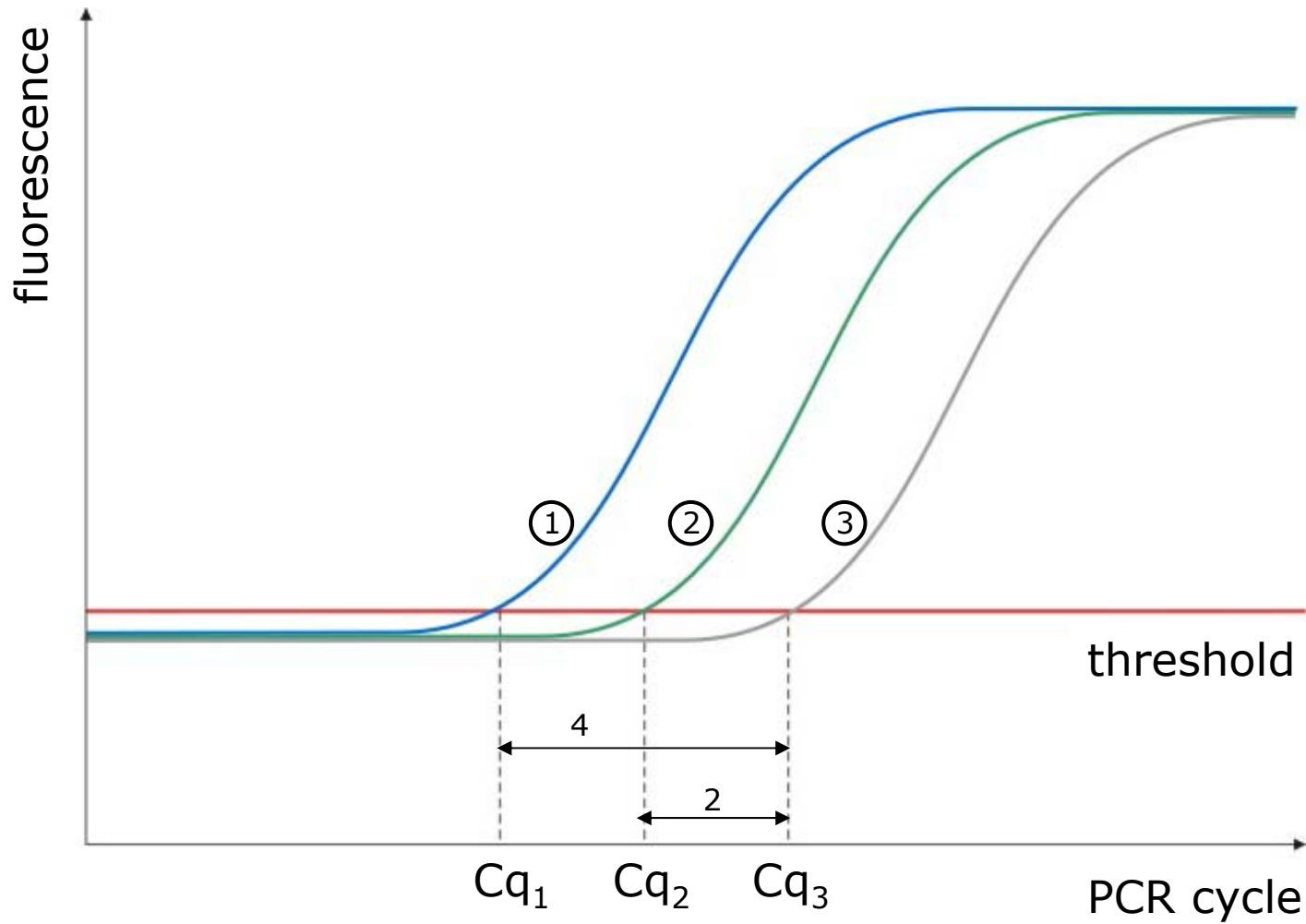


Advanced and universally applicable models for relative quantification

Jan Hellemans
CMGG UGent

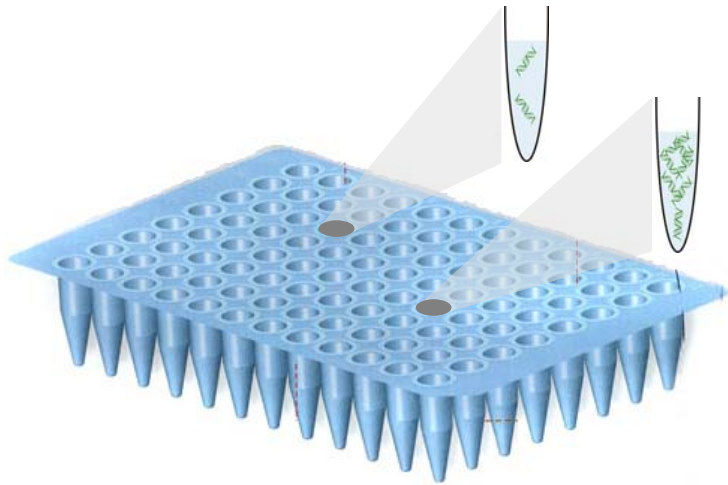


$$RQ = 2^{\Delta Cq}$$

$$RQ_{1/3} = 2^4 = 16$$

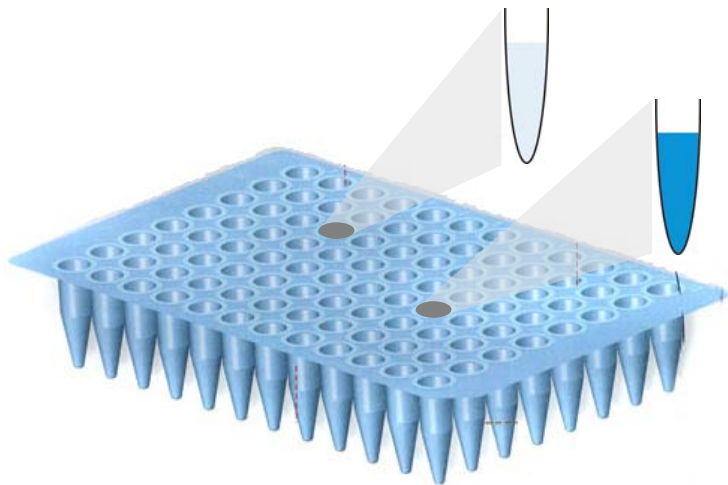
$$RQ_{2/3} = 2^2 = 4$$

$$RQ_{3/3} = 2^0 = 1$$



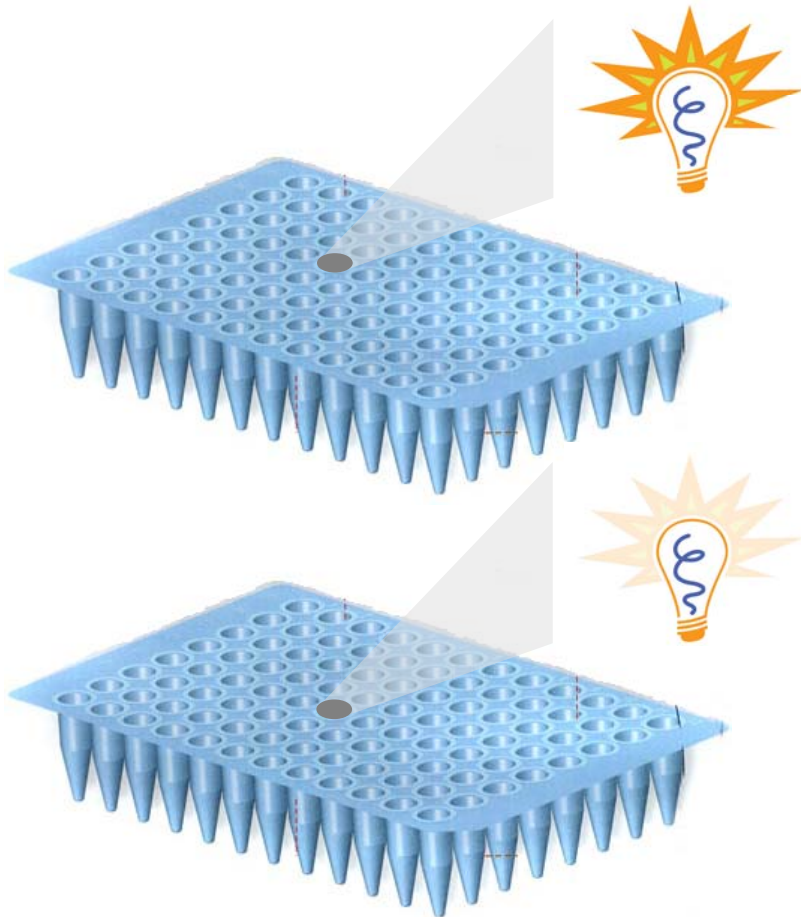
Differences in RQ due to:

- different gene expression level
= variation of interest



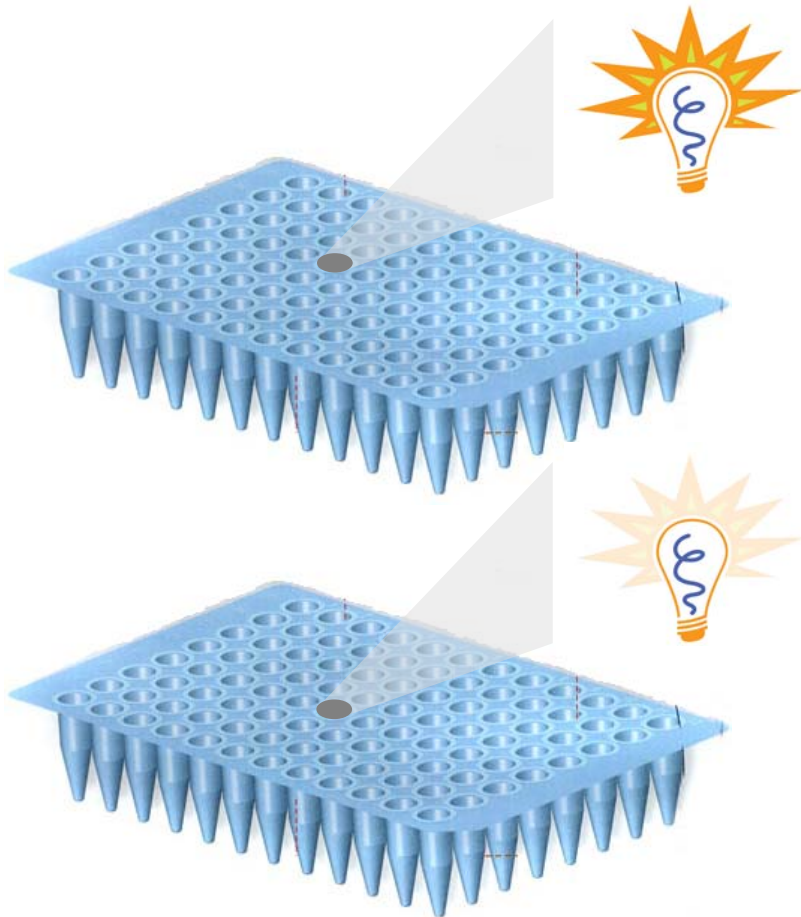
Differences in RQ due to:

- different gene expression level
- different total starting amount
 - ≠ cell numbers
 - ≠ RNA isolation
 - ≠ RNA degradation
 - ≠ cDNA synthesis



Differences in RQ due to:

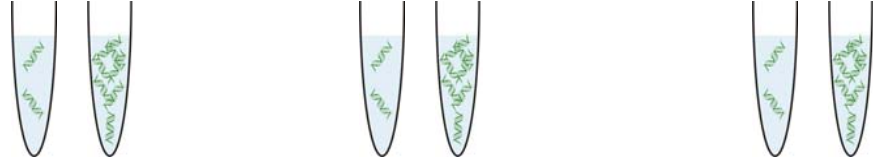
- different gene expression level
- different total starting amount
- run dependent differences
 - \neq data analysis settings



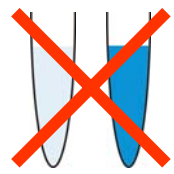
Differences in RQ due to:

- different gene expression level
- different total starting amount
- run dependent differences
 - ≠ data analysis settings
 - instrument related variation
 - ≠ reagents
 - ≠ optical properties of plastics

Cq → RQ → NRQ → CNRQ



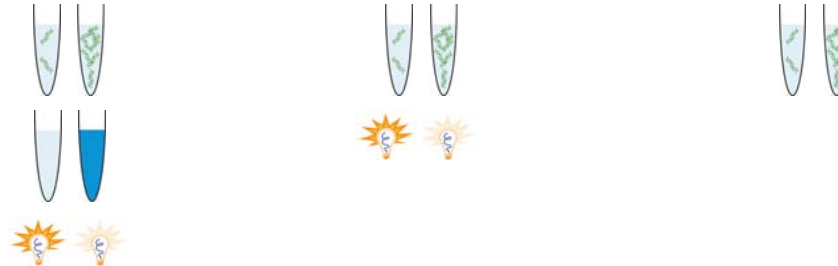
Normalization



Inter-run calibration



Avoid variation
> Minimize & correct variation
>>> Ignore variation



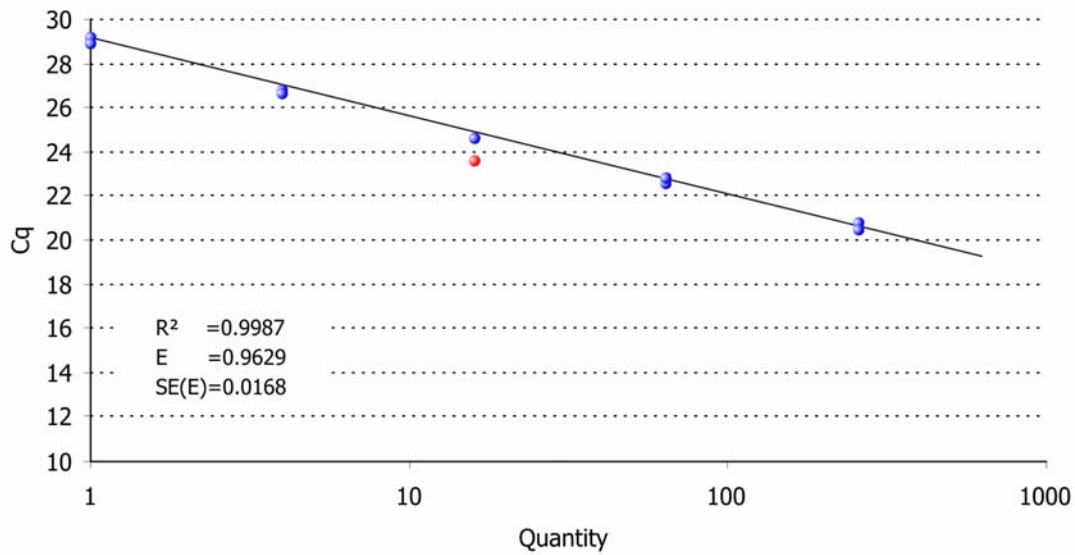
$$RQ = 2^{\Delta Cq}$$

$$RQ = E^{\Delta Cq}$$



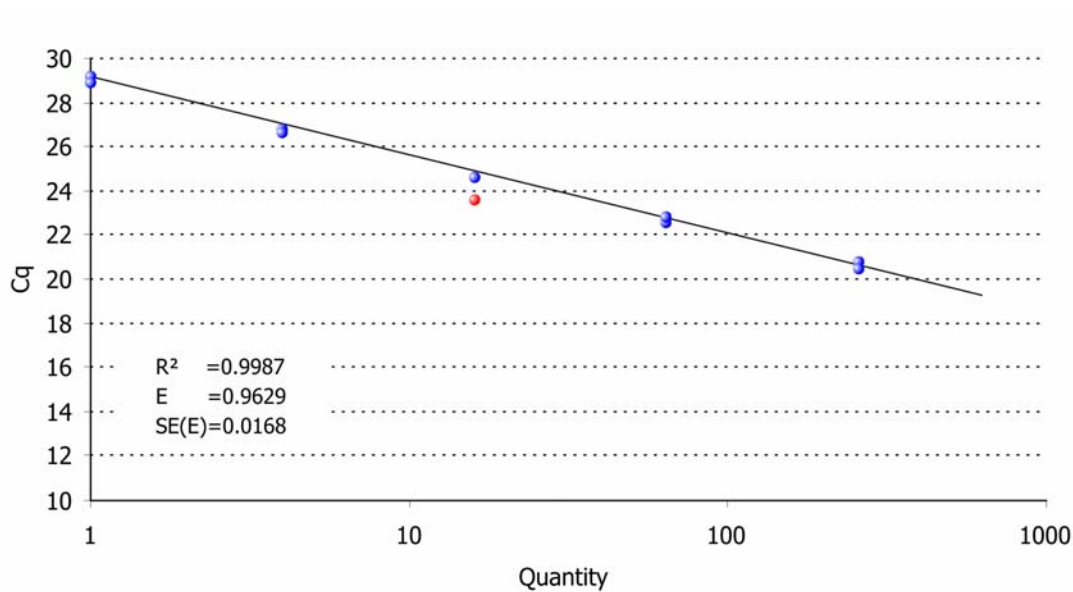
Calculate gene specific amplification efficiency

- Dilution series
- Average of well specific amplification efficiencies



$$slope = \frac{\sum_{a=1}^n (Q_a - \bar{Q})(C_{Q_a} - \bar{C}_Q)}{\sum_{a=1}^n (Q_a - \bar{Q})^2}$$

$$E = 10^{\left(\frac{-1}{slope}\right)}$$



$$SE(slope) = \frac{s_e}{s_x(n-1)}$$

$$s_e = \sqrt{\frac{\sum_{a=1}^n (Cq_{a,measured} - Cq_{a,predicted})^2}{n-2}}$$

$$s_x = \sqrt{\frac{1}{n-1} \sum_{a=1}^n (Q_a - \bar{Q})^2}$$

$$SE(E) = \frac{E \cdot \ln(10) \cdot SE(slope)}{slope^2}$$

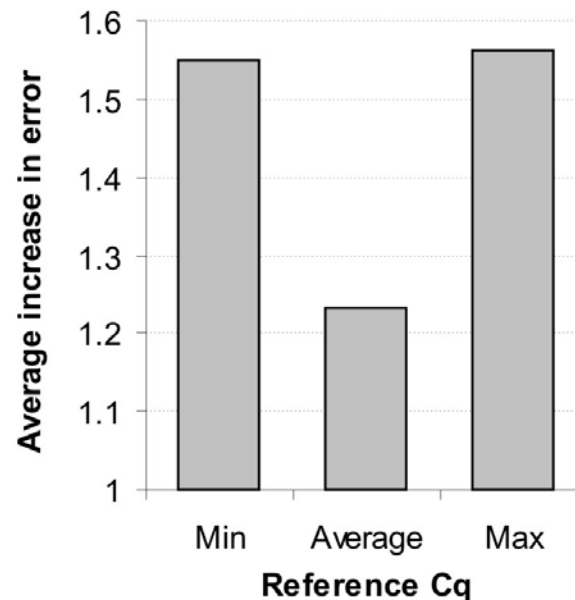
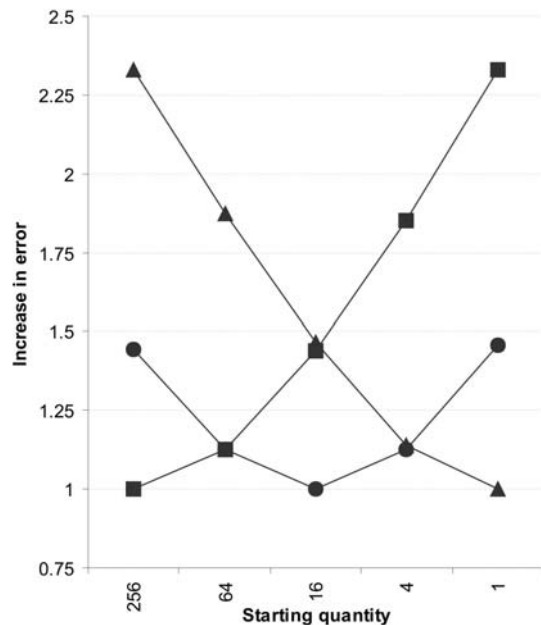
Decrease SE(E)

- ➔ Increase number of dilution points (n)
- ➔ Increase range of dilution

$$RQ = E^{\Delta Cq} \quad SE(RQ) = \sqrt{RQ^2 \left[\left(\frac{\Delta Cq \cdot SD(E)}{E} \right)^2 + (\ln(E) \cdot SD(Cq_{ref}))^2 \right]}$$

$$\Delta Cq = Cq_{ref} - Cq_{soi}$$

- Choice of Cq_{ref} does not affect the RQ ratio between samples
- Choice of Cq_{ref} does affect the error on RQ (if $SE(E)$ is taken into account)
- Minimize error: $Cq_{ref} = \text{average } Cq$

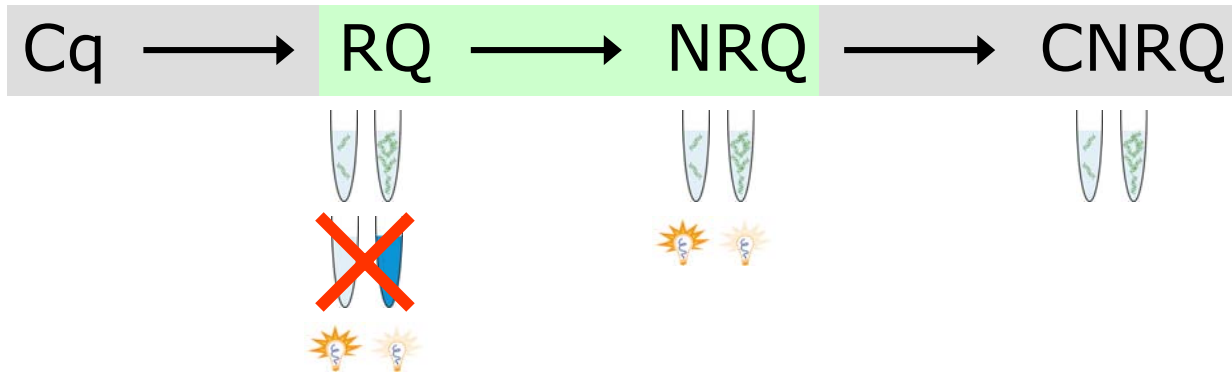


$$RQ = E^{\Delta Cq} \quad SE(RQ) = \sqrt{RQ^2 \left[\left(\frac{\Delta Cq \cdot SD(E)}{E} \right)^2 + (\ln(E) \cdot SD(Cq_{ref}))^2 \right]}$$

$$\Delta Cq = Cq_{ref} - Cq_{soi}$$

- Choice of Cq_{ref} does not affect the RQ ratio between samples
- Choice of Cq_{ref} does affect the error on RQ (if $SE(E)$ is taken into account)
- Minimize error: $Cq_{ref} = \text{average } Cq$

Calculate the correct error on RQ
without overestimating it



$$NRQ = \frac{RQ_{goi}}{RQ_{ref}}$$

$$NRQ = \frac{RQ_{goi}}{\sqrt[n]{\prod_x RQ_{ref,x}}}$$

Multiple reference genes
→ More accurate results

Multiple reference genes

→ quality control on reference gene stability

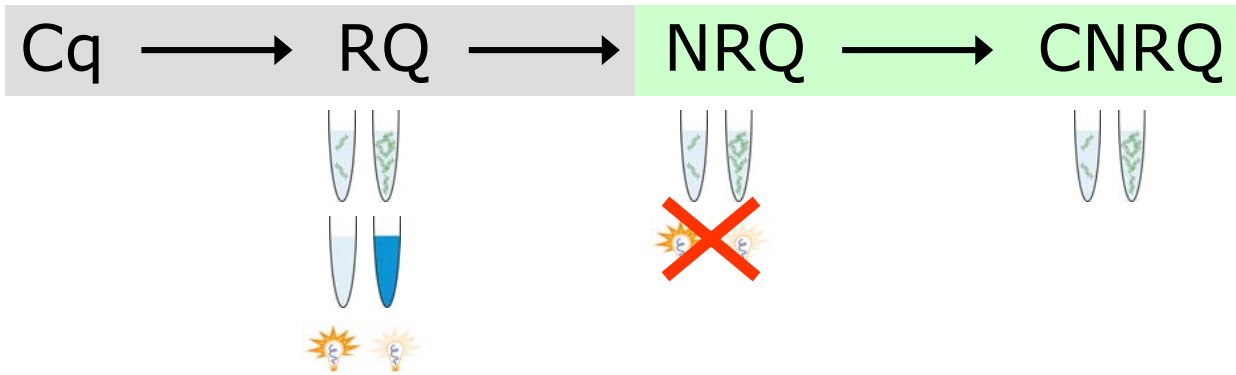
- coefficient of variation of NRQ_{ref}
 - 1 reference gene → $NRQ_{ref}=1$
 - multiple reference genes → NRQ_{ref} fluctuate around 1
 - $CV = \frac{SE(\overline{NRQ})}{\overline{NRQ}}$
- geNorm M value

Detect & exclude unstable reference genes
→ prevent skewing of results

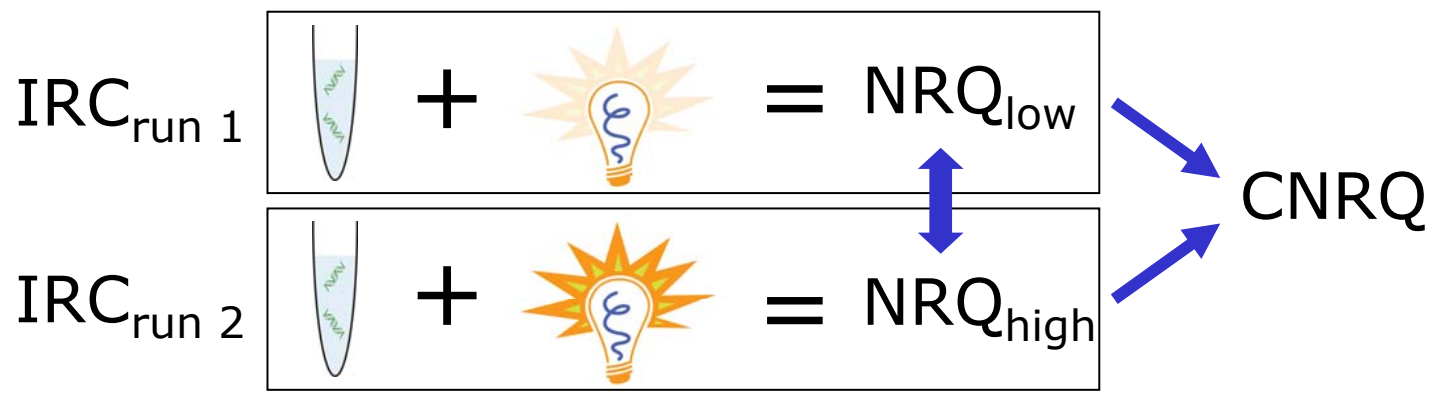
tissue type	gene	CV (%)	M	mean CV%	mean M
neuroblastoma	<i>UBC</i>	31.84	0.740	30.89	0.703
	<i>SDHA</i>	27.40	0.660		
	<i>HPRT1</i>	37.11	0.736		
	<i>GAPDH</i>	27.21	0.675		
fibroblast	<i>YHWAZ</i>	18.19	0.408	14.81	0.365
	<i>HPRT1</i>	8.84	0.308		
	<i>GAPDH</i>	17.40	0.378		
leukocyte	<i>B2M</i>	15.76	0.400	15.81	0.394
	<i>UBC</i>	15.79	0.389		
	<i>YWHAZ</i>	15.89	0.393		
bone marrow	<i>YWHAZ</i>	17.77	0.383	15.47	0.372
	<i>UBC</i>	13.60	0.356		
	<i>RPL13A</i>	15.03	0.376		
normal pool	<i>TBP</i>	47.51	1.099	43.73	0.925
	<i>HPRT1</i>	46.99	0.988		
	<i>HMBS</i>	31.16	0.849		
	<i>SDHA</i>	49.50	0.869		
	<i>GAPDH</i>	43.50	0.819		

Sample panel

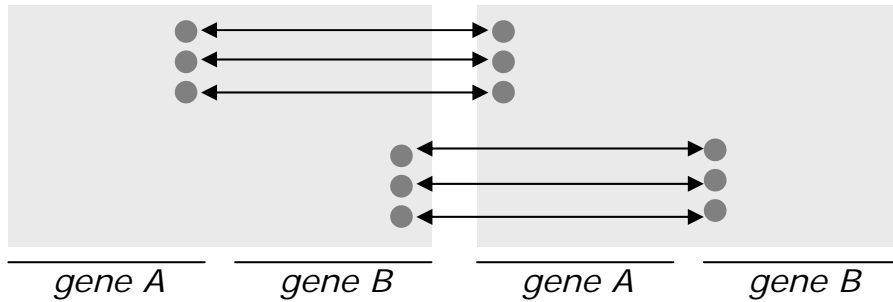
- Homogenous
 - CV: 25%
 - M: 0.5
- Heterogenous
 - CV: 50%
 - M: 1.0



- Inter-run differences are generally ignored
- Inter-run differences can be corrected for with IRCs
inter-run calibrator = identical sample measured for the same gene in different runs



- Inter-run calibration needs to be performed for each gene separately



- Inter-run calibration on the level of

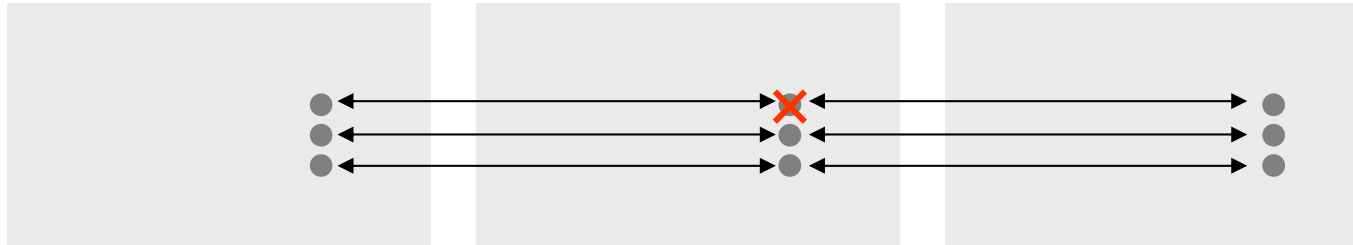
- Cq values ()

- NRQ values ()

➔ Enables comparison with an IRC from a fresh cDNA synthesis

$$CNRQ = \frac{NRQ_{soi}}{\sqrt[n]{\prod_x NRQ_{irc,x}}}$$

- Limitation: same set of IRCs required in all runs



$$\sqrt[3]{\prod_x^3 NRQ_{irc,x}}$$

~~$$\sqrt[3]{\prod_x^3 NRQ_{irc,x}}$$~~

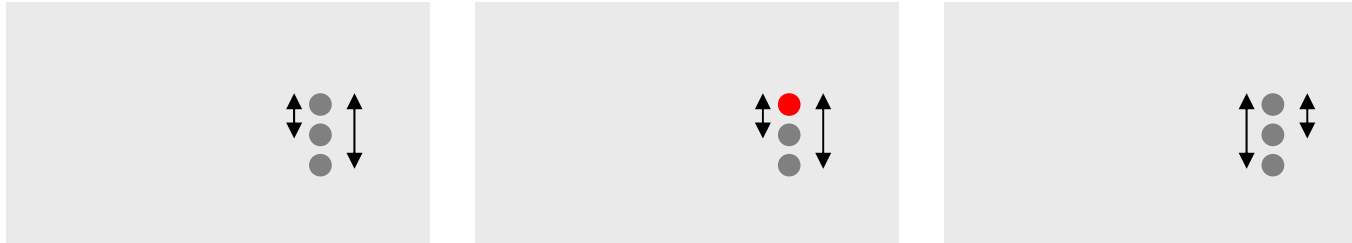
$$\sqrt[3]{\prod_x^3 NRQ_{irc,x}}$$

$$\sqrt[2]{\prod_x^2 NRQ_{irc,x}}$$

Solutions:

- No calibration for that gene in that run
- Exclude this IRC in all runs
- Complex inter-run calibration

Complex inter-run calibration: based on multiple imputation



1. Same NRQ ratio between identical samples in different runs
2. Use this ratio to impute missing data (in several ways)
3. Calculate calibration factor (multiple times)
4. Average all instances of the calibration factor

Multiple inter-run calibrators

- more accurate results
- less increase in error
- allows quality control

$$s_{e,jl} = \sqrt{\frac{\sum_{i=1}^h (Cq_{qil,measured} - Cq_{qil,predicted})^2}{h-2}} \quad (\text{formula 2})$$

$$s_{x,jl} = \sqrt{\frac{1}{h-1} \sum_{q=1}^h (Q_{qjl} - \bar{Q}_{jl})^2} \quad (\text{formula 3})$$

$$SE(\text{slope}_{jl}) = \frac{s_{e,jl}}{s_{x,jl}(h-1)} \quad (\text{formula 4})$$

The base for exponential amplification E , and its standard error $SE(E)$ are calculated from these values:

$$E_{jl} = 10^{\left(\frac{1}{\text{slope}_{jl}}\right)} \quad (\text{formula 5})$$

$$SE(E_{jl}) = \frac{E_{jl} \cdot \ln(10) \cdot SE(\text{slope}_{jl})}{\text{slope}_{jl}^2} \quad (\text{formula 6})$$

Conversion of Cq values into relative quantities

Step 1

Calculation of the average Cq value for all replicates of the same gene/sample combination jk within a given run l :

$$\bar{Cq}_{jkl} = \frac{\sum_{i=1}^n Cq_{ijkil}}{n} \quad (\text{formula 7})$$

$$SE(Cq_{jkl}) = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (Cq_{ijkil} - \bar{Cq}_{jkl})^2} \quad (\text{formula 8})$$

Step 2

Transformation of mean Cq value into RQ using the gene specific PCR efficiency E_{jl} , with minimization of the overall error:

$$Cq_{reference,jl} = \bar{Cq}_{jl} = \frac{\sum_{k=1}^s Cq_{jkl}}{s} \quad (\text{formula 9})$$

$$\Delta Cq_{jkl} = Cq_{reference,jl} - Cq_{jkl} \quad (\text{formula 10})$$

$$RQ_{jkl} = E_{jl}^{\Delta Cq_{jkl}} \quad (\text{formula 11})$$

$$SE(RQ_{jkl}) = \sqrt{RQ_{jkl}^2 \left[\left(\frac{\Delta Cq_{jkl}}{E_{jl}} \cdot SD(E_{jl}) \right)^2 + (1 + \ln(E_{jl}))^2 \cdot SD(Cq_{jkl})^2 \right]} \quad (\text{formula 12})$$

Normalization: inter-run calibration

The procedures for normalization and inter-run calibration are highly analogous and are therefore described in parallel.

Step 1

Calculation of the normalization factor NF for sample k based on the RQs of the reference genes p .

Step 1'

Calculation of the calibration factor CF for gene j in run l based on the NRQs of the IRCs m :

$$NF_k = \sqrt{\prod_{p=1}^f RQ_{pk}} \quad (\text{formula 13})$$

$$CF_{jl} = \sqrt{\prod_{m=1}^c NRQ_{jlm}} \quad (\text{formula 13'; for definition of NRQ, see formula 15})$$

$$SE(NF_k) = NF_k \sqrt{\sum_{p=1}^f \left(\frac{SE(RQ_{pk})}{f \cdot RQ_{pk}} \right)^2} \quad (\text{formula 14})$$

$$SE(CF_{jl}) = CF_{jl} \sqrt{\sum_{m=1}^c \left(\frac{SE(NRQ_{jlm})}{c \cdot NRQ_{jlm}} \right)^2} \quad (\text{formula 14'})$$

Step 2

Conversion of RQs into NRQs.

Step 2'

Conversion of NRQs into CNRQs:

$$NRQ_{jk} = \frac{RQ_{jk}}{NF_k} \quad (\text{formula 15})$$

$$CNRQ_{jkl} = \frac{NRQ_{jkl}}{CF_{jl}} \quad (\text{formula 15'})$$

$$SE(NRQ_{jk}) = NRQ_{jk} \sqrt{\left(\frac{SE(NF_k)}{NF_k} \right)^2 + \left(\frac{SE(RQ_{jk})}{RQ_{jk}} \right)^2} \quad (\text{formula 16})$$

$$SE(CNRQ_{jkl}) = CNRQ_{jkl} \sqrt{\left(\frac{SE(CF_{jl})}{CF_{jl}} \right)^2 + \left(\frac{SE(NRQ_{jkl})}{NRQ_{jkl}} \right)^2} \quad (\text{formula 16'})$$

Coefficient of variation of NRQs of a reference gene

Step 1

Calculation of the mean NRQ for all samples k and a given reference gene p :

$$\bar{NRQ}_p = \frac{\sum_{k=1}^s NRQ_{pk}}{s} \quad (\text{formula 17})$$

$$s_{e,jl} = \sqrt{\frac{\sum_{i=1}^h (Cq_{qil,measured} - Cq_{qil,predicted})^2}{h-2}} \quad (\text{formula 2})$$

$$s_{x,jl} = \sqrt{\frac{1}{h-1} \sum_{q=1}^h (Q_{qjl} - \bar{Q}_{jl})^2} \quad (\text{formula 3})$$

$$SE(\text{slope}_{jl}) = \frac{s_{e,jl}}{s_{x,jl}(h-1)} \quad (\text{formula 4})$$

The base for exponential amplification E , and its standard error $SE(E)$ are calculated from these values:

$$E_{jl} = 10^{\left(\frac{1}{\text{slope}_{jl}}\right)} \quad (\text{formula 5})$$

$$SE(E_{jl}) = \frac{E_{jl} \cdot \ln(10) \cdot SE(\text{slope}_{jl})}{\text{slope}_{jl}^2} \quad (\text{formula 6})$$

Conversion of Cq values into relative quantities

Step 1

Calculation of the average Cq value for all replicates of the same gene/sample combination jk within a given run l :

$$\bar{Cq}_{jkl} = \frac{\sum_{i=1}^n Cq_{ijkil}}{n} \quad (\text{formula 7})$$

$$SE(Cq_{jkl}) = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (Cq_{ijkil} - \bar{Cq}_{jkl})^2} \quad (\text{formula 8})$$

Step 2

Transformation of mean Cq value into RQ using the gene specific PCR efficiency E_{jl} , with minimization of the overall error:

$$Cq_{reference,jl} = \bar{Cq}_{jl} = \frac{\sum_{k=1}^s Cq_{jkl}}{s} \quad (\text{formula 9})$$

$$\Delta Cq_{jkl} = Cq_{reference,jl} - Cq_{jkl} \quad (\text{formula 10})$$

$$RQ_{jkl} = E_{jl}^{\Delta Cq_{jkl}} \quad (\text{formula 11})$$

$$SE(RQ_{jkl}) = \sqrt{RQ_{jkl}^2 \left[\left(\frac{\Delta Cq_{jkl}}{E_{jl}} \cdot SD(E_{jl}) \right)^2 + (1 + \ln(E_{jl}))^2 \cdot SD(Cq_{jkl})^2 \right]} \quad (\text{formula 12})$$

Normalization: inter-run calibration

The procedures for normalization and inter-run calibration are highly analogous and are therefore described in parallel.

Step 1

Calculation of the normalization factor NF for sample k based on the RQs of the reference genes p .

Step 1'

Calculation of the calibration factor CF for gene j in run l based on the NRQs of the IRCs m :

$$NF_k = \sqrt{\prod_{p=1}^f RQ_{pk}} \quad (\text{formula 13})$$

$$CF_{jl} = \sqrt{\prod_{m=1}^c NRQ_{jlm}} \quad (\text{formula 13'; for definition of NRQ, see formula 15})$$

$$SE(NF_k) = NF_k \sqrt{\sum_{p=1}^f \left(\frac{SE(RQ_{pk})}{f \cdot RQ_{pk}} \right)^2} \quad (\text{formula 14})$$

$$SE(CF_{jl}) = CF_{jl} \sqrt{\sum_{m=1}^c \left(\frac{SE(NRQ_{jlm})}{c \cdot NRQ_{jlm}} \right)^2} \quad (\text{formula 14'})$$

Step 2

Conversion of RQs into NRQs.

Step 2'

Conversion of NRQs into CNRQs:

$$NRQ_{jk} = \frac{RQ_{jk}}{NF_k} \quad (\text{formula 15})$$

$$CNRQ_{jkl} = \frac{NRQ_{jkl}}{CF_{jl}} \quad (\text{formula 15'})$$

$$SE(NRQ_{jk}) = NRQ_{jk} \sqrt{\left(\frac{SE(NF_k)}{NF_k} \right)^2 + \left(\frac{SE(RQ_{jk})}{RQ_{jk}} \right)^2} \quad (\text{formula 16})$$

$$SE(CNRQ_{jkl}) = CNRQ_{jkl} \sqrt{\left(\frac{SE(CF_{jl})}{CF_{jl}} \right)^2 + \left(\frac{SE(NRQ_{jkl})}{NRQ_{jkl}} \right)^2} \quad (\text{formula 16'})$$

Coefficient of variation of NRQs of a reference gene

Step 1

Calculation of the mean NRQ for all samples k and a given reference gene p :

$$\bar{NRQ}_p = \frac{\sum_{k=1}^s NRQ_{pk}}{s} \quad (\text{formula 17})$$

Method

qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data

Jan Hellemans, Geert Mortier, Anne De Paepe, Frank Speleman and Jo Vandesompele

Address: Center for Medical Genetics, Ghent University Hospital, De Pintelaan, B-9000 Ghent, Belgium.

Correspondence: Jo Vandesompele. Email: Joke.Vandesompele@UGent.be

Published: 9 February 2007

Genome Biology 2007, 8:R19 (doi:10.1186/gb-2007-8-2-r19)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2007/8/2/R19>

Received: 31 August 2006

Revised: 7 December 2006

Accepted: 9 February 2007

© 2007 Hellemans et al.; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

qBase

- Implements most discussed methods
- Interest demonstrated by downloads
 - >2700
 - >70 countries
- Has a number of limitations
- → professional qBasePlus

- Advantages of qBasePlus (Biogazelle)
 - Platform independent (Windows, Mac, Linux, ...)
 - Appealing graphical interface with calculations on the fly
 - Direct import from instrument software export files
 - Data exchange and online sharing
 - Error estimation on PCR efficiency
 - Quality control
 - Raw data
 - Normalization
 - Inter-run calibration
 - Inter-run calibration for any experimental setup
 - Biostatistical analysis

Project Explorer Open ▾

- ⊕ My Project nr 0
- ⊕ My Project nr 2
- ⊖ My Project nr 1
 - Annotations
 - ⊕ Default experiment settings
 - ⊖ Experiments
 - ⊖ Experiment nr 2
 - Annotations
 - ⊖ Settings
 - Calculation parameters
 - Quality control settings
 - ⊕ Targets
 - ⊕ Samples
 - ⊕ Runs
 - ⊕ Quality control
 - ⊕ Intermediate results
 - ⊖ Analysis
 - CNRQ matrix
 - Correlations
 - Heat map
 - Multi-gene
 - Conclusions
 - ⊕ Experiment nr 0
 - ⊕ Experiment nr 3
 - ⊕ Experiment nr 1
 - Project Conclusions

Experiment Calculation Parameters ⊗

Calculation Pathway

a - c - d

Amplification Efficiency Types

One default amplification efficiency for all targets

Target specific amplification efficiencies

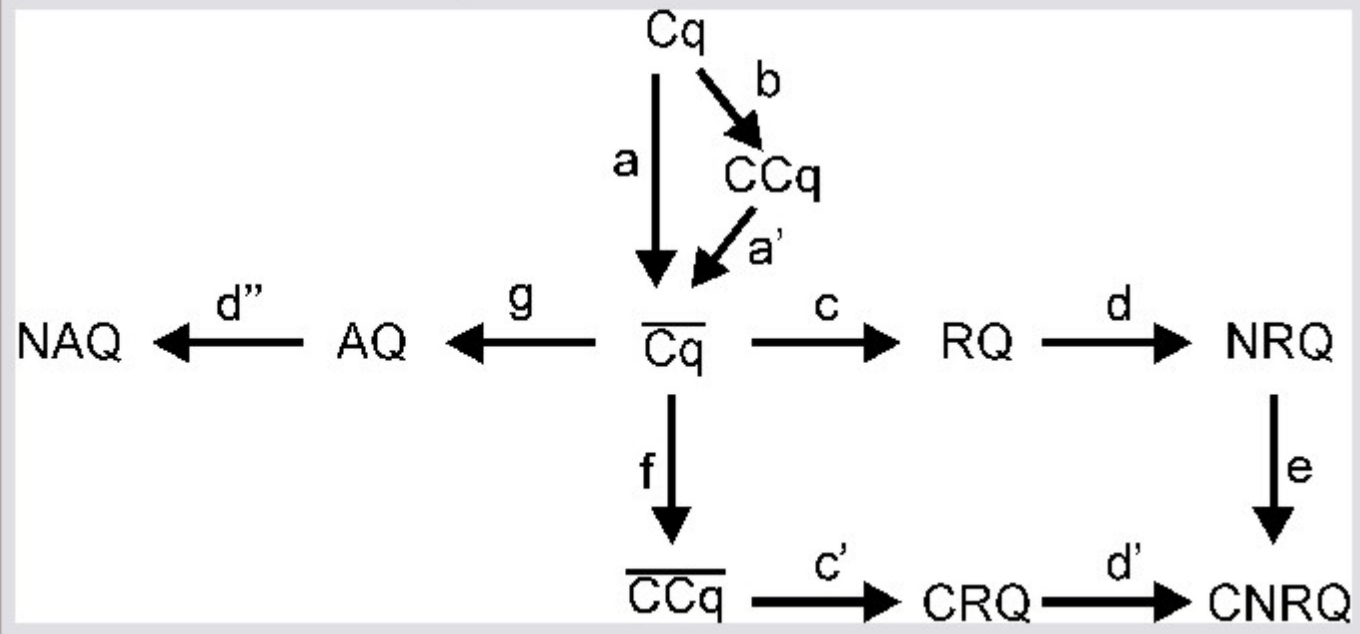
Target and run specific amplification efficiencies

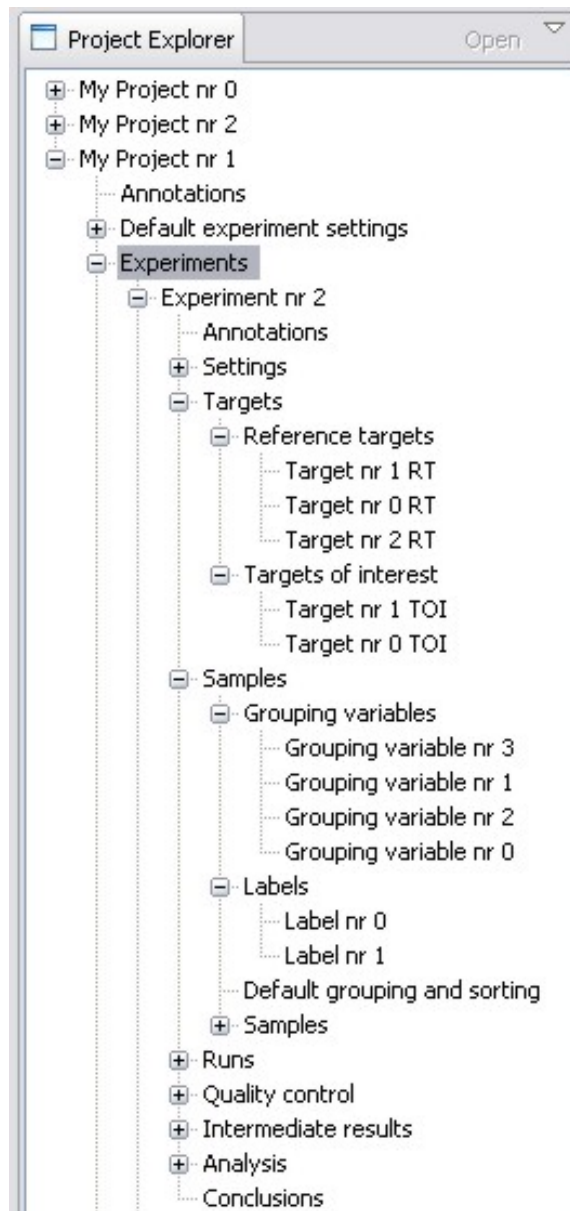
Sample, target and run specific amplification efficiencies

Well specific amplification efficiencies

Number of PCR Replicates

Number of Wells :





Project Explorer Open

- My Project nr 0
- My Project nr 2
- My Project nr 1
 - Annotations
 - Default experiment settings
 - Experiments
 - Experiment nr 2
 - Annotations
 - Settings
 - Calculation parameters
 - Quality control settings
 - Targets
 - Reference targets
 - Target nr 1 RT
 - Target nr 0 RT
 - Target nr 2 RT
 - Targets of interest
 - Samples
 - Runs
 - Quality control
 - Intermediate results
 - Analysis
 - CNRQ matrix
 - Correlations
 - Heat map
 - Multi-gene
 - Conclusions

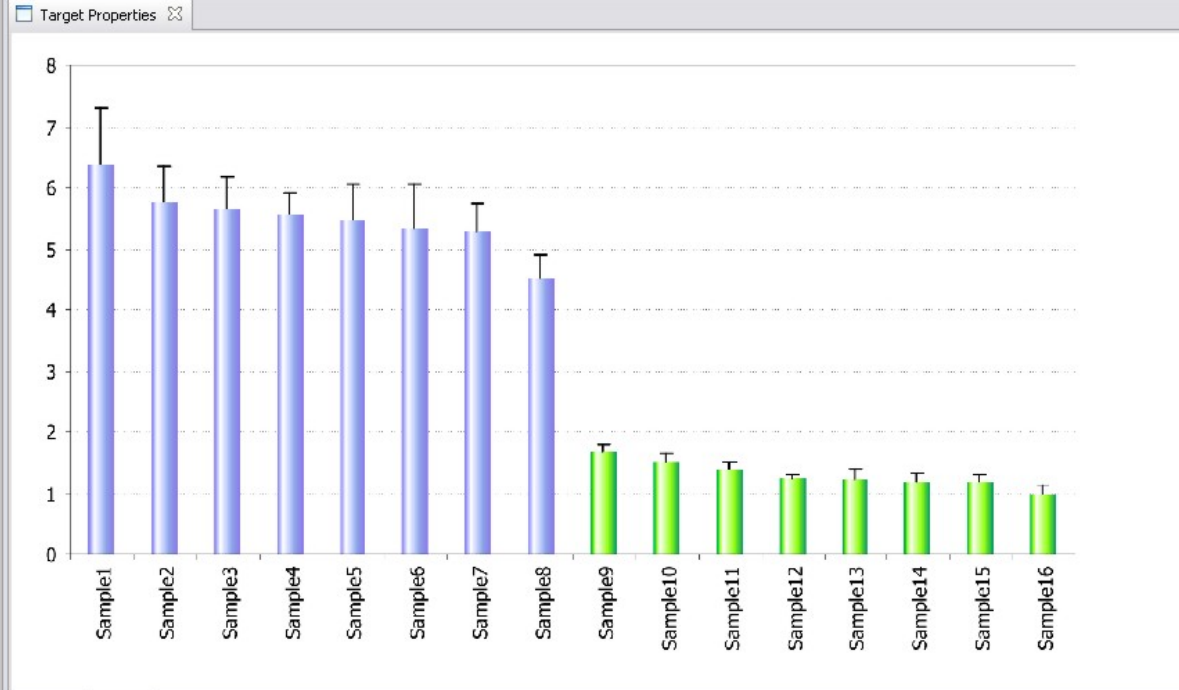
- Experiment nr 0
- Experiment nr 3
- Experiment nr 1
- Project Conclusions

Experiment Annotations

Experiment Name :

E :

SE (E) :



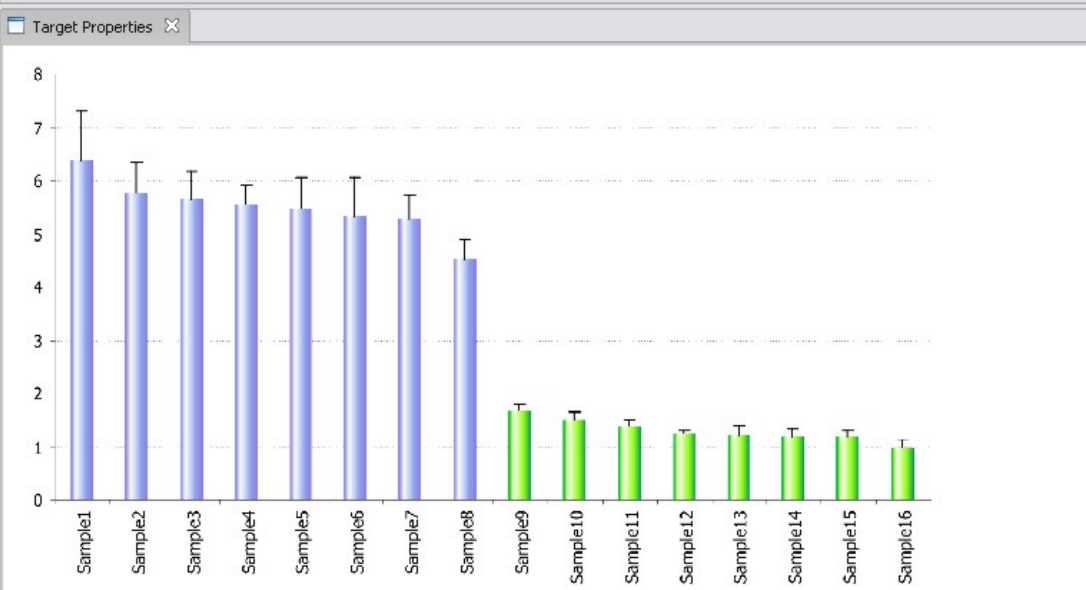
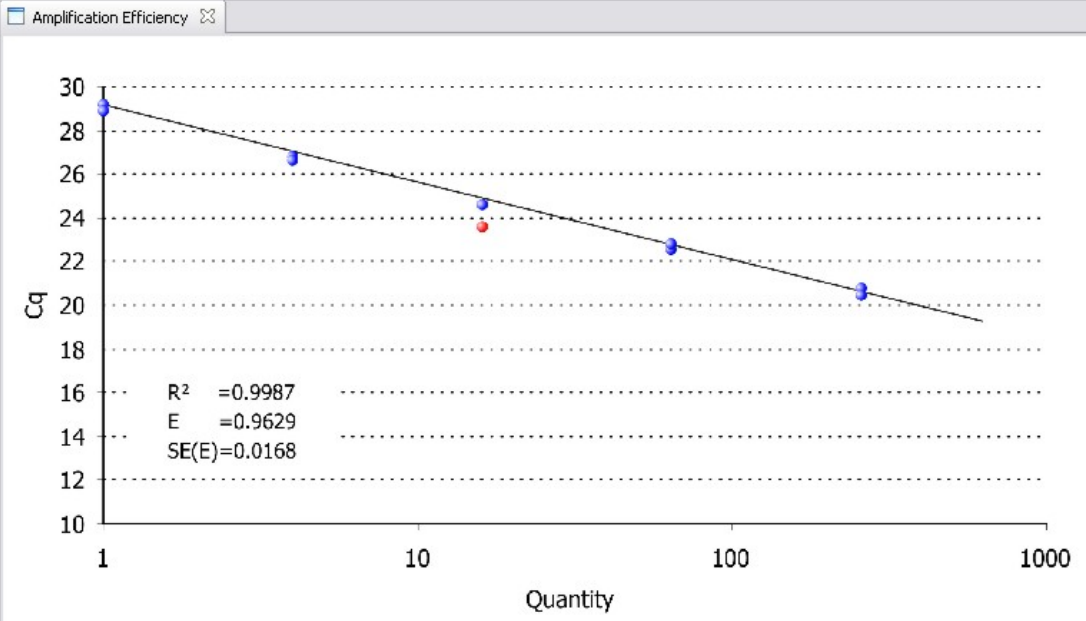
Properties Results

Run

	Target nr 1 RT	Target nr 0 RT	Target nr 2 RT	Target nr 1 TOI	Target nr 0 TOI
Sample nr 2	0,927 ± 0,312	1,572 ± 1,346	0,686 ± 0,456	0,675 ± 0,896	1,154 ± 0,383
Sample nr 5	1,379 ± 0,91	1,098 ± 0,872	0,661 ± 0,759	0,682 ± 0,793	1,324 ± 0,748
Sample nr 4	0,725 ± 0,213	0,911 ± 0,83	1,513 ± 0,471	1,466 ± 0,453	0,755 ± 0,256
Sample nr 3	1,079 ± 0,905	0,636 ± 0,295	1,457 ± 1,573	1,481 ± 0,788	0,867 ± 0,658
Sample nr 1	0,996 ± 0,403	0,619 ± 0,456	1,623 ± 0,935	1,362 ± 1,125	0,721 ± 0,389
Sample nr 0	1,004 ± 1,038	1,616 ± 0,7	0,616 ± 0,401	0,734 ± 0,577	1,387 ± 0,539

Project Explorer Open

- ⊕ My Project nr 0
- ⊕ My Project nr 2
- ⊖ My Project nr 1
 - Annotations
 - ⊕ Default experiment settings
 - ⊖ Experiments
 - ⊖ Experiment nr 2
 - Annotations
 - ⊖ Settings
 - Calculation parameters
 - Quality control settings
 - ⊖ Targets
 - ⊖ Reference targets
 - Target nr 1 RT
 - Target nr 0 RT
 - Target nr 2 RT
 - ⊕ Targets of interest
 - ⊕ Samples
 - ⊕ Runs
 - ⊕ Quality control
 - ⊖ Intermediate results
 - Amplification efficiency
 - Normalization
 - Interrun calibration
 - ⊖ Analysis
 - CNRQ matrix
 - Correlations
 - Heat map
 - Multi-gene
 - Conclusions
 - ⊕ Experiment nr 0
 - ⊕ Experiment nr 3
 - ⊕ Experiment nr 1
 - Project Conclusions



Properties Results

Acknowledgement



Jo Vandesompele



Kristel Van Steen
Peter Konings



info@biogazelle.com

