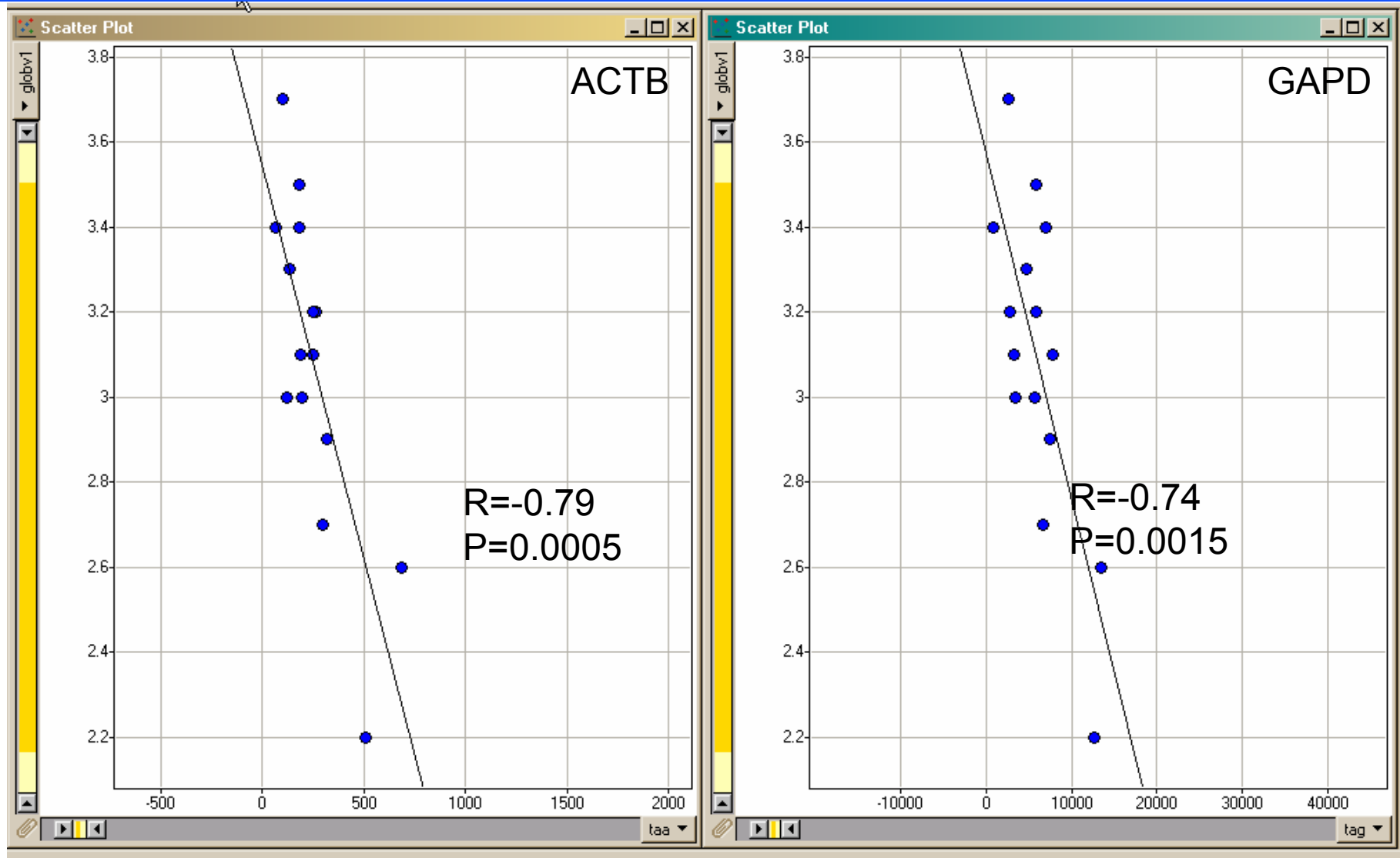
Normal Ranges of Transcript Abundance (TA) Levels Across Subjects

Target gene cDNA molecules / 10⁶ reference gene cDNA molecules

ACTB Gene	#Obs	Mean	SD	1SD Range		%Obs in Range	2SD Range		%Obs in Range	3SD Range		%Obs in Range
				Lower	Upper		Lower	Upper		Lower	Upper	
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%

Association of "Gene L" Expression with Number of Globulin

Globulin



Transcript Abundance

Summary: Clinical Endpoint that Correlate with Gene Transcript

<u>Clinical Endpoint</u>	<u>Target Gene</u>	<u>Reference Gene</u>	<u>Pearson's Corr.</u>	<u>Significance</u>
No. Eosinophils	Gene B	ACTB	R = 0.85	P < 0.0001
No. Eosinophils	Gene B	GAPD	R = 0.86	P < 0.0001
Body Mass Index	Gene B	ACTB	R = -0.85	P < 0.0001
Body Mass Index	Gene B	GAPD	R = -0.82	P = 0.0002
No. Globulin	Gene L	ACTB	R = -0.79	P = 0.0005
No. Globulin	Gene L	GAPD	R = -0.74	P = 0.0015
MCH	Gene F	GAPD	R = 0.78	P = 0.0006
Cardiac PP	Gene Q	GAPD	R = 0.82	P = 0.0002
Cardiac PP	Gene R	GAPD	R = 0.82	P = 0.0002
Cardiac QRS	Gene L	ACTB	R = 0.79	P = 0.0005
Cardiac QRS	Gene L	GAPD	R = 0.74	P = 0.0017
Cardiac QRS	Gene Q	ACTB	R = 0.77	P = 0.0008
Cardiac RR	Gene Q	GAPD	R = 0.83	P = 0.0001
Cardiac RR	Gene R	GAPD	R = 0.77	P = 0.0008
Heart Rate	Gene Q	GAPD	R = -0.82	P = 0.0004
Pulse Rate	Gene Q	GAPD	R = -0.82	P = 0.0002
Pulse Rate	Gene R	GAPD	R = 0.77	P = 0.0007
Race	Gene C	ACTB	R = 0.88	P = 0.0008
Age	Gene N	ACTB	R = 0.76	P = 0.0009
Gender	Gene E	ACTB	R = 0.80	P = 0.0004

Implications for Clinical Endpoint

- TAM with *StaRT-PCR*[™] appears normally distributed across subjects
 - assume similar distribution shapes for other phenotypes
- Multiple normalizer genes
 - GAPD slightly more variable than ACTB
 - better utility for different genes
- Good test-retest however some genes problematic due to very low response (Gene S, Gene R)
- Some significant clinical endpoint associations with phenotype
 - Gene B with BMI (<0.0001) and Eosinophils (<0.0001)
 - Gene L with QRS (<0.001) and globulin (<0.001)
 - Genes Q, R & L with several cardiac measures (~0.0003)
 - Gene N with age (0.0009), Gene E with gender (0.0004) & Gene C with race (0.0008)
 - All clinical endpoints must be clinically validated
- A clinical study would require ~17 subjects per group to detect a 1 SD difference in group mean response at $\alpha=0.05$ with 80% power

Summary - Future implementation of *StaRT-PCR*[™] technology for clinical research

■ *StaRT-PCR*[™] technology appears to be amenable to utilization as a clinical assessment

- Fast & Reliable
- Reproducible
- Standardized & Quantitative multi-gene expression measurement
- Integrated Quality Control for each gene

■ Variation attributed to analytical factors (assay replicates, sample-to-sample and visit-to-visit)

- less than 5% of total combined variation

■ Variation attributed to both the analytical and biological components

- Between subject & within subject residual variation equally account for almost all of the TA variation

■ Study Conclusion

- Using the described *StaRT-PCR*[™] methods, it is possible to define normal ranges of TA measurements values in whole blood.