Estimation of sample specific efficiency

methods and applications

Ales Tichopad

Tzachi Bar



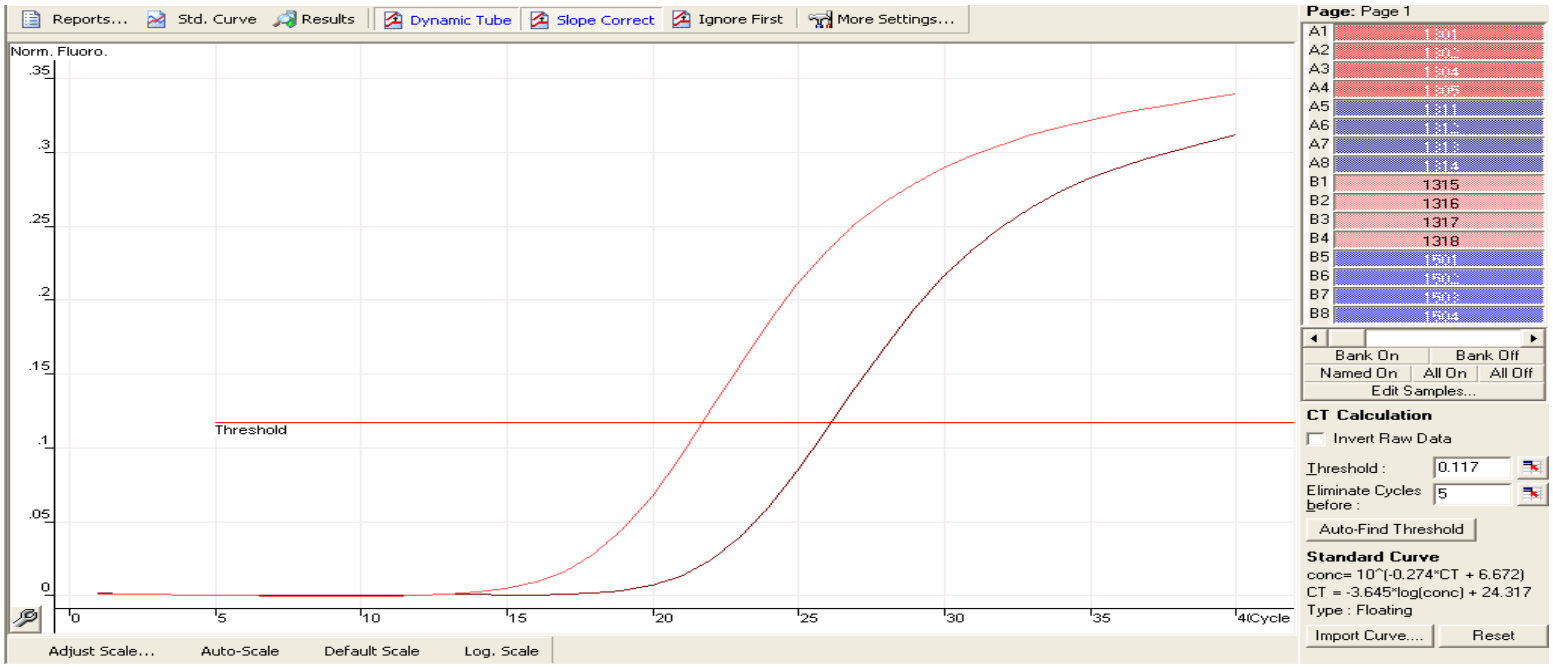
LabonNet

Quality Assurance for real – time PCR

Outline

- The amplification efficiency
- Why to estimate amplification efficiencies?
- Amplification efficiency estimation
- Comparison of different efficiency estimation methods
- Applications
- Conclusion

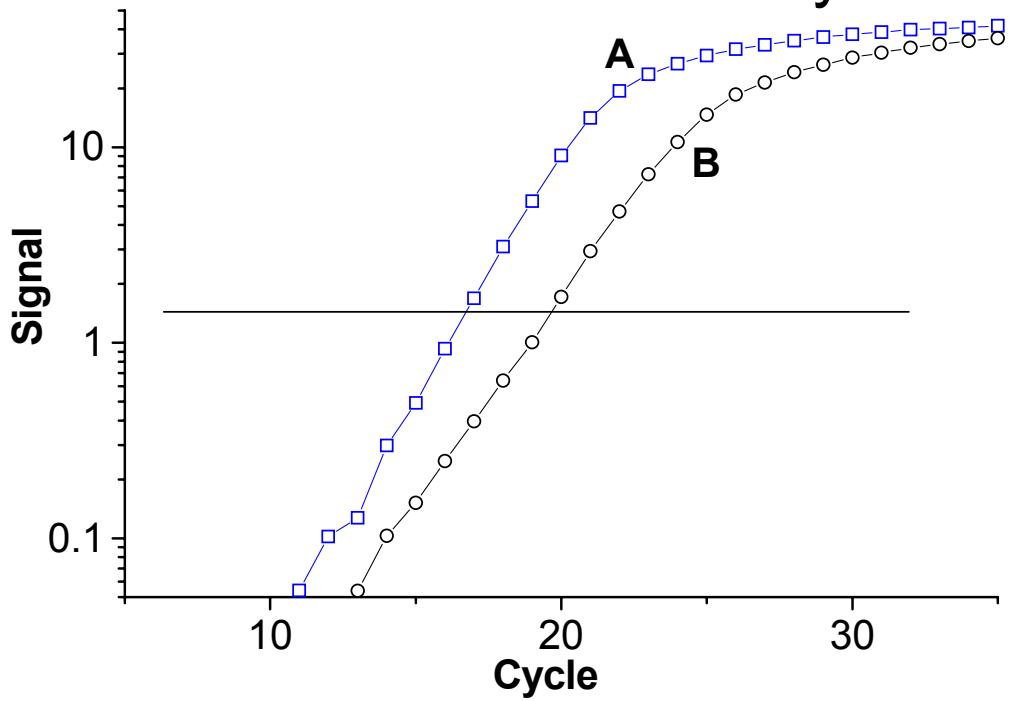
PCR efficiencies of the same gene in compared samples must be similar



What if the efficiencies are **NOT** similar?

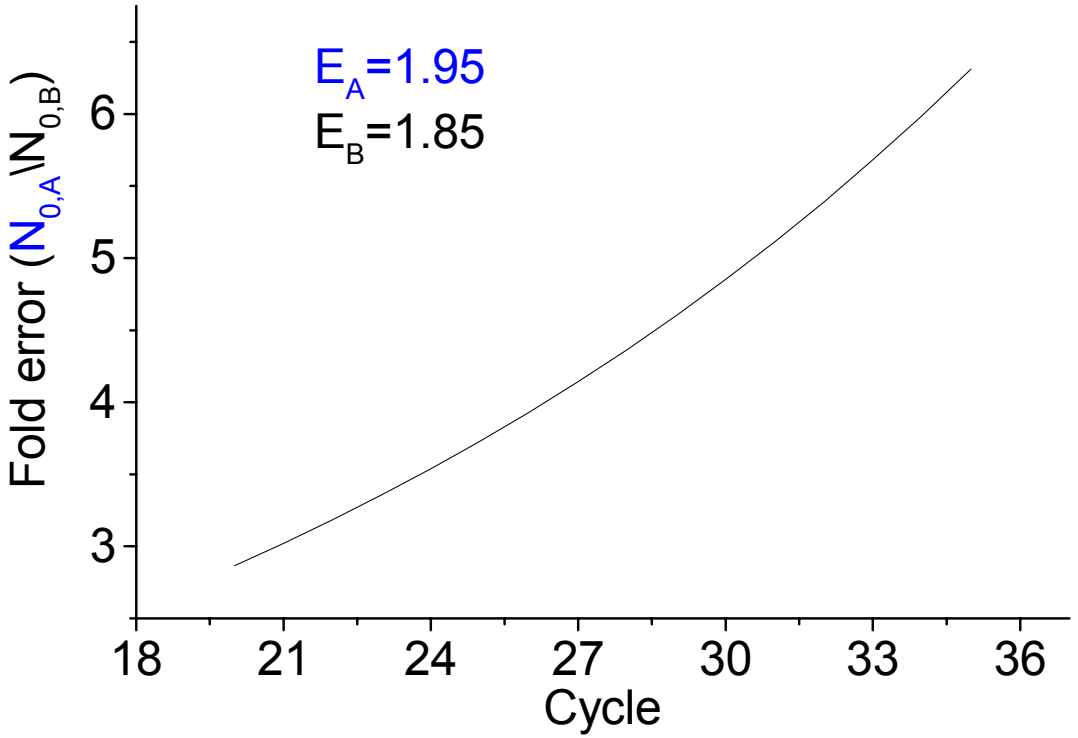
Equal initial copy number

Different PCR efficiency



Estimation of error due to efficiency dissimilarity

constant CT, different efficiencies



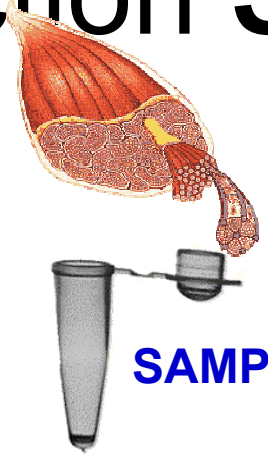
Do reference genes compensate for different efficiencies between samples?

- Not always
- Target and reference gene might have different amplification efficiencies $E_{\text{target gene}} \neq E_{\text{ref. gene}}$
- Difference in E between target and reference gene (ΔE) doesn't stay constant but differs from sample to sample – interaction.

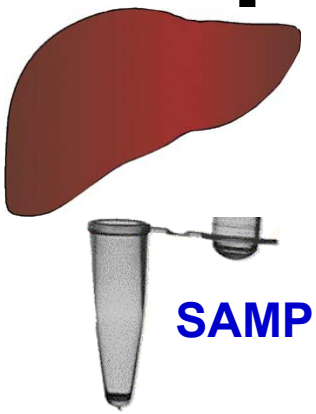
  $\Delta E = E_{\text{target gene}} - E_{\text{ref. Gene}}$

  $\Delta E_{\text{sample 1}} \neq \Delta E_{\text{sample 2}}$
- i.e. **Interaction** *Sample*Sequence*

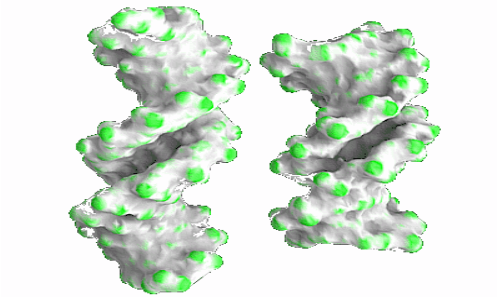
Interaction **Sample*Sequence**



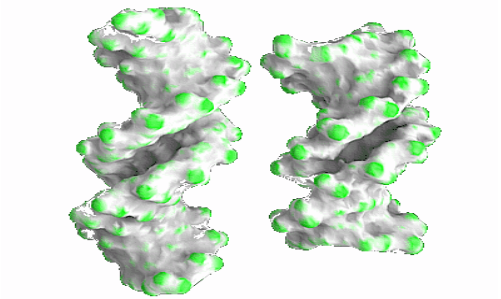
SAMPLE 1



SAMPLE 2



Target **Reference**



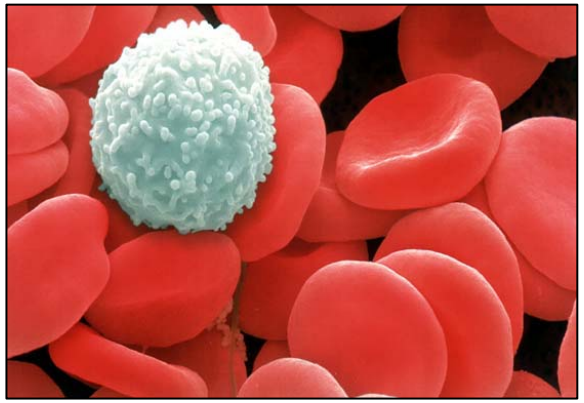
Target **Reference**

n=10	n=10	n=10	n=10
E= 1.85 > E=1.79 **ΔE=0.06**		E=1.80 < E=1.81 **ΔE=-0.01**	

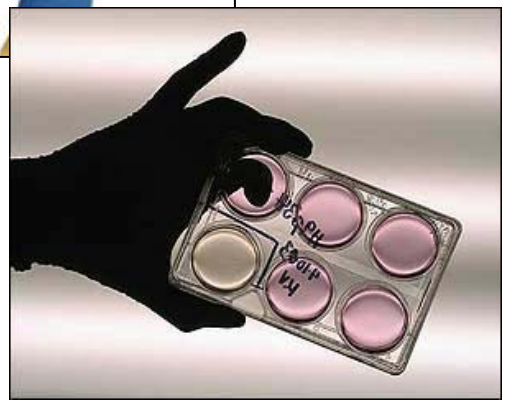


versus





versus





versus



APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Oct. 1997, p. 3741–3751
0099-2240/97/$04.00+0
Copyright © 1997, American Society for Microbiology
Vol. 63, No. 10

MINIREVIEW

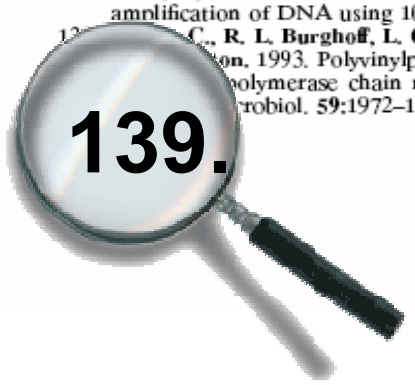
Inhibition and Facilitation of Nucleic Acid Amplification

IAN G. WILSON*

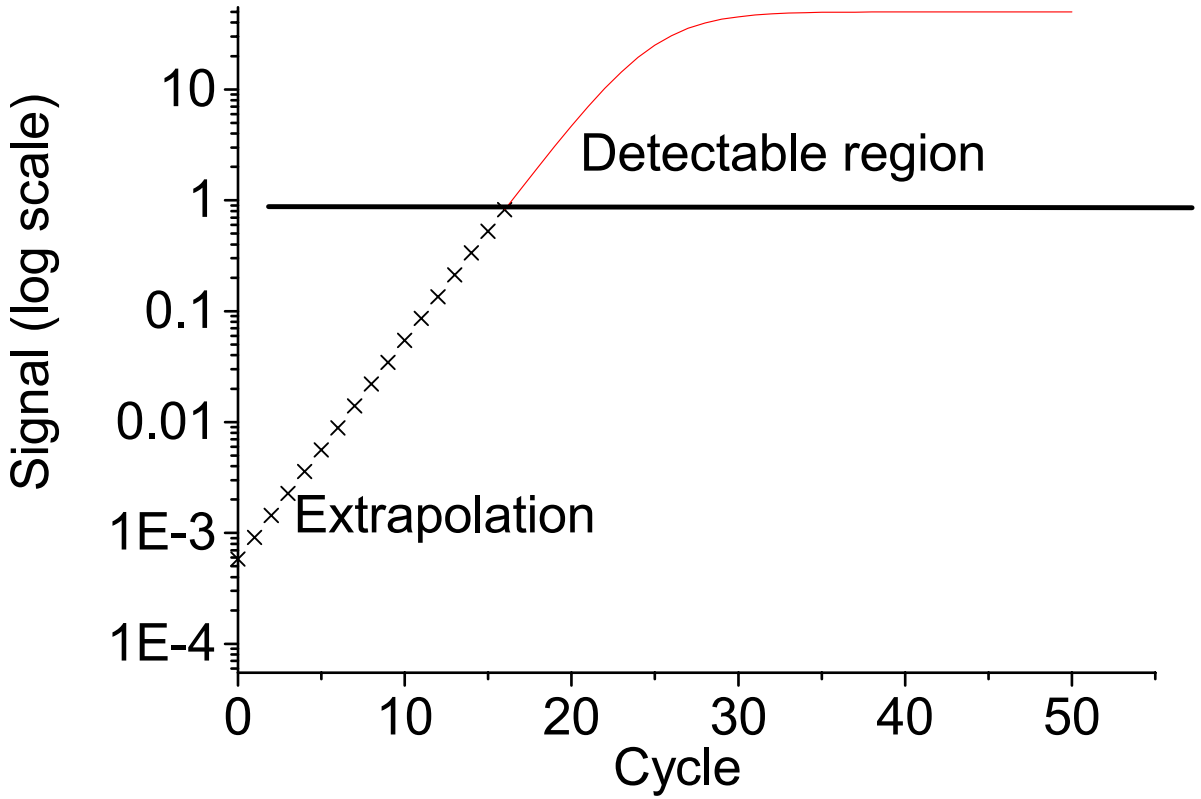
Northern Ireland Public Health Laboratory, Bacteriology Department, Belfast City Hospital, Belfast BT9 7AD, United Kingdom

Factors that inhibit the amplification of nucleic acids by PCR ... sensitivity, specificity, and reproducibility have been reported ...

Sussman (ed.), New techniques in food and beverage microbiology. Scientific Publishers, Blackwell Oxford, United Kingdom.
135. **Witham, P. K., C. T. Yamashiro, K. J. Livak, and C. A. Batt.** 1996. A PCR-based assay for the detection of *Escherichia coli* Shiga-like toxin genes in ground beef. Appl. Environ. Microbiol. **62**:1347–1353.
136. **Wu, P., S. Daniel-Issakani, K. LaMarco, and B. Strulovici.** 1997. An automated high throughput filtration assay: application to polymerase inhibitor identification. Anal. Biochem. **245**:226–230.
137. **Xia, J. Q., C. V. Yason, and F. S. B. Kibenge.** 1995. Comparison of dot blot hybridization, polymerase chain reaction, and virus isolation for detection of bovine herpesvirus-1 (BHV-1) in artificially infected bovine semen. Can. J. Vet. Res. **59**:102–109.
138. **Yoon, C., and D. A. Glawe.** 1993. Pretreatment with RNase to improve PCR amplification of DNA using 10-mer primers. BioTechniques **6**:908–910.
13[illegible] **C., R. L. Burghoff, L. G. Keim, V. Minak-Bernero, J. R. Lute, and [illegible]on.** 1993. Polyvinylpyrrolidone-agarose gel electrophoresis purification of polymerase chain reaction-amplifiable DNA from soils. Appl. [illegible]robiol. **59**:1972–1974.



The ideal solution: quantification by sample specific amplification



The signal from the detectable region is extrapolated and used to draw conclusions about the initial amount of DNA.

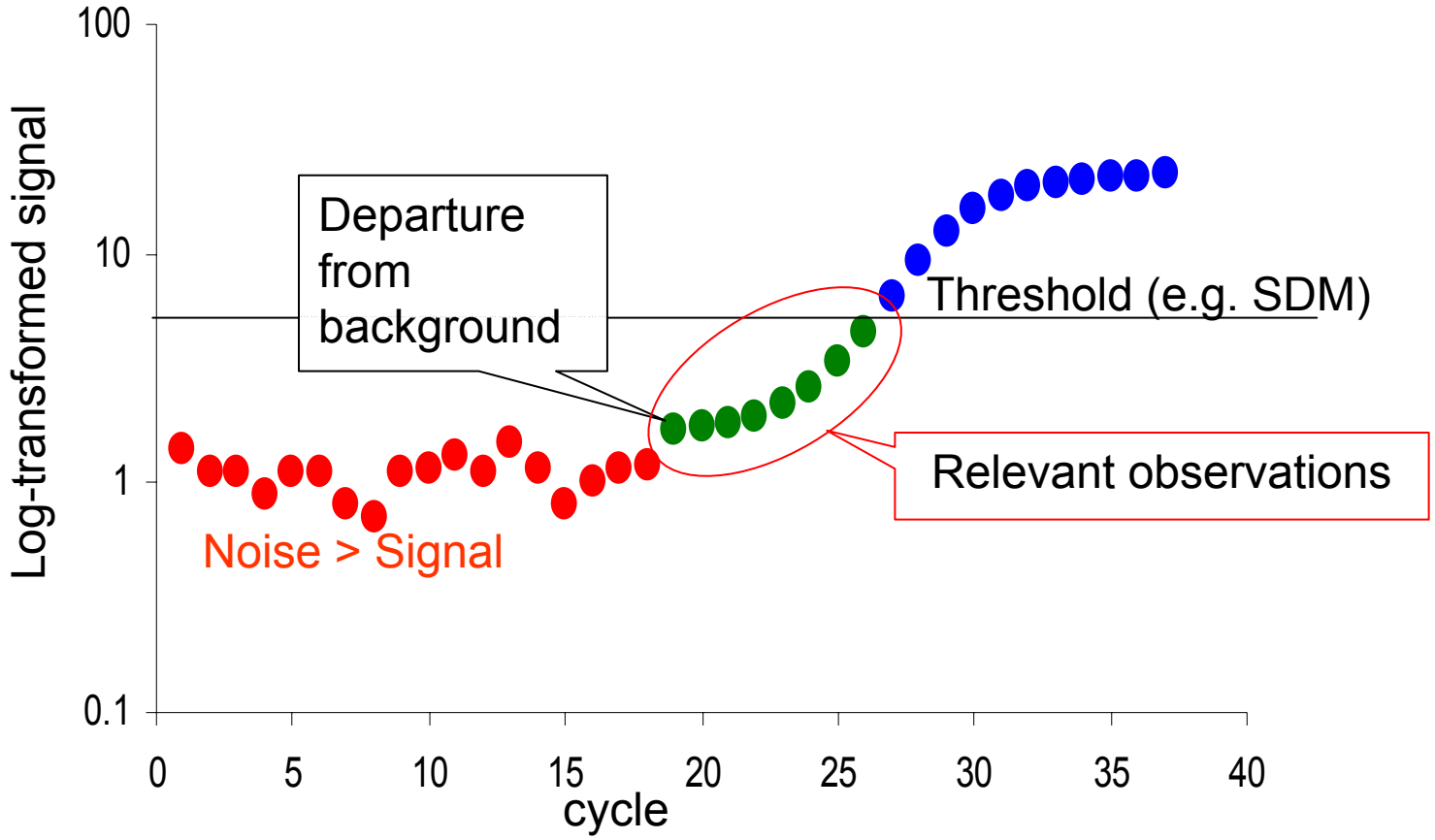
Efficiency estimation methods:

- Bar T & Muszta A 2005 BioTechinques 39: 333-340
- Liu W & Saint DA 2002a Anal Biochem 302
- Liu W & Saint DA 2002b Biochem Biophys Res Commun 294
- Marino JH et al. 2003 J Immunol Methods 283
- Peccoud J & Jacob C 1998 In Gene quantification (ed. F. Ferre)
- Peirson SN et al. 2003 Nucleic Acids Res 31
- Ramakers C et al. 2003 Neurosci Lett 339
- Rutledge RG 2004 Nucleic Acids Res 32
- Tichopad A et al. 2003 Nucleic Acids Res 31
- Wilhelm JA et al. 2003 Biotechniques 34

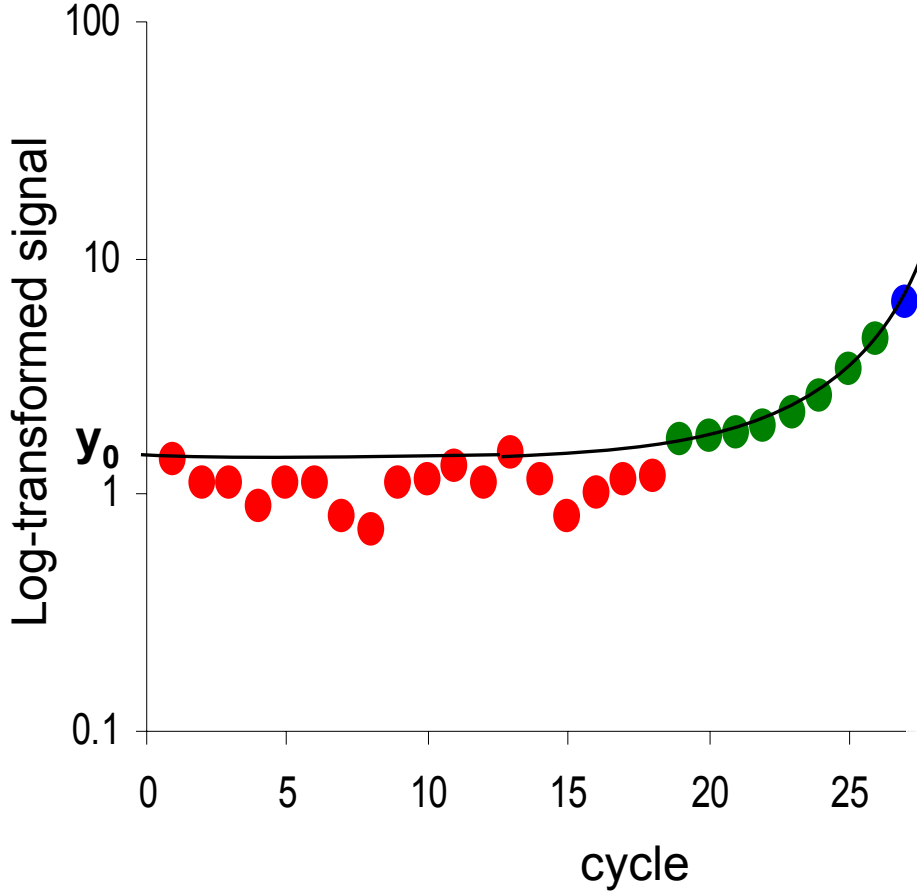
Where is the problem in efficiency estimation?

- Early signal is hidden in the noise
- Modeling the background signal
- Where on the amplification curve the efficiency should be estimated?
- Variable Signal / DNA proportion

Where to estimate efficiency by exponential model?



What model to fit



DNA amplification

$y = a * E^{cycle}$

Signal amplification

$\gamma = \gamma_0 + \alpha * \varepsilon^{cycle} + \delta$

γ – Signal

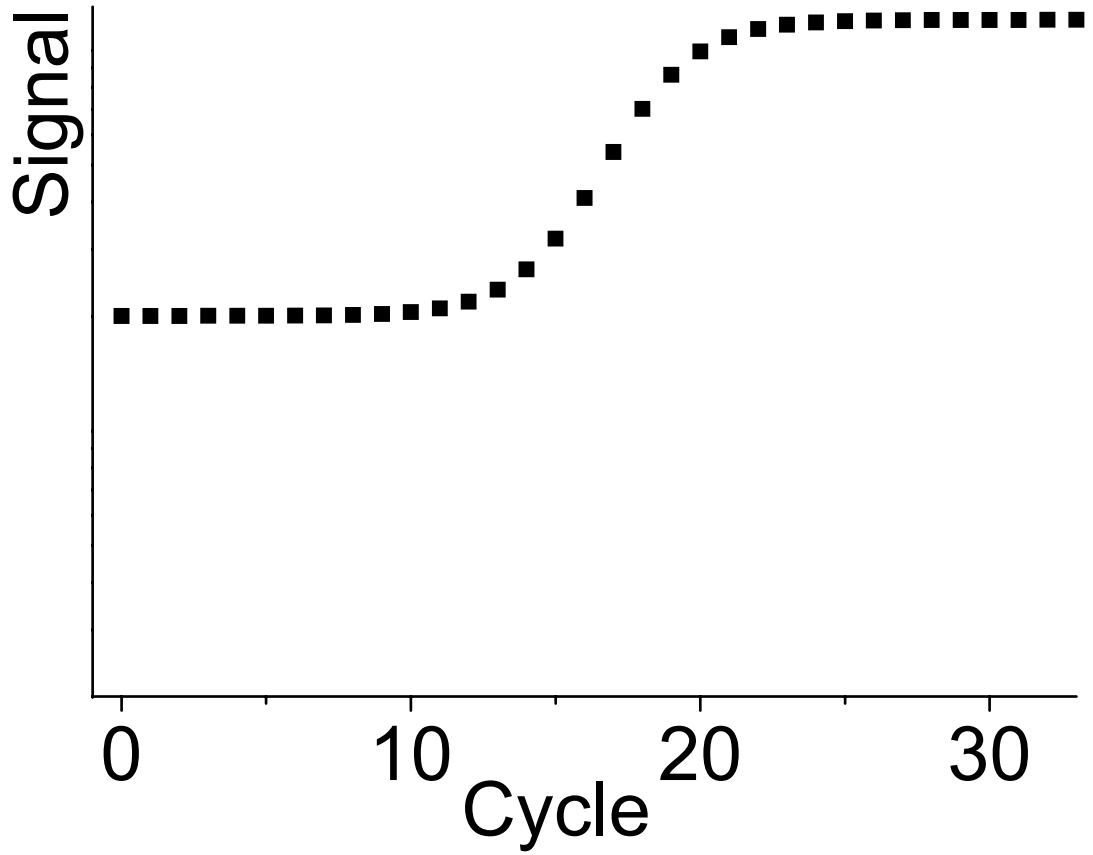
γ_0 – Background signal

α – Signal by DNA

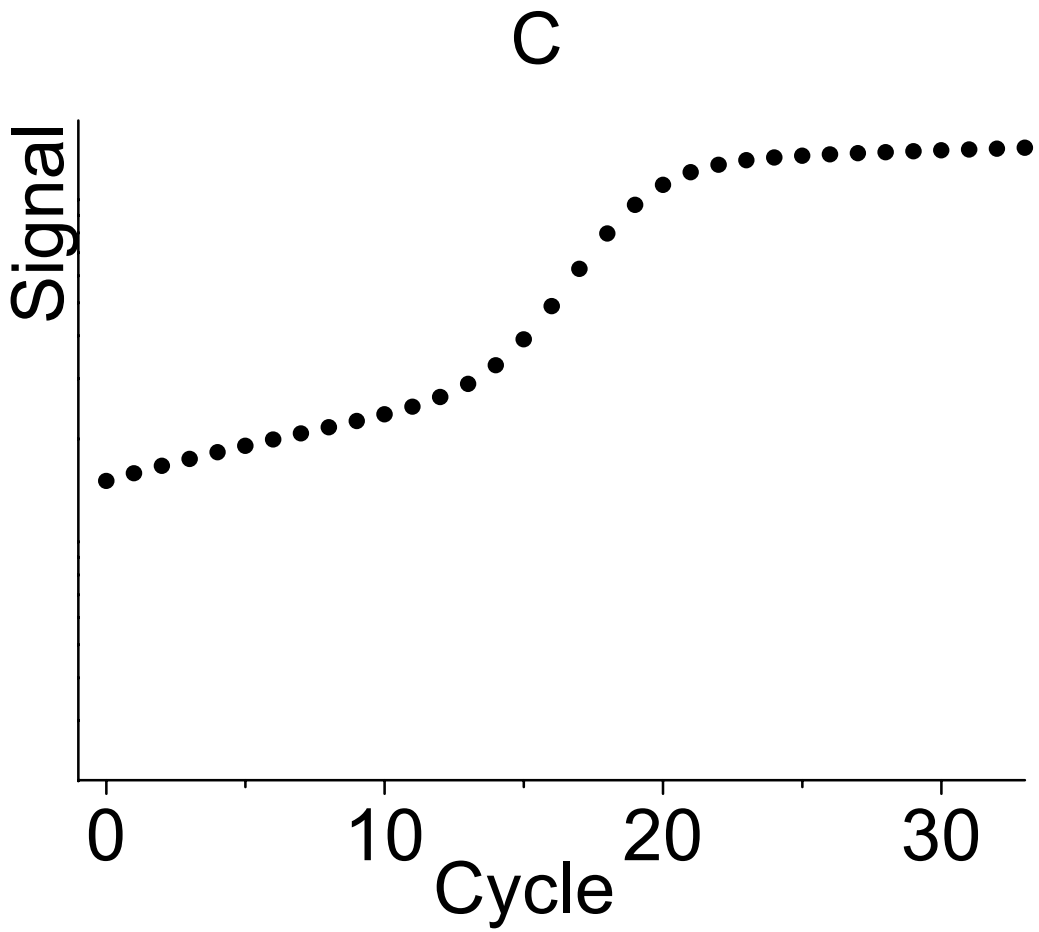
ε – Efficiency

δ - noise

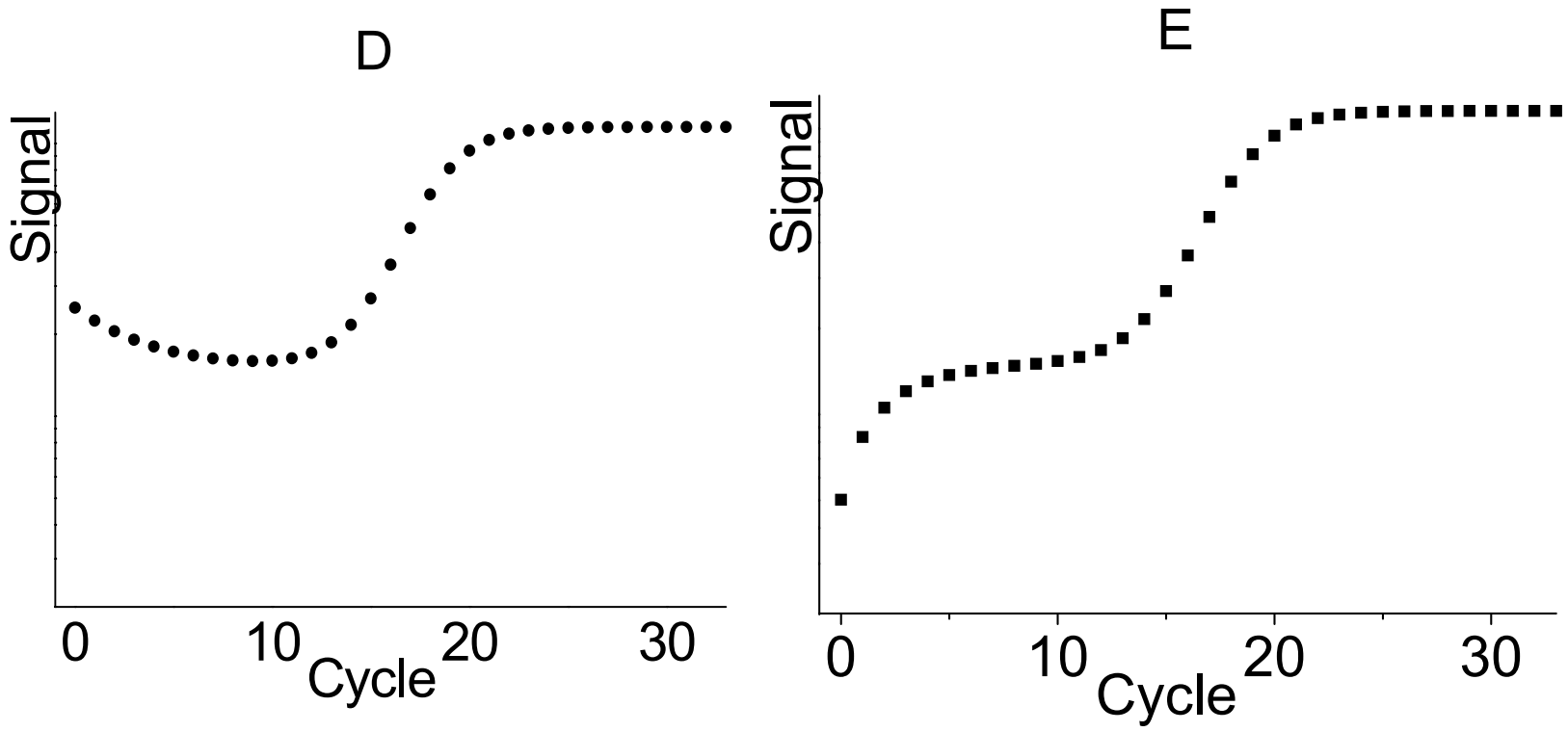
Constant background



Linear background

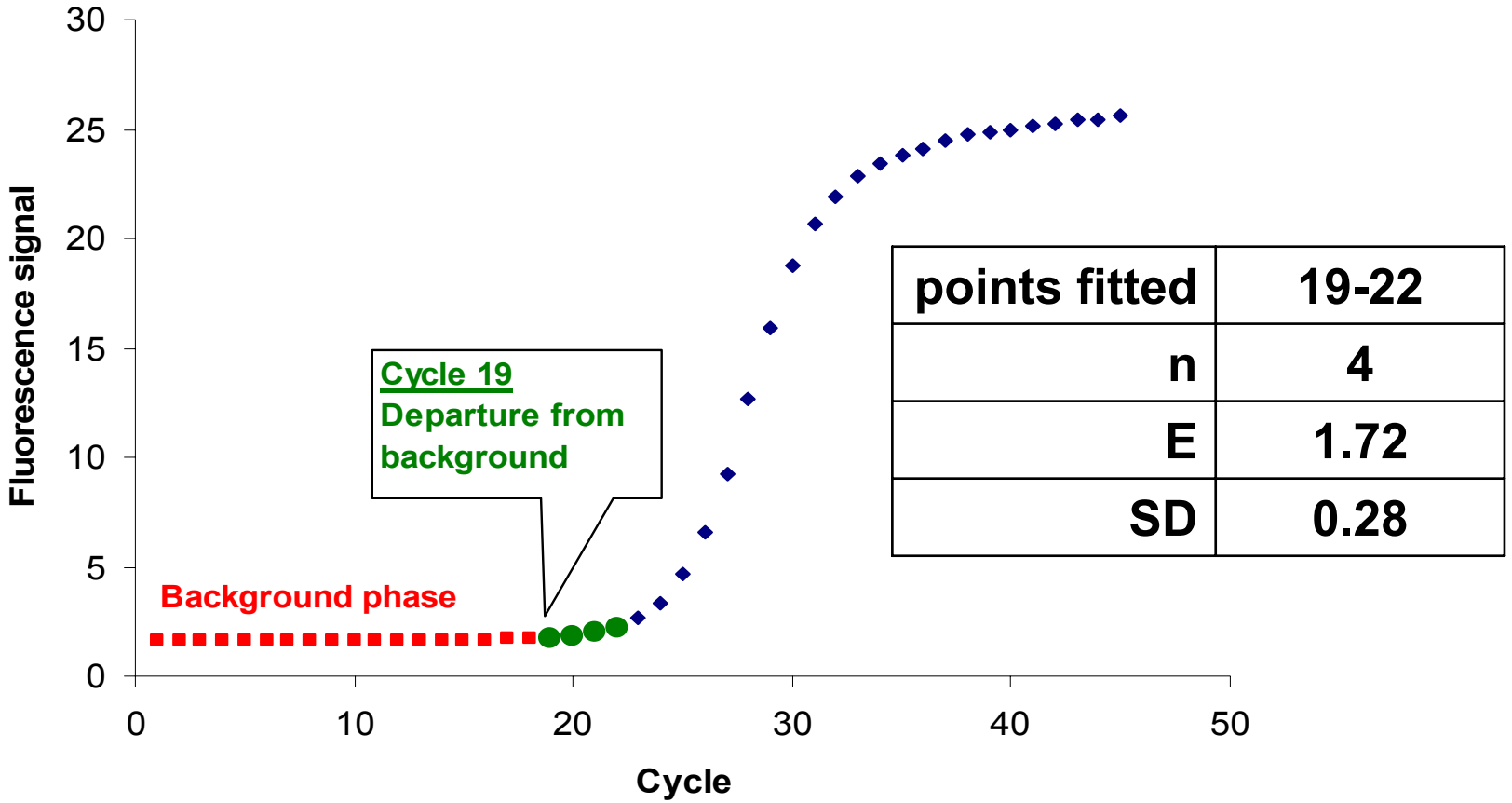


Non-linear background



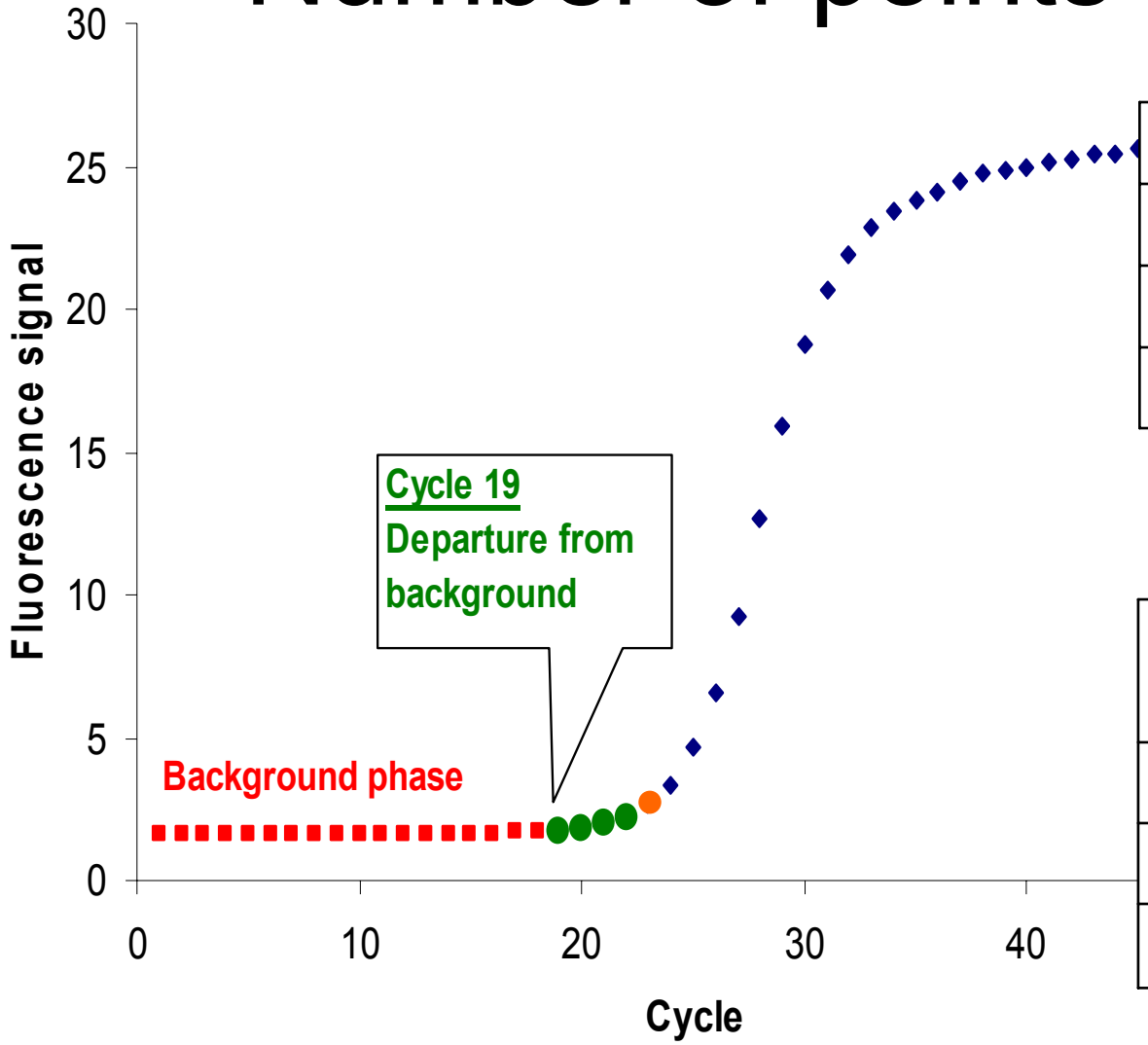
Where to fit?

Number of points fitted



points fitted	19-22
n	4
E	1.72
SD	0.28

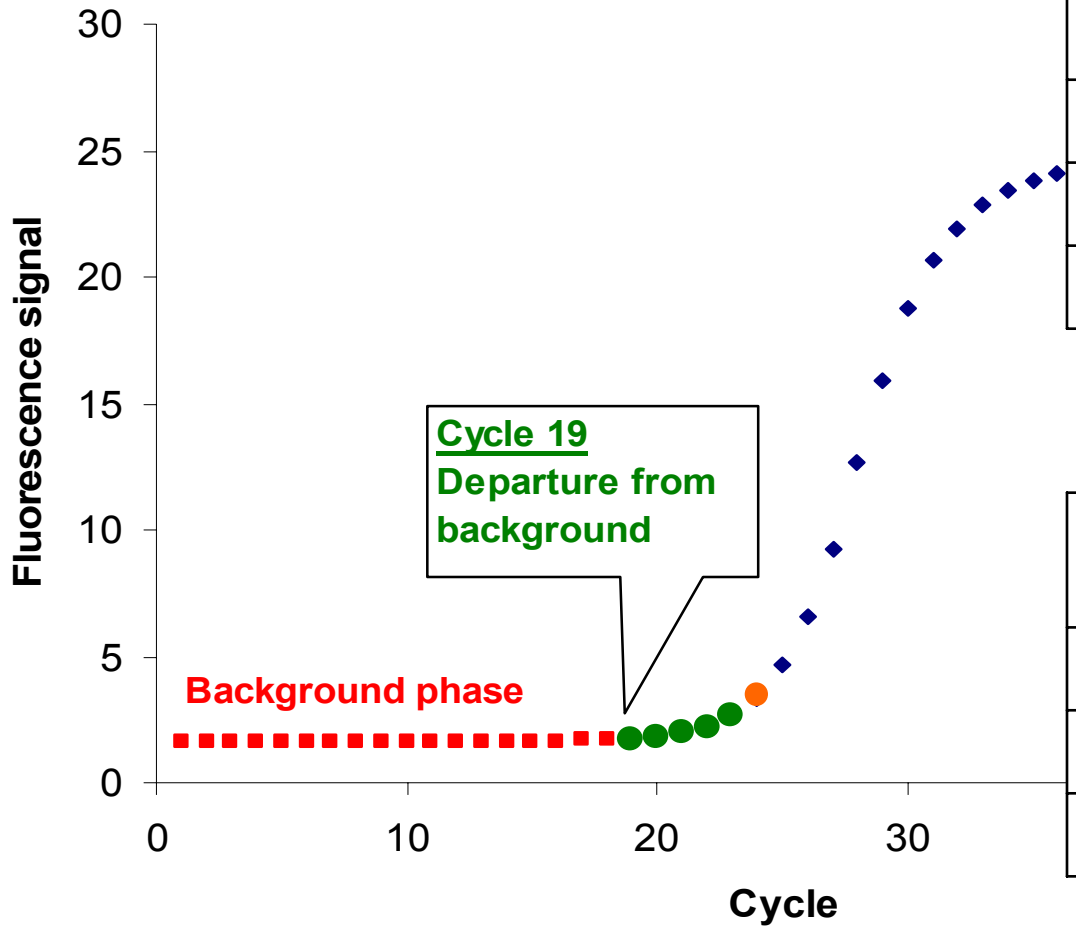
Number of points fitted



points fitted	19-22
n	4
E	1.86
SD	0.28

points fitted	19-23
n	5
E	1.85
SD	0.17

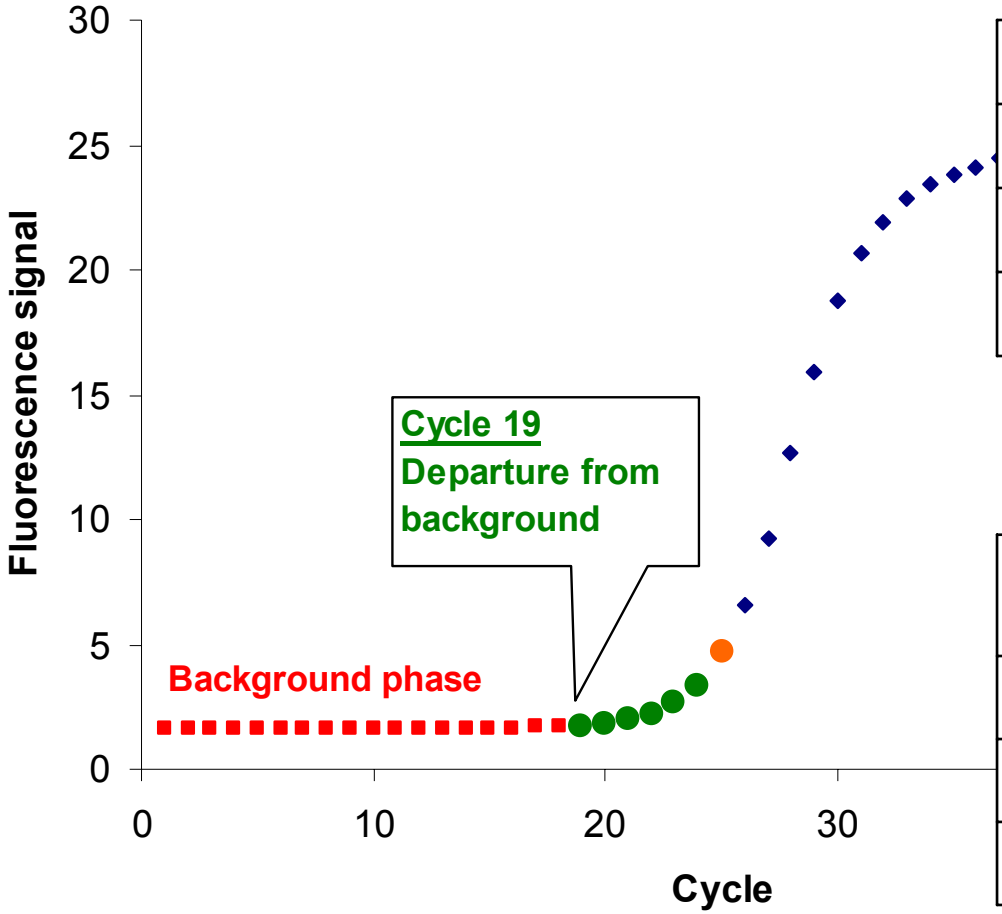
Number of points fitted



points fitted	19-23
n	5
E	1.85
SD	0.17

points fitted	19-24
n	6
E	1.76
SD	0.04

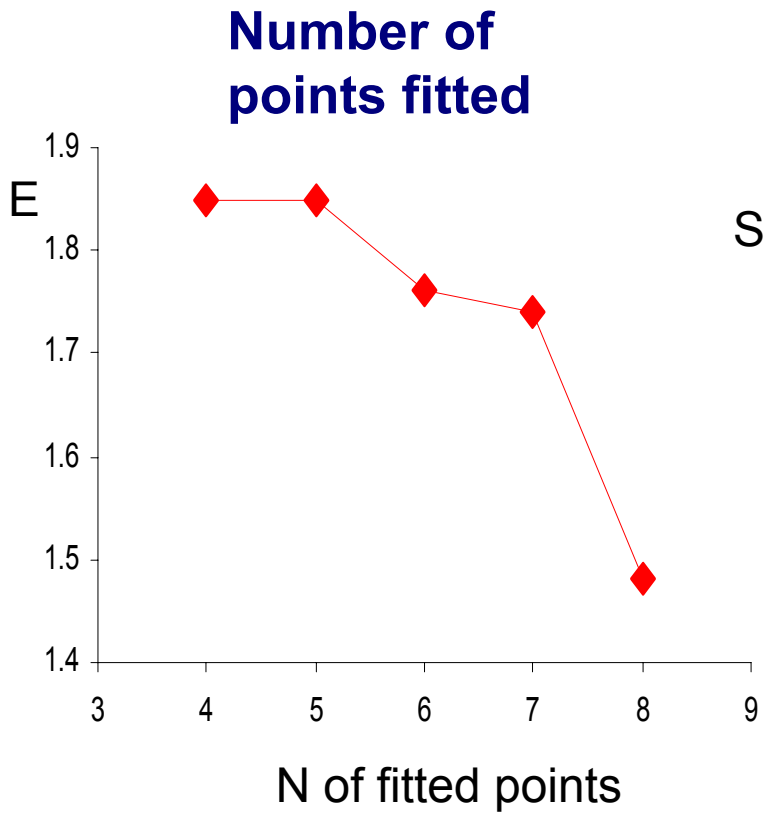
Number of points fitted

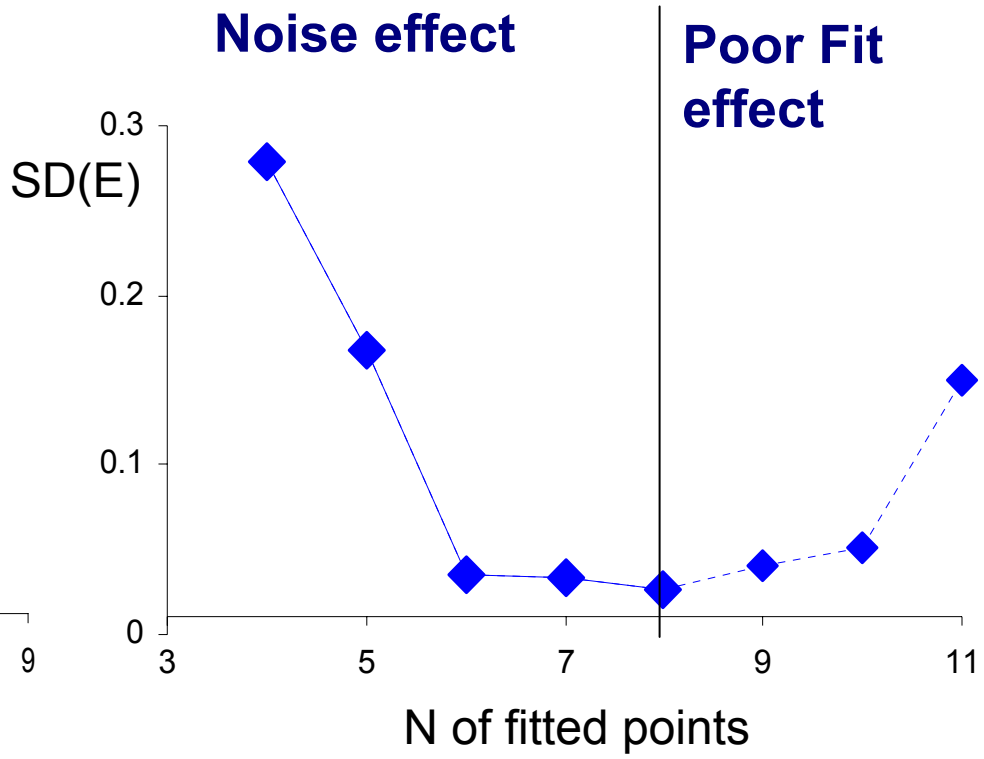


points fitted	19-24
n	6
E	1.76
SD	0.036

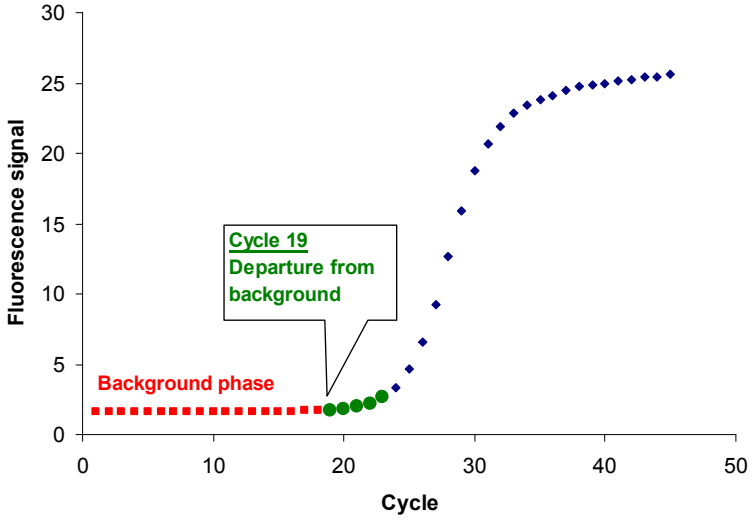
points fitted	19-25
n	7
E	1.74
SD	0.033

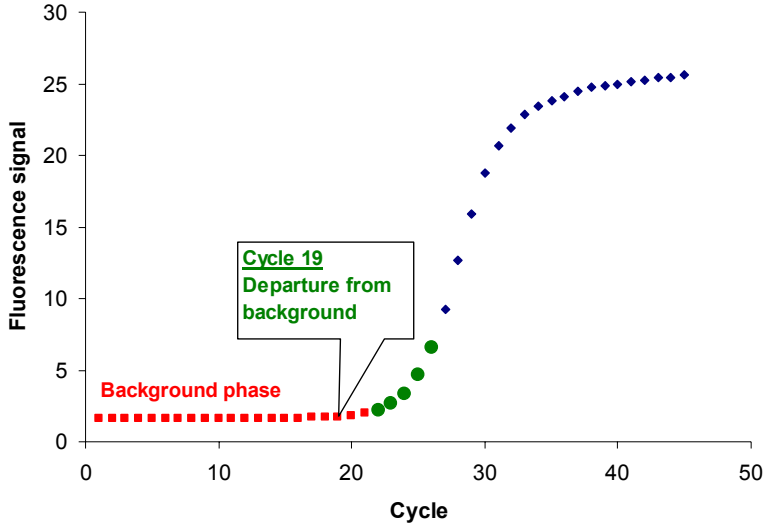
Where to fit?


Number of points fitted
E
1.9
1.8
1.7
1.6
1.5
1.4
3
4
5
6
7
8
9
N of fitted points


Noise effect
Poor Fit effect
SD(E)
0.3
0.2
0.1
0
3
5
7
9
11
N of fitted points

Where to fit? Effect of noise



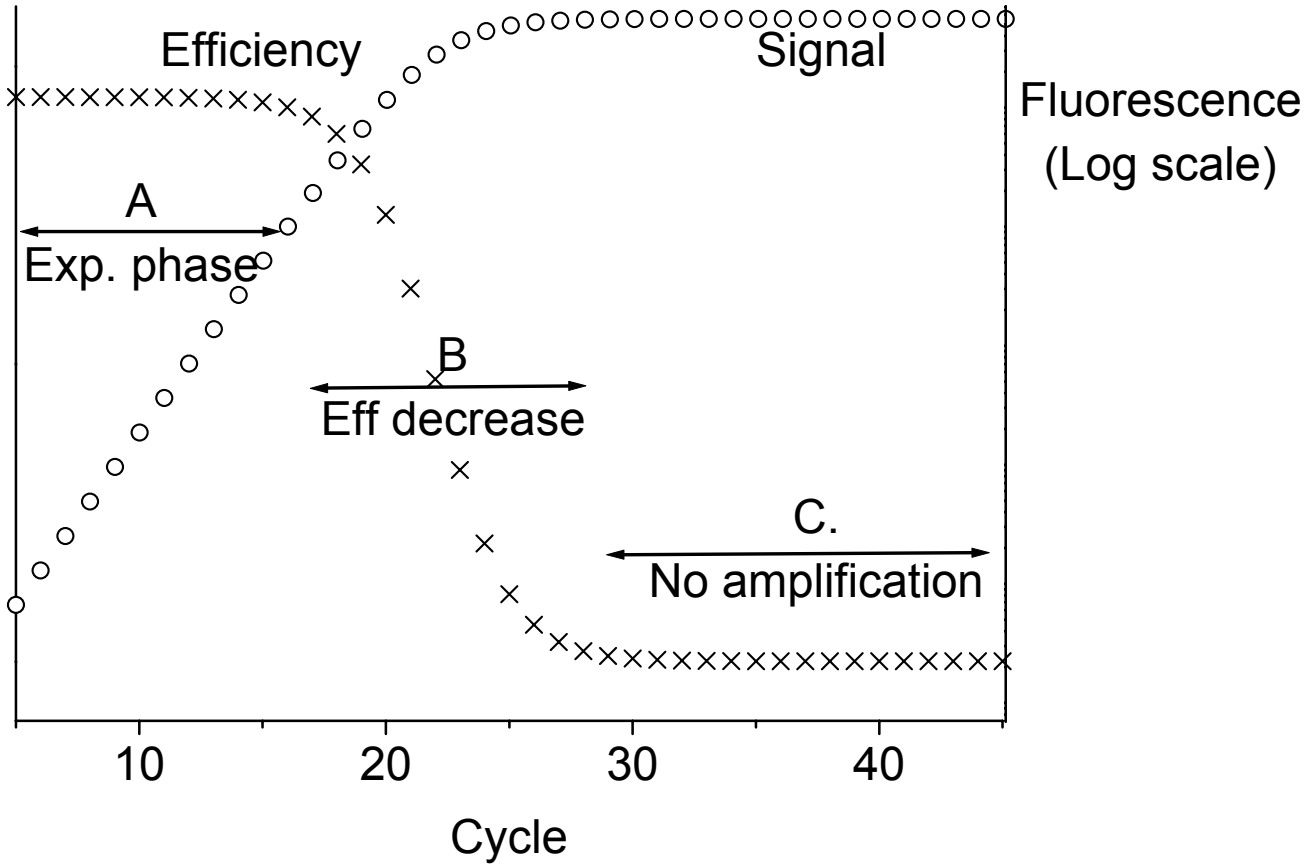


points fitted	19-23
n	5
E	1.85
SD	0.165

=
>
<

points fitted	22-26
n	5
E	1.70
SD	0.027

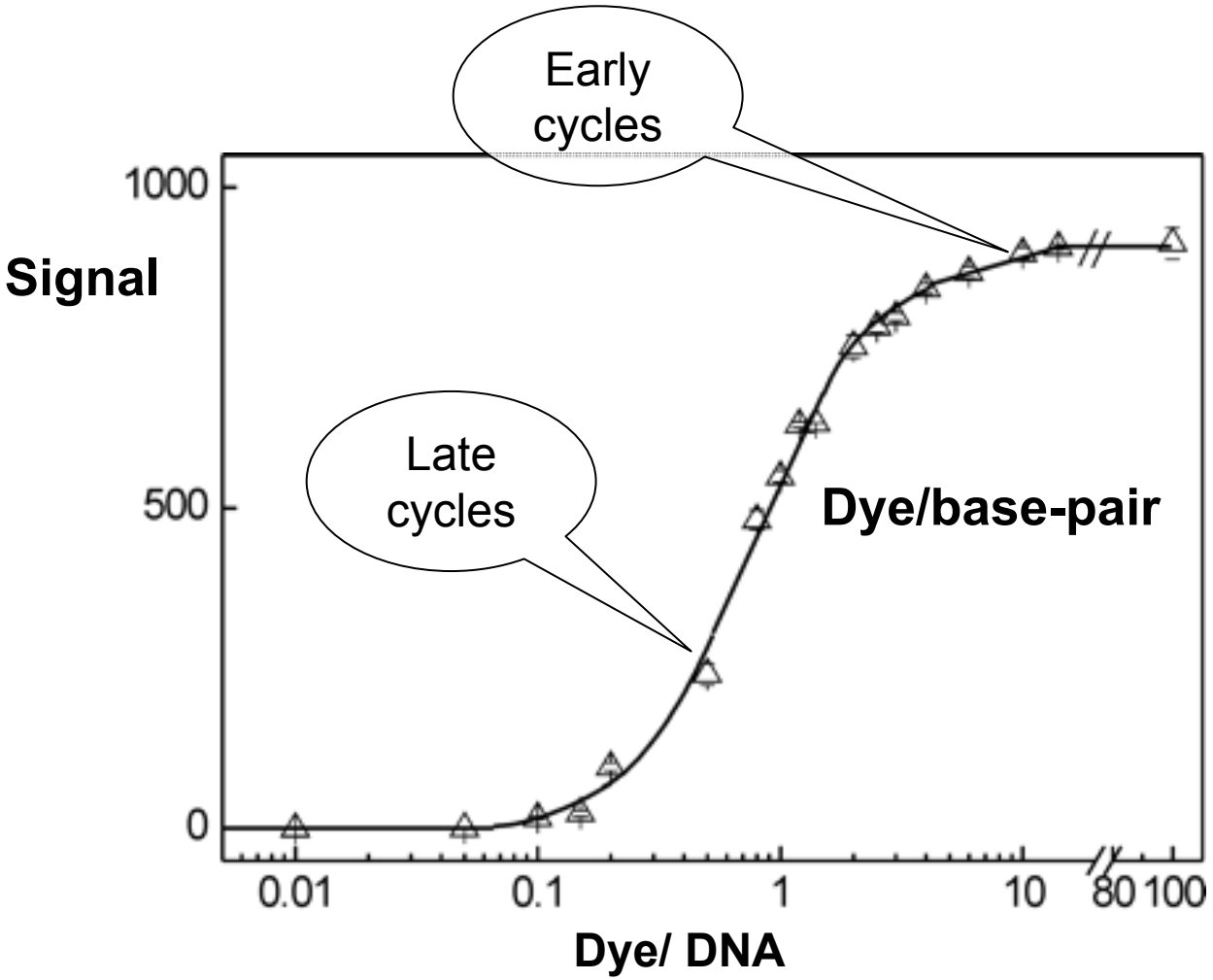
PCR kinetics



Summary

- Efficiency decreases as the reaction progresses.
- Is PCR exponential or logistic process ?
- How to evaluate amplification efficiency in early cycles?

SYBR Green Signal might not be proportional to DNA amount



…And might thus cause inaccurate efficiency estimation.

Do the efficiency estimation methods deliver their promise?

Is the estimated efficiency precise enough?

How close is the estimated efficiency to the efficiency calculated from the standard curve?

What can we do with the estimated efficiency?

Comparison of different efficiency estimation methods

- 53 standard curves, 8-21 samples each (mean= 14.4, median=15)
- Total: 762 samples
- $R^2 > 0.99$
- Minimal span of dilution - 256 fold
- Most common span of dilution – 10^4 or larger

Compared methods

- Tichopad et al., NAR **31**(20): e122
- Ramakers et al., *Neurosci. Lett.*, **339**, 62–66 (LinRegPCR)
- Peirson et al., NAR **31**(14): e73 (DART-PCR)
- Wilhelm et al., BioTechniques **34**, 324-32 (SoFAR)
- Liu & Saint, Biochem. Biophys. Res. Commun. **294**(2), 347-53 (4P)

Precision

Median SD(E)

	Stand[1]	LinReg[2]	DART[3]	SoFAR[4]	4P[5]
Rotogene	0.076	0.102	0.118	0.076	0.071*
LightCycler	0.036	0.043	0.034*	0.055	0.044
iCycler	0.067	0.134	0.146	0.098	0.055*
Total	0.062	0.101	0.11	0.081	0.058

*Most precise efficiency estimation method on this machine
Machine producing the most precise eff for a given method

[1]Tichopad et al., NAR **31**(20): e122
[2]Ramakers et al., *Neurosci. Lett.*, **339**, 62–66
[3]Peirson et al., NAR **31**(14): e73
[4]Wilhelm et al., BioTechniques **34**, 324-32
[5]Liu & Saint, Biochem. Biophys. Res. Commun. **294**(2), 347-53

Closeness to standard curve value

Distance from standard curve value

	Stand[1]	LinReg[2]	DART[3]	SoFAR[4]	4P[5]
Rotogene	-0.47	-0.26	-0.24*	-0.31	-0.33
LightCycler	-0.3	-0.08	-0.06	-0.06*	-0.26
iCycler	-0.27	-0.11	-0.11	-0.12	0.03*
Average	-0.33	-0.15	-0.14*	-0.17	-0.16

*Closest efficiency estimation method on this machine

Machine producing the closest efficiency for a given eff estimation method

[1]Tichopad et al., NAR **31**(20): e122

[2]Ramakers et al., *Neurosci. Lett.*, **339**, 62–66

[3]Peirson et al., NAR **31**(14): e73

[4]Wilhelm et al., BioTechniques **34**, 324-32

[5]Liu & Saint, Biochem. Biophys. Res. Commun. **294**(2), 347-53

Application 1: Quantification

N_0 = No. of molecules at cycle 0

N_{CT} = No. of molecules at threshold =10^9

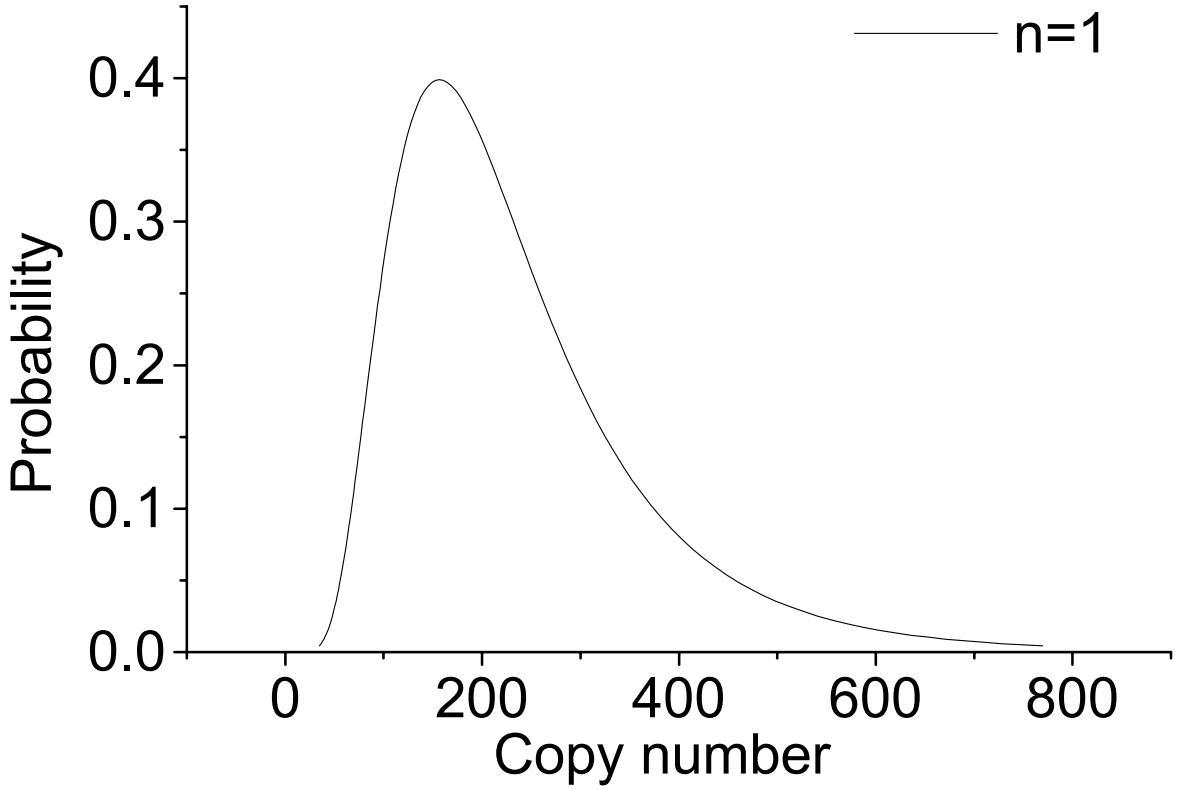
SD = Standard Deviation

$$N_0 = \frac{N_{CT}}{((1+E)-3*SD(E))^{CT}}$$

$$0 \le E \le 1$$

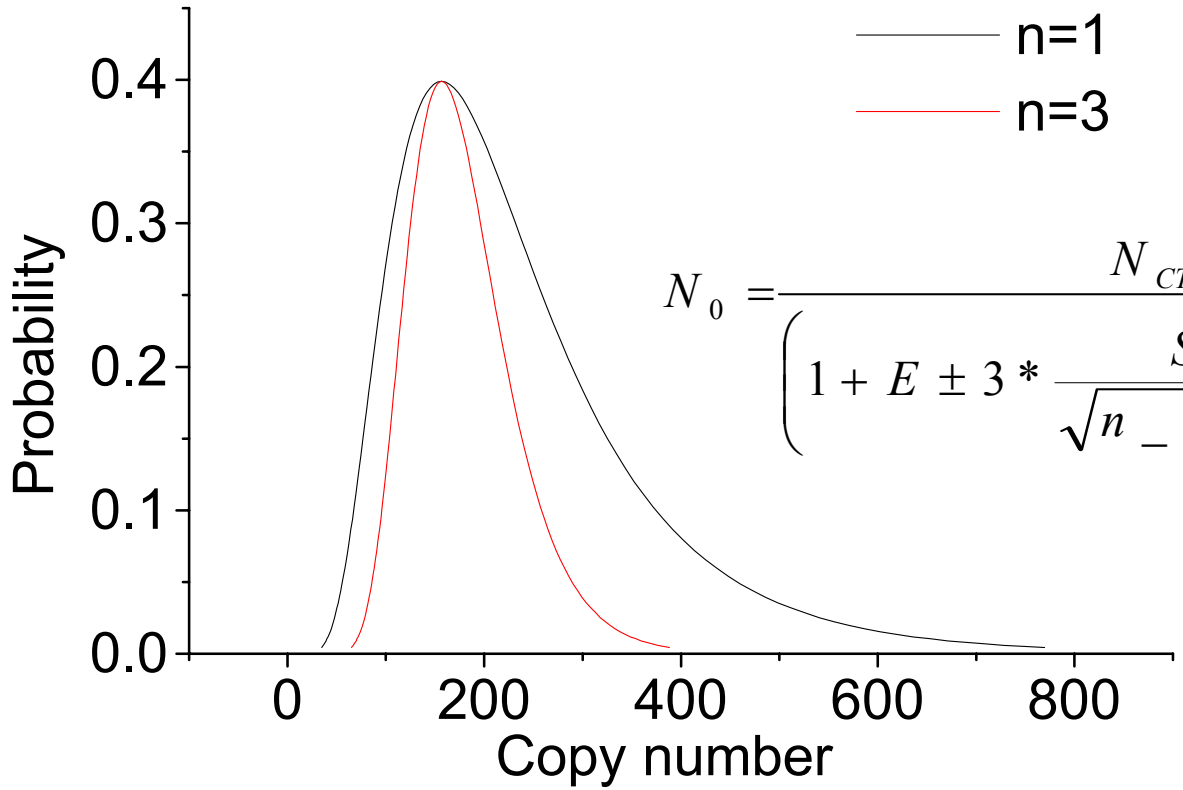
N_0 Distribution

CT=28, $N_{CT}=10^9$, E=0.95, SD(E)= 0.035



N_0 Distribution

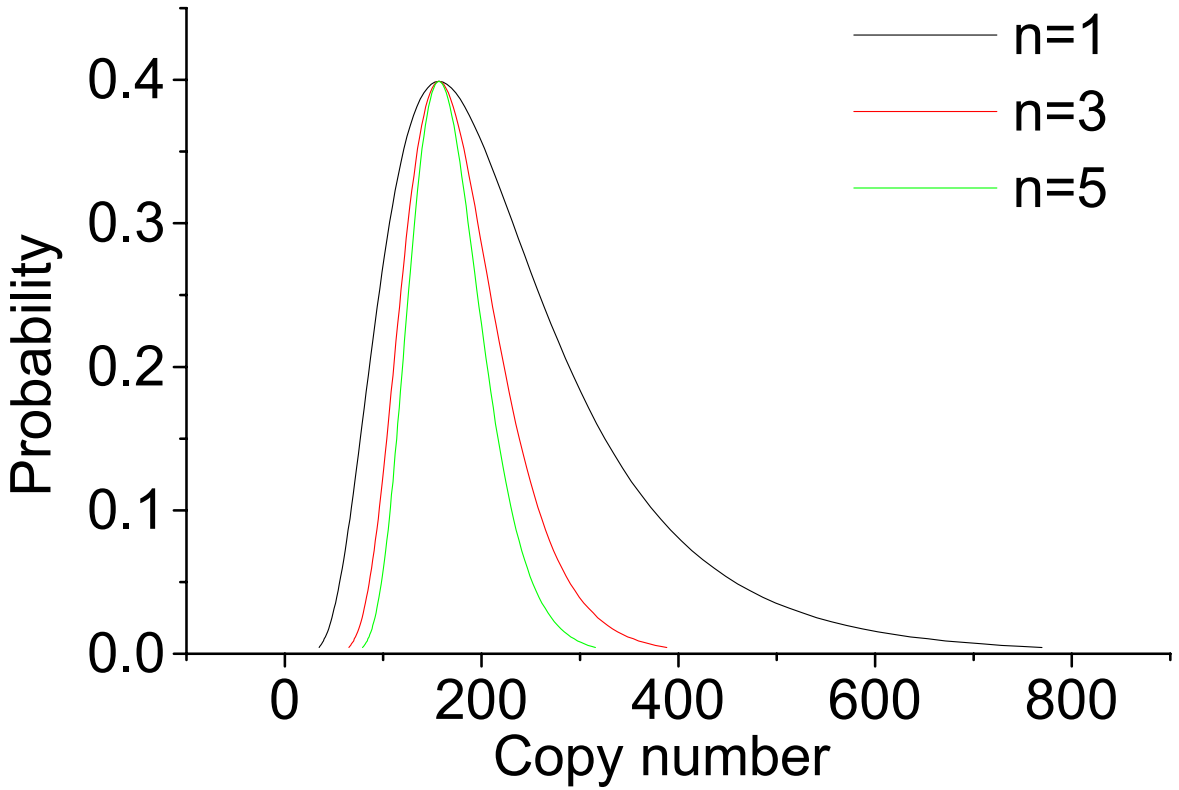
CT=28, $N_{CT}=10^9$, E=0.95, SD(E)= 0.035



$$N_0 = \frac{N_{CT}}{\left(1 + E \pm 3 * \frac{SD(E)}{\sqrt{n_PCR_rep}}\right)^{CT}}$$

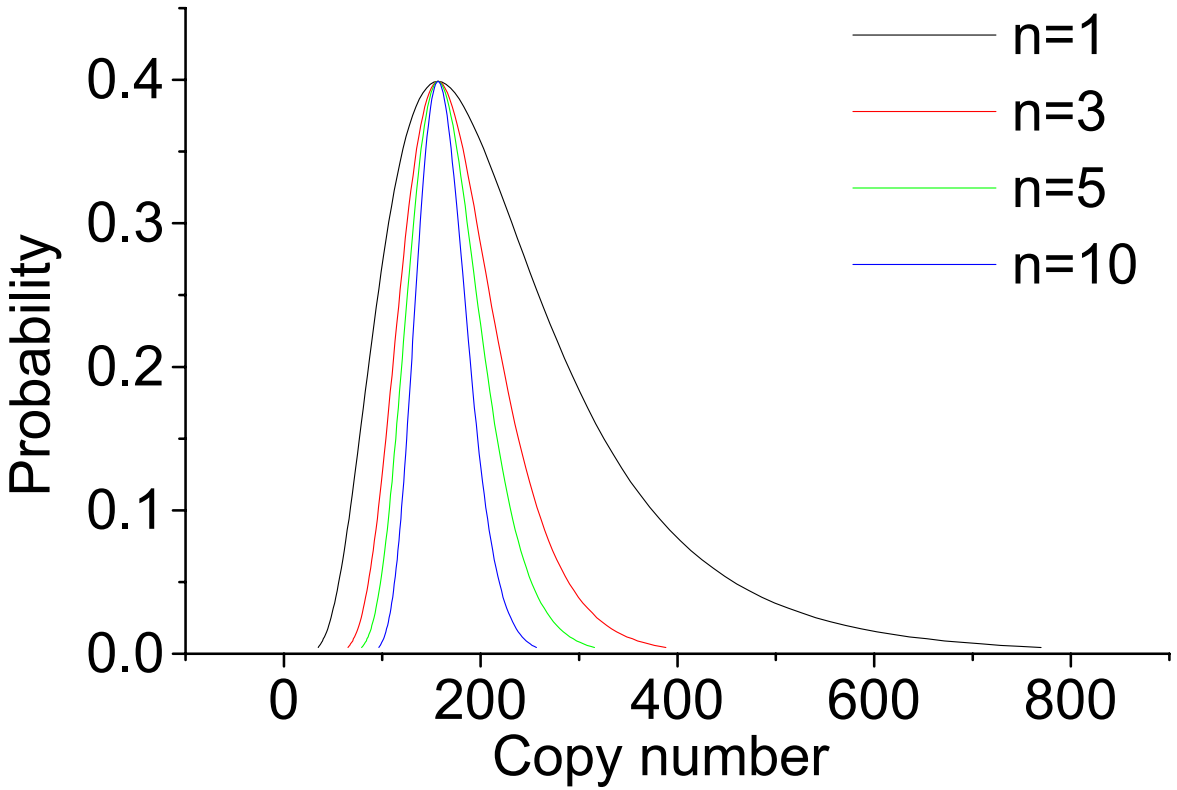
N_0 Distribution

CT=28, $N_{CT}=10^9$, E=0.95, SD(E)= 0.035



N_0 Distribution

CT=28, $N_{CT}=10^9$, E=0.95, SD(E)= 0.035



Precise quantification based on sample specific efficiency requires at least 3 PCR replicates

Replacement of standard curve value

- Gentle et al. 2001
- Peirson et al. 2003
- Tichopad et al. 2003
- Peyton et al. 2004

Replacement of standard curve value

Distance from standard curve value

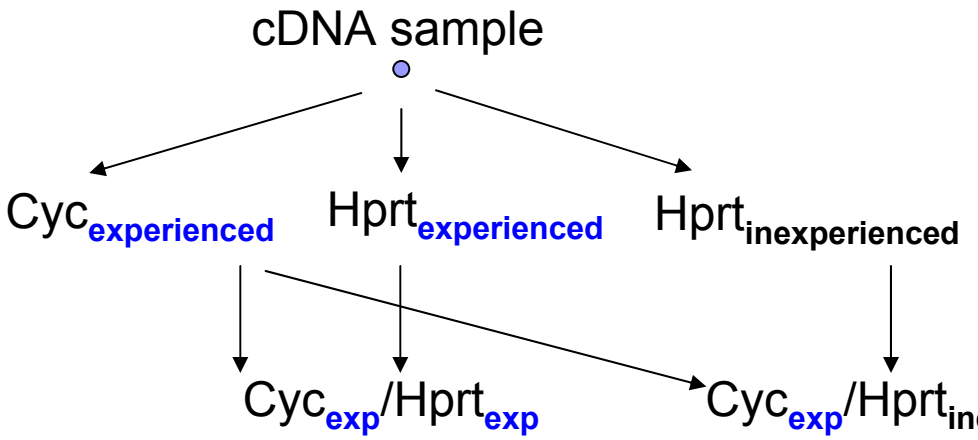
	Stand[1]	LinReg[2]	DART[3]	SoFAR[4]	4P[5]
Average	**-0.33**	**-0.15**	**-0.14***	**-0.17**	**-0.16**

$$\text{Ratio} = \frac{E_{\text{target}}^{\Delta CT_{\text{target}}\,(\text{control}-\text{test})}}{E_{\text{reference}}^{\Delta CT_{\text{reference}}\,(\text{control}-\text{test})}}$$

Using inaccurate efficiency affects the accuracy of quantification.

Kinetics Quality Assessment: Experimental design

15 sets of 5 replicate animals
(total of 75 cDNA samples)



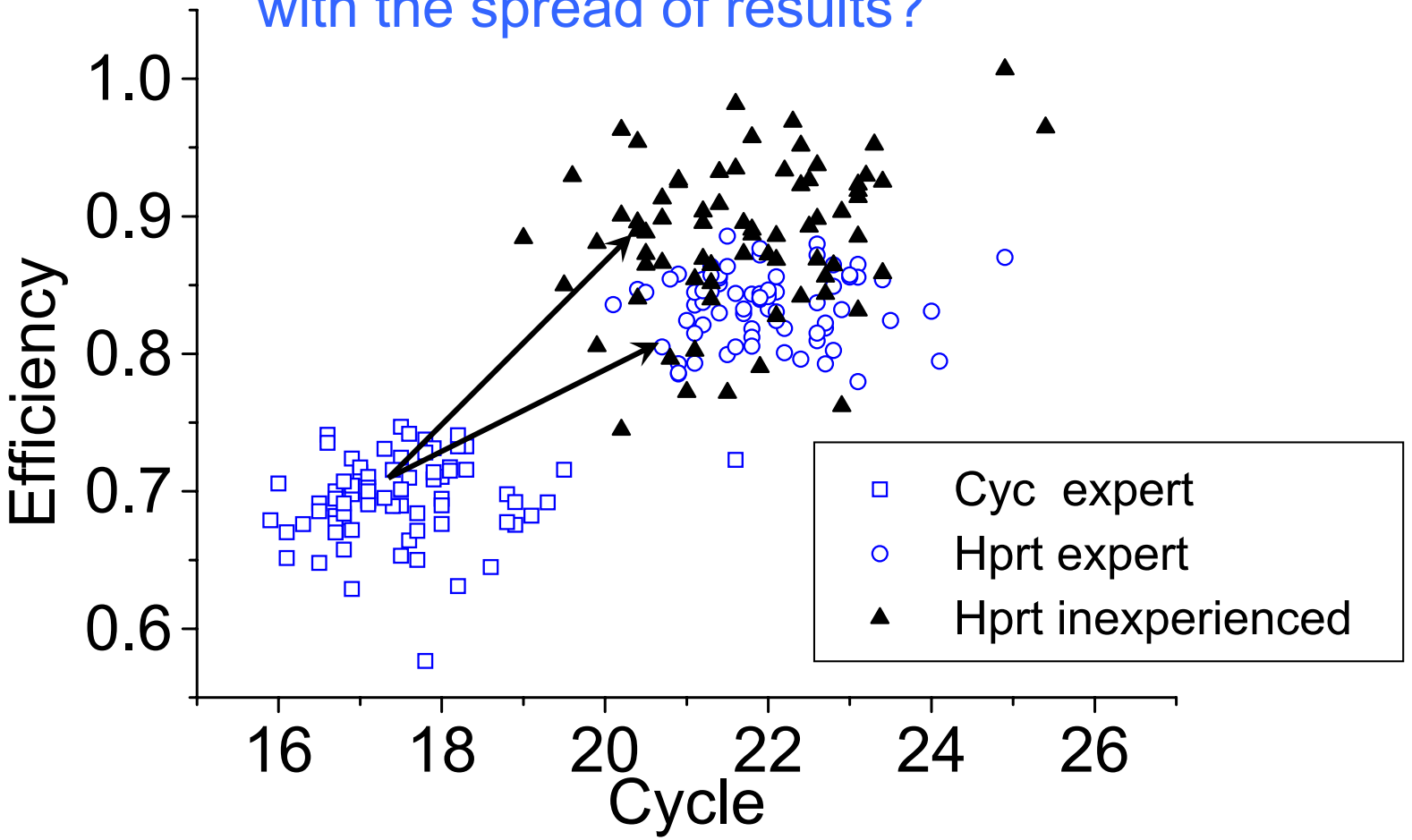
Spread of results

15 sets of 5 + 15 set of 5 = 30 sets of 5

*Bar and Muszta, Biotechniques, 39(3), 2005, 333-340

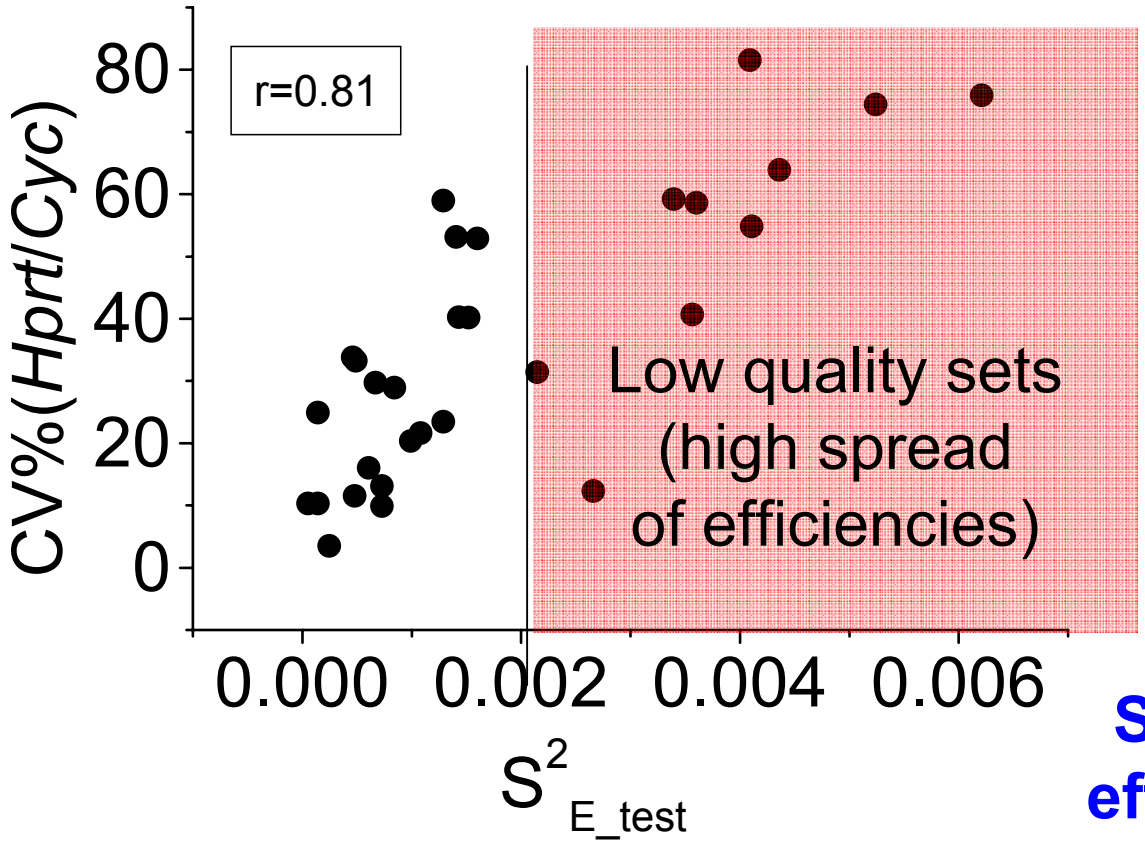
Efficiency spread of samples

Will the spread of efficiencies be associated with the spread of results?



Variance test

Spread of quantities



Spread of efficiencies

Critical S^2_E = average of $S^2_{E_ref}$ from many PCR systems

Do low quality sets contain more outliers?

% of outliers in **high quality** sets (similar efficiencies)	% of outliers in **low quality** sets (dissimilar efficiencies)
4.9% (*16%)	38% (*43%)

*In parenthesis – results obtained with average efficiency of the standard curve

Conclusions:

Sample specific efficiencies may be used:

1. For quantification, but need adjustment and use of replicates.
2. To replace standard curve efficiency after adjustment.
3. For kinetics QA.

Acknowledgments:

- Stuart Peirson for help with DART
- Jochen Wilhelm for help with SoFAR
- Mikael Kubista and Michael Pfaffl for insightful discussions and comments



?