

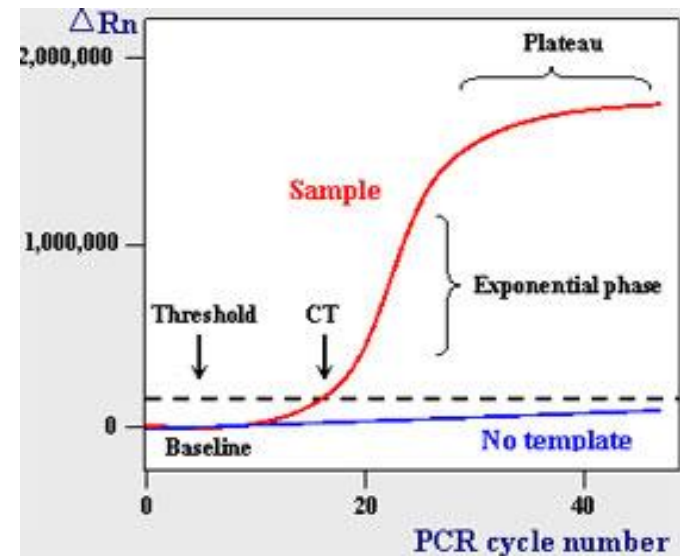
# Estimation of Power and Analysis of qPCR Data with Normal Mixed Models

Auli Partanen, Éva Tas, Juha Akkila, Sami Hokkanen  
Orion Corporation Orion Pharma, Finland

**Statistikot Suomen Lääketeollisuudessa syysseminaari**  
**Espoo, Finland on November 3<sup>th</sup>, 2008**

# qPCR

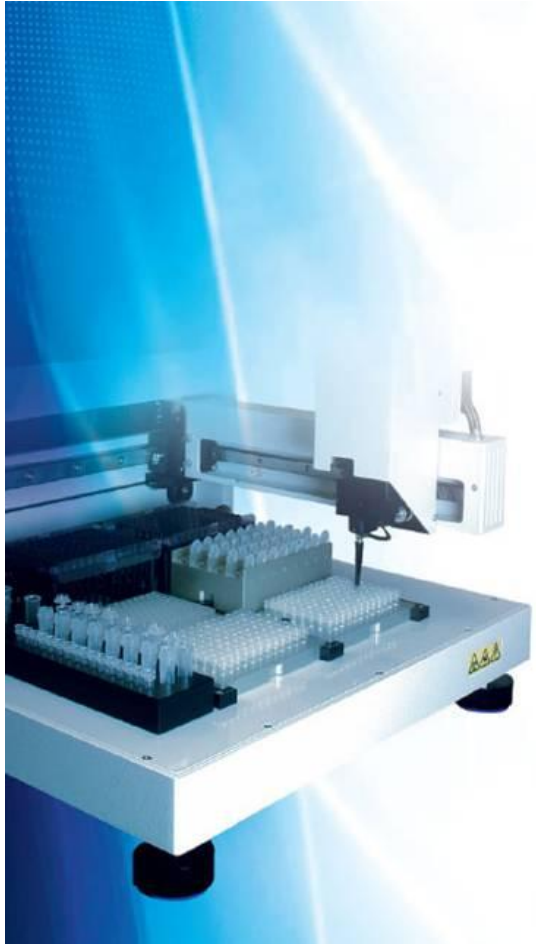
- Is used to quantify DNA or messenger RNA (mRNA) in a sample
- When combined with reverse transcriptase (RT-PCR), relative gene expressions between tissues or genes can be compared
- Measure fluorescence reporter



<http://www.rt-pcr.com/>

<http://www.gene-quantification.info/>

# Standardizing measurements



- Significant problems caused by
  - variability of RNA templates, assay designs and protocols
  - various data normalization
  - data analysis strategies
- Which are tried to control by
  - consistently using standard chemistries, protocols and reaction conditions
  - pipetting robot
  - repeated measurements (technical repeats)
  - all measurements (to be compared) at the same time
  - controlling efficiency in each run

# Aim of study and Study design

- **Aim of study is to evaluate relative expression ratio between genes and treatments**
- **18 animal tissue samples in 4 groups**
  1. control (n=4)
  2. disease model (n=5)
  3. disease model with study drug treatment (n=4)
  4. disease model with reference drug treatment (n=5)
- Expression levels of **3 target genes** and **12 potential reference genes** were analyzed by quantitative RT-PCR in **3 replicates** for each sample.

# Study design

	Original data		Selected data		Analysis data	
<b>Samples</b>	18		8		8	
<b>Treatments</b>	4		2		2	
<b>Genes</b>	3 targets ----- 12 references	X 3 technical repeats	1 target ----- 3 references	X 3 technical repeats	1 target ----- Mean* over references	Mean* over technical repeats
<b>N</b>	810		96		16	

\*) Arithmetic mean

# Analysis Data

Treatment	Gene	Sample	Ct	Subgroup
Control	Target A	28	35.6	1
		32	38.1	1
		33	37.3	1
		34	35.7	1
Control	Reference	28	24.7	2
		32	28.3	2
		33	26.5	2
		34	25.1	2
Study drug treatment	Target A	66	37.1	3
		69	36.3	3
		70	36.6	3
		71	37.1	3
Study drug treatment	Reference	66	24.5	4
		69	24.9	4
		70	23.8	4
		71	25.4	4

# Definitions

- For reference gene:  $\text{mean } r_{\text{control}}$  and  $\text{mean } r_{\text{treatment}}$
- $\Delta\text{Ct}_{\text{ref}} = \text{mean } r_{\text{control}} - \text{mean } r_{\text{treatment}}$
- For target gene:  $\text{mean } t_{\text{control}}$  and  $\text{mean } t_{\text{treatment}}$
- $\Delta\text{Ct}_{\text{target}} = \text{mean } t_{\text{control}} - \text{mean } t_{\text{treatment}}$
- $\Delta\Delta\text{Ct} = \Delta\text{Ct}_{\text{ref}} - \Delta\text{Ct}_{\text{target}}$
- Expression ratio = 
$$\frac{(E_{\text{target}})^{\Delta\text{Ct}_{\text{target}}}}{(E_{\text{ref}})^{\Delta\text{Ct}_{\text{ref}}}} = 2^{-\Delta\Delta\text{Ct}_{\text{adjusted}}}$$

# PAE and adjusted $\Delta\Delta Ct$

- qPCR data analysis is based on the assumption that PCR products double each cycle (AE=2).
- When the AE (Amplification Efficiency) is not 2, Ct -values are recommended to be adjusted.
- We used percentile AE (PAE) instead of AE

$$AE = 2^{PAE}$$

$$PAE = \log_2(AE)$$

- $\Delta\Delta Ct_{\text{adjusted}} = PAE_{\text{ref}} * \Delta Ct_{\text{ref}} - PAE_{\text{target}} * \Delta Ct_{\text{target}}$
- Efficiency can be estimated for a group of reactions or a single reaction by simple regression model.

# Normal Mixed Model with Gene as a repeated factor

- Effects

- gene (target, reference)
- treatment (study drug, control)
- gene by treatment interaction
- sample (sample number)
- residual

- Mixed Model in SAS

```
PROC MIXED; CLASS gene treatment sample;  
MODEL Ct = gene treatment gene*treatment;  
REPEATED gene / SUBJECT = sample TYPE = UN;
```

# Estimation of $\Delta\Delta Ct$ based on Mixed Model parametrization (A) and (B) ?

## Mixed Model in SAS (parametrization A)

```
PROC MIXED; CLASS gene treatment sample;
MODEL Ct = gene treatment gene*treatment;
REPEATED gene / SUBJECT = sample TYPE = UN;
```

```
 $\Delta\Delta Ct$ : ESTIMATE gene*treatment +1 -1 -1 +1;
```

## Mixed Model in SAS (parametrization B)

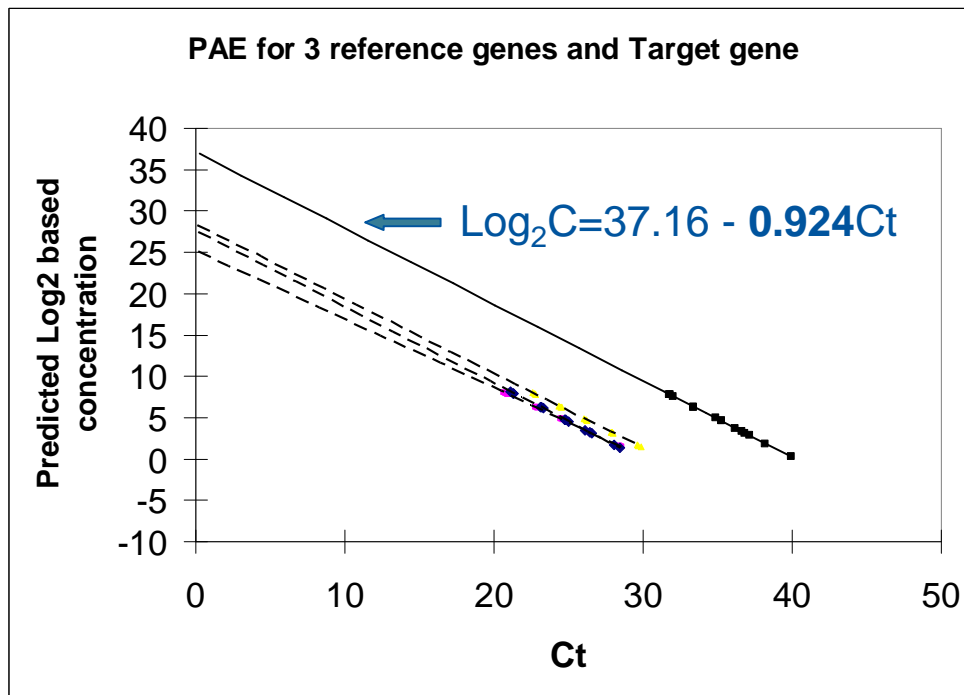
```
PROC MIXED; CLASS subgroup sample;
MODEL Ct = subgroup;
REPEATED / SUBJECT = sample TYPE = UN;
```

```
 $\Delta\Delta Ct_{adjusted}$ : ESTIMATE subgroup '+PAEtarget -PAEtarget -PAEref +PAEref' ;
```

- (subgroup is a categorical variable with 4 classes)

# Estimation of PAE

PAE estimates are based on the data of 5 different dilutions (and 3 technical repeats per dilution) over the pooled samples treated by 4 treatments.



Gene	PAE	
Reference 1	0.831	Mean PAE of reference genes: 0.888
Reference 2	0.929	
Reference 3	0.903	
Target A	<b>0.924</b>	

# Table with descriptive $\Delta\Delta\text{Ct}$ and model estimates

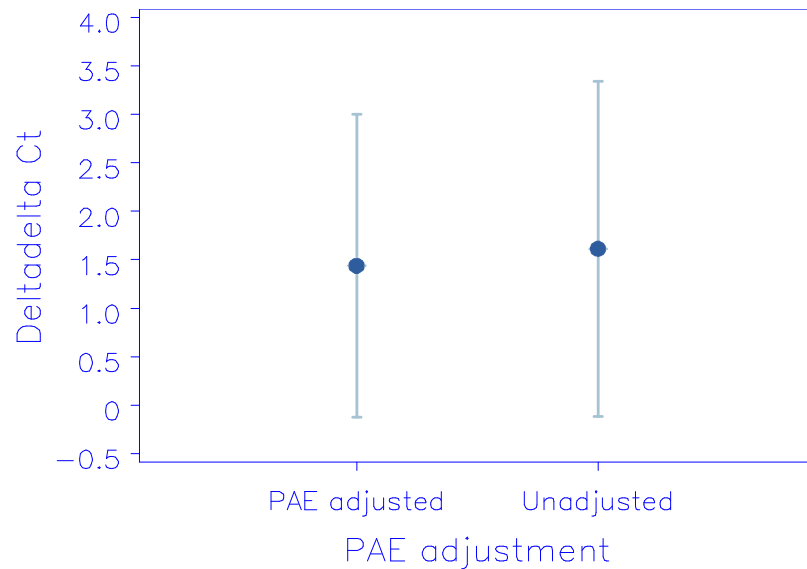
---

	$\Delta\Delta\text{Ct}(\text{SE})$	
	Descriptive statistics	Normal mixed model
<b>PAE Adjusted</b>	1.437 (0.708)	1.439 (0.729)
<b>Unadjusted</b>	1.613 (0.786)	1.614 (0.806)

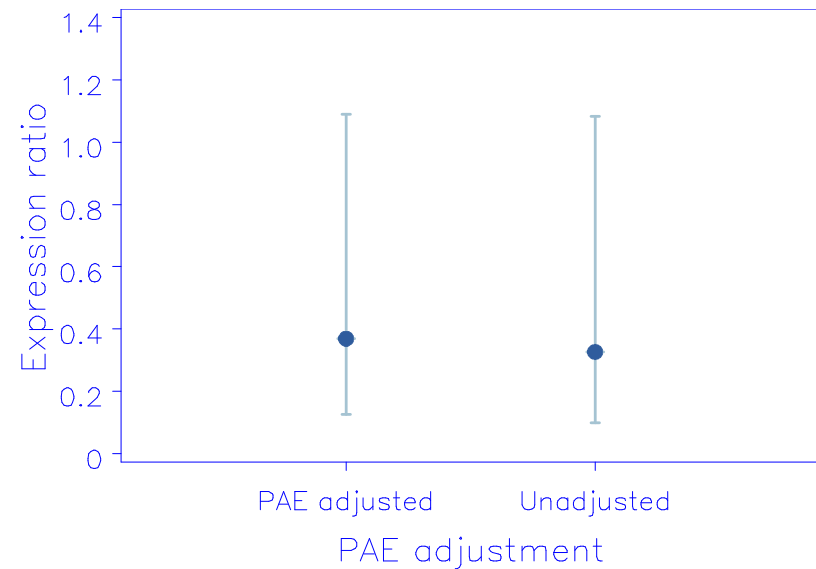
---

# $\Delta\Delta Ct$ and expression ratio with 95% CIs

## $\Delta\Delta Ct$ s



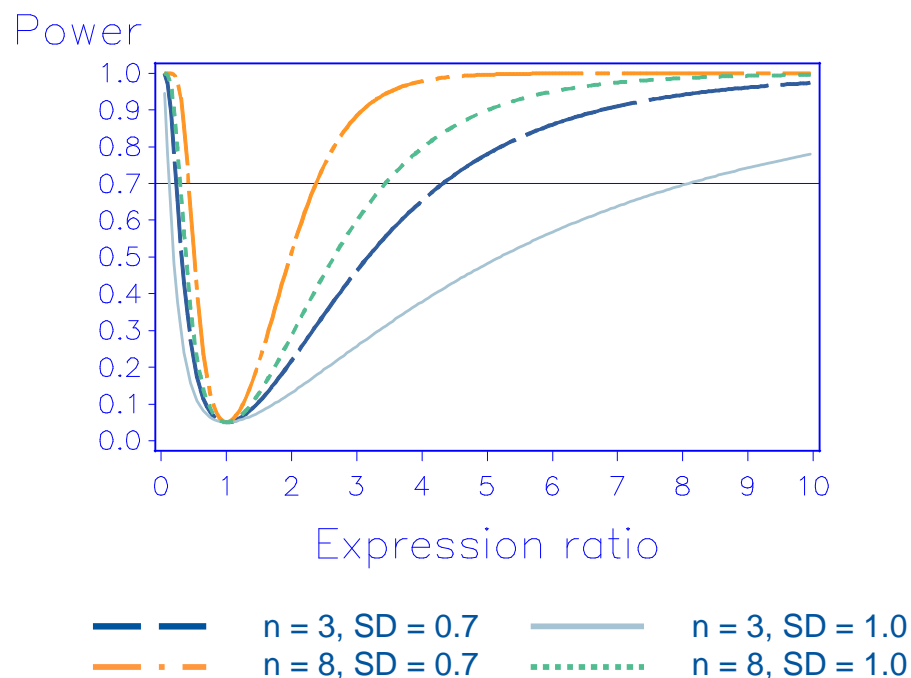
## Expression ratios



$$\text{Expression ratio} = 2^{-\Delta\Delta Ct}$$

# Power

Power curves for the expression ratio between 0 and 10 when type I error is 0.05



Power calculation is based on

- noncentral t-distribution and the log<sub>2</sub> transformed expression ratio ( $\Delta\Delta Ct$ )
- variance was estimated as the sum of four equal variances
- degrees of freedom was estimated as a sum of n subtracted by a number of groups

# Conclusion

- Importance of power calculation in study planning phase.
- Normal mixed model works well with qPCR data and enables the dependence between genes.
- With balanced, complete data these two methods give similar results.
- Mixed model gives more accurate estimates with unbalanced data.
- Data transfer from instrument to analysis software is challenging.

# References

- [1] Brown, H. and Prescott, R. (2006): Applied mixed models in medicine.
- [2] Yuan, J., Reed, A., Chen, F. and Stewart, C. N.: Statistical analysis of real-time PCR data. BMC Bioinformatics 2006, 7:85.
- [3] Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D. and Schabenberger O. (2006): SAS® for Mixed Models, Second Edition. Cary, NC: SAS Institute Inc.
- [4] Yuan, J., Wang, D. and Stewart, C. N.: Statistical methods for efficiency adjusted real-time PCR quantification. Biotechnology journal 2008, 3, 112-123.

# Thank you!

## Questions?