

Análise da expressão gênica por meio da PCR quantitativa

Dr^a Patrícia Natália Silva

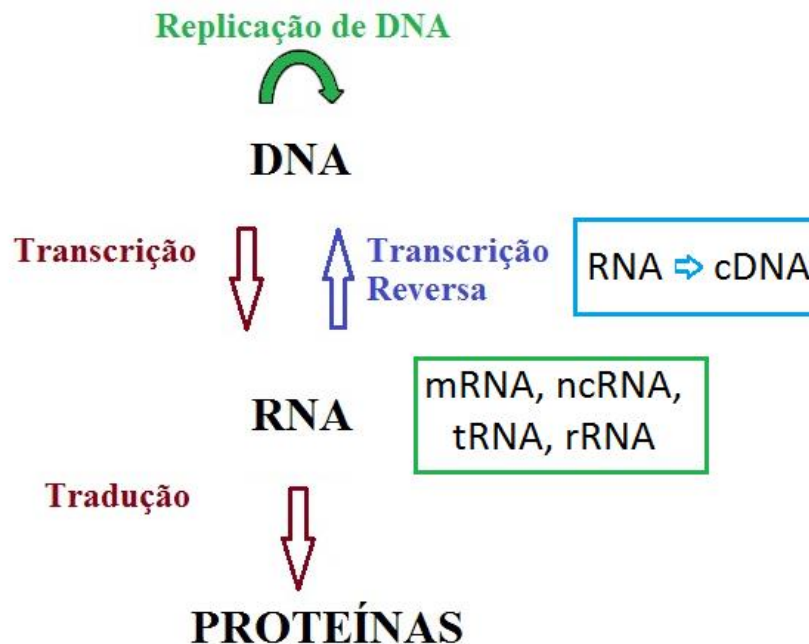
Curso de Introdução à análise bioinformática aplicada à genética

Sumário

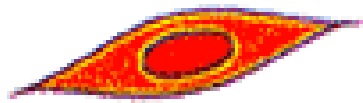
- Revisão sobre qPCR em expressão gênica
- Resumo de metodologias SYBR[®] e Taqman[®]
- Conceitos importantes
 - *Baseline, threshold, threshold cycle*, normalização e eficiência de reação
- Métodos de Análises:
 - $\Delta\Delta CT$ e curva padrão relativa
- Análises estatísticas
 - *Outliers*, comparação entre RQ, $2^{-\Delta CT}$ e ΔCT

Expressão gênica

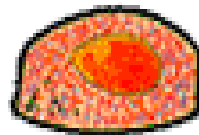
- Expressão gênica: processo pelo qual a informação de um gene é utilizado na síntese de um produto funcional



Expressão gênica



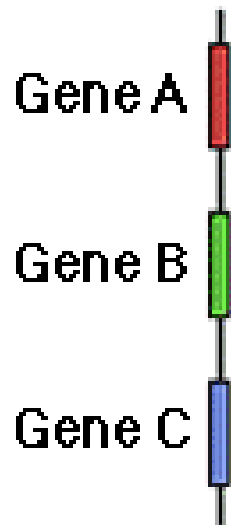
Muscle cell



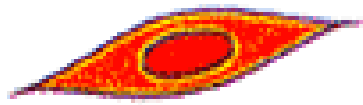
Skin cell



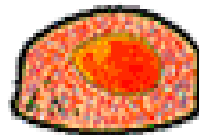
Nerve cell



Expressão gênica



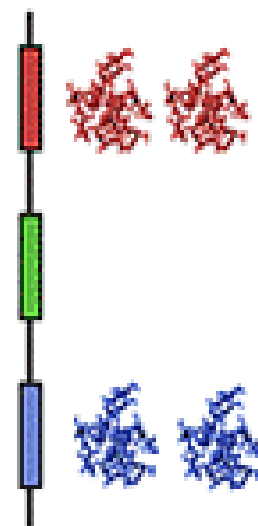
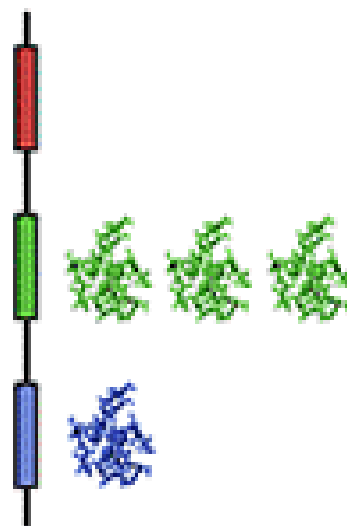
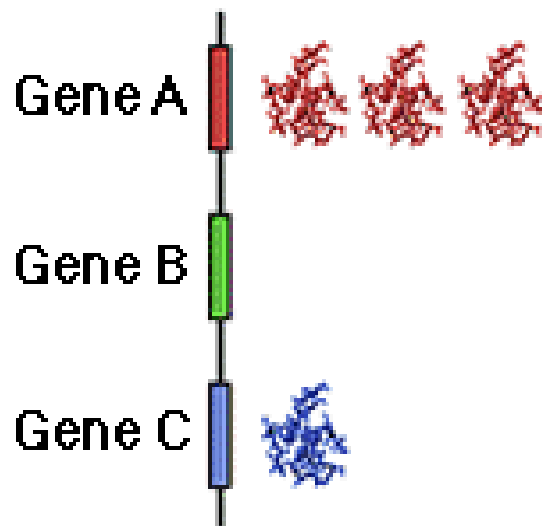
Muscle cell



Skin cell



Nerve cell



PCR quantitativa em Tempo Real

- Método usado para medir a quantidade de cDNA inicial amplificado por PCR
- Há uma relação quantitativa entre a quantidade inicial da amostra alvo e a quantidade de produto da PCR após um determinado número de ciclos

$$P = T (1+E)^n$$
$$T_{\text{inicial}} = \frac{P_{\text{final}}}{(1+E)^n}$$

P = Produto final

T = Template no início da reação

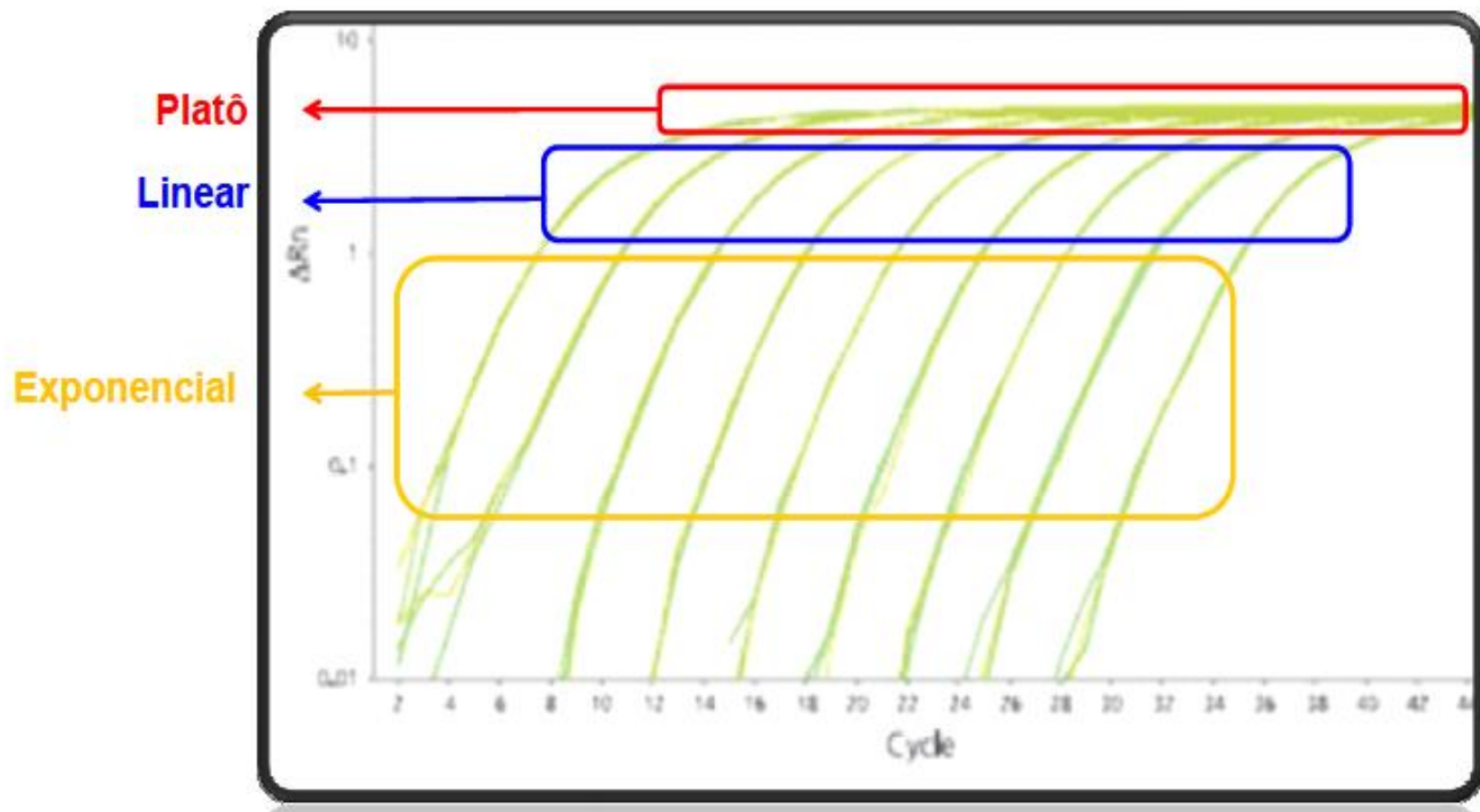
n = Número de ciclos

E = Eficiência

Taxa de Duplicação de DNA por PCR

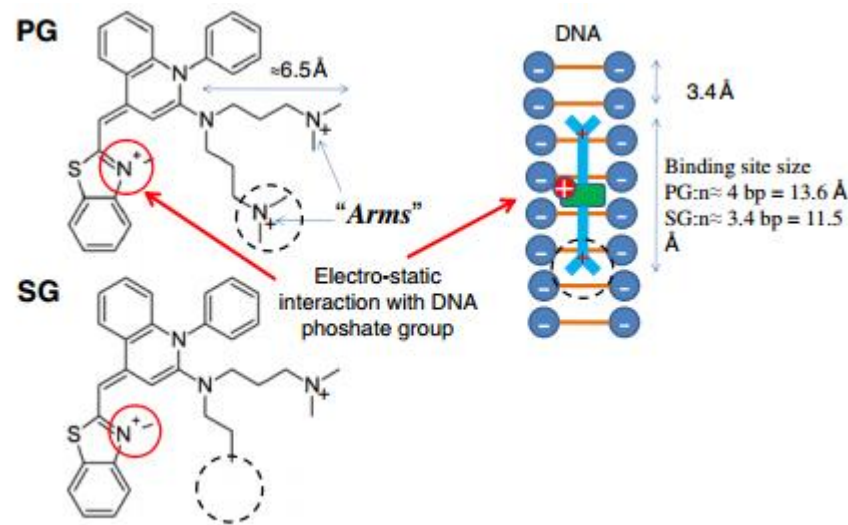


Fases de uma PCR



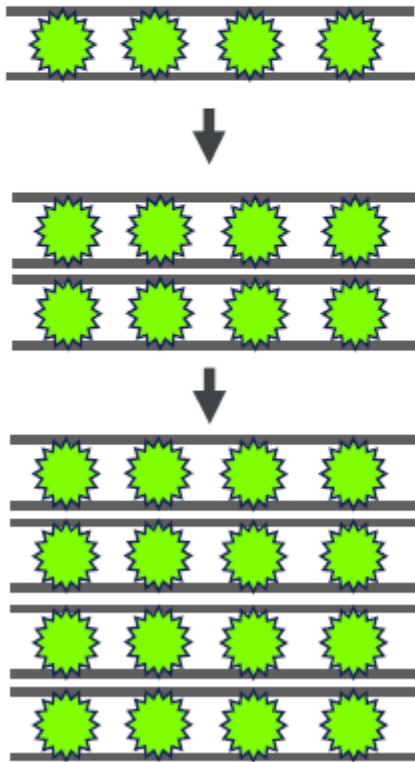
Sistemas de detecção da expressão gênica

- Agentes intercalantes de DNA
- SYBR[®] Green
 - Afinidade por DNA dupla fita
 - 25 vezes mais sensível que o Brometo de Etídeo



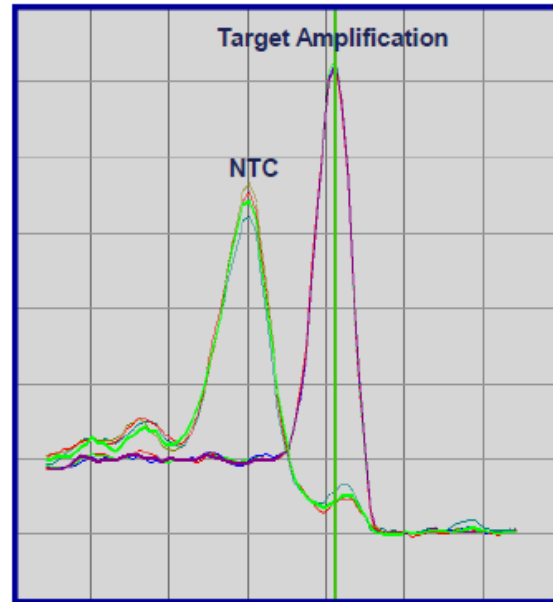
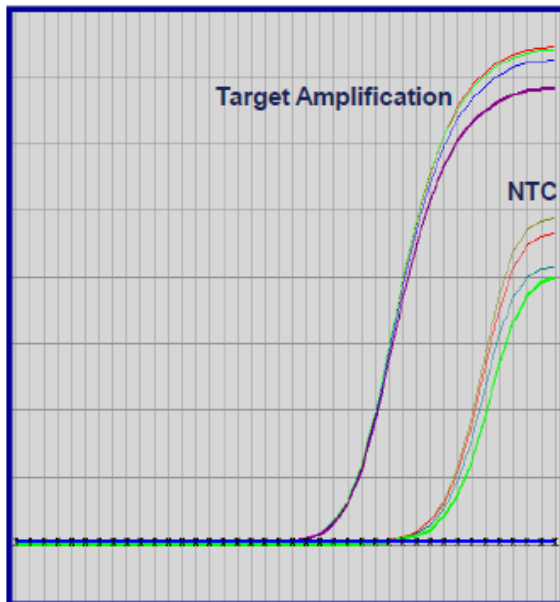
Agentes Intercalantes de DNA

Aumento do sinal é proporcional o aumento do produto



SYBR[®]Green

- Curva de dissociação (Curva de *melting*)
- T_m (Temperatura de *melting*)
 - Temperatura onde metade do produto de PCR está dissociado (desnaturado)



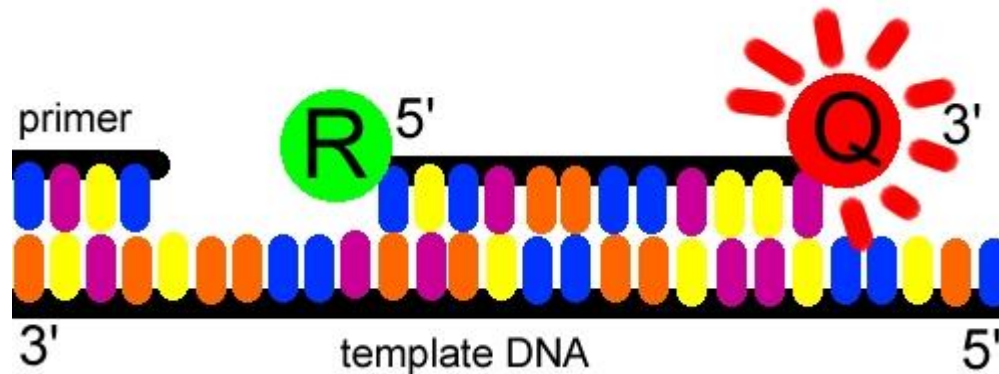
SYBR[®]Green

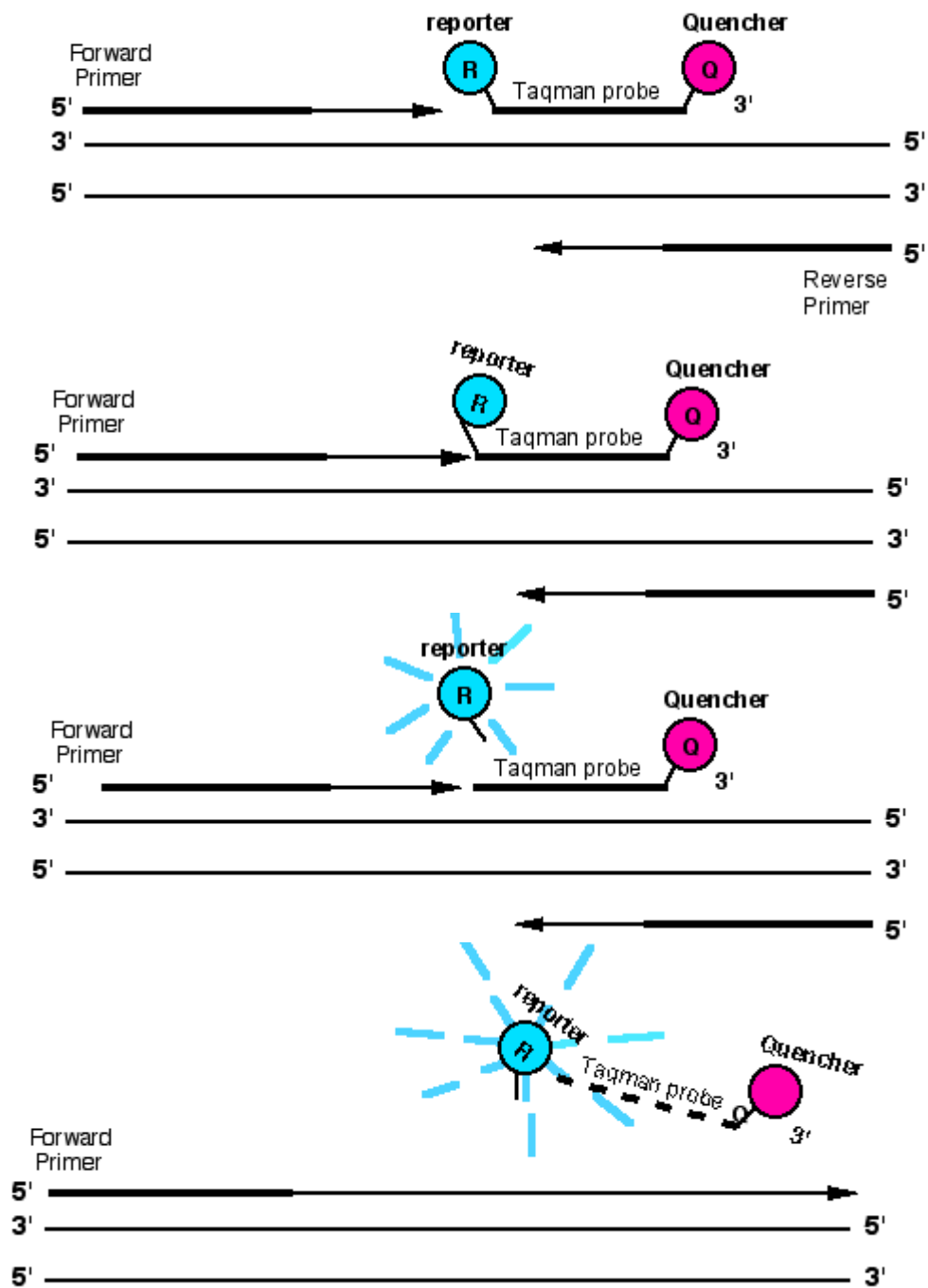
- Liga-se inespecificamente a qualquer dupla fita
- Produtos não específicos geram sinal
- Resultados quantitativos incorretos
- Desenho de *primers* para regiões desfavoráveis à síntese
- Menos dispendiosa

Sondas TaqMan[®]

- Sondas lineares (hidrólise)
- Além dos *primers*:

Sonda com *Quencher* e *Reporter*





TaqMan[®]

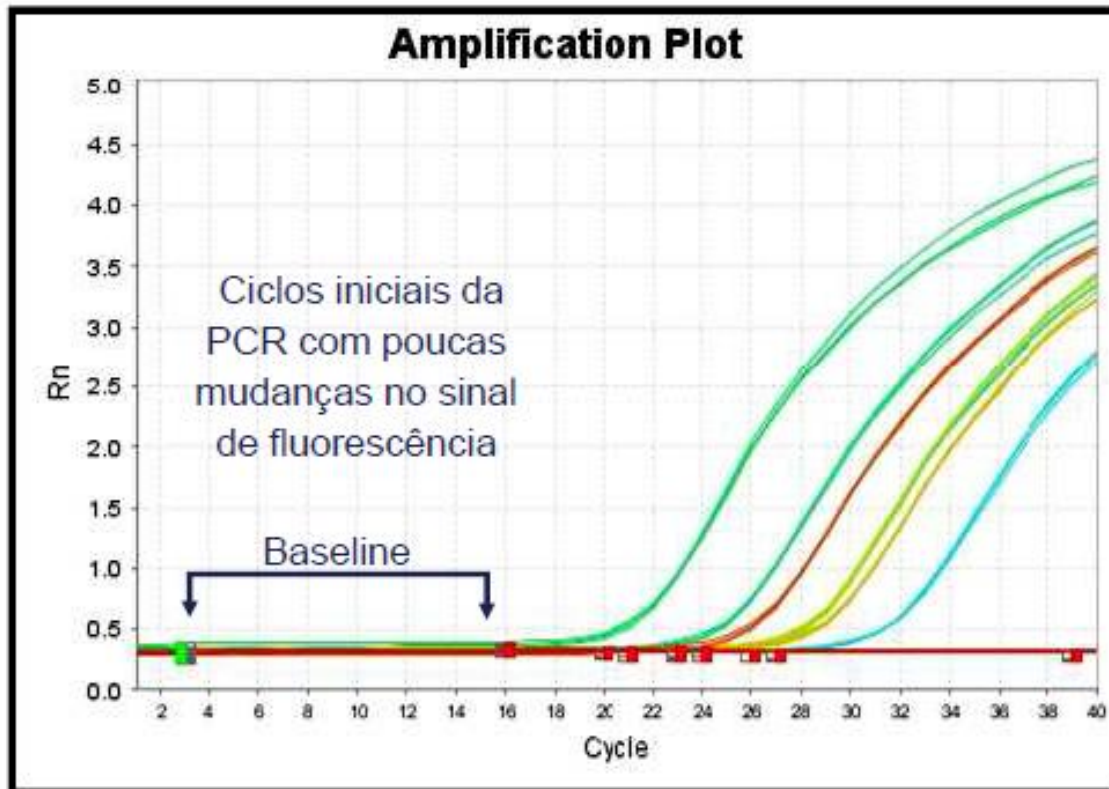
- Altamente específico e sensível
- Não permite, nem precisa de dissociação
- Permite Multiplex
- Mais dispendiosa \$



Conceitos chave

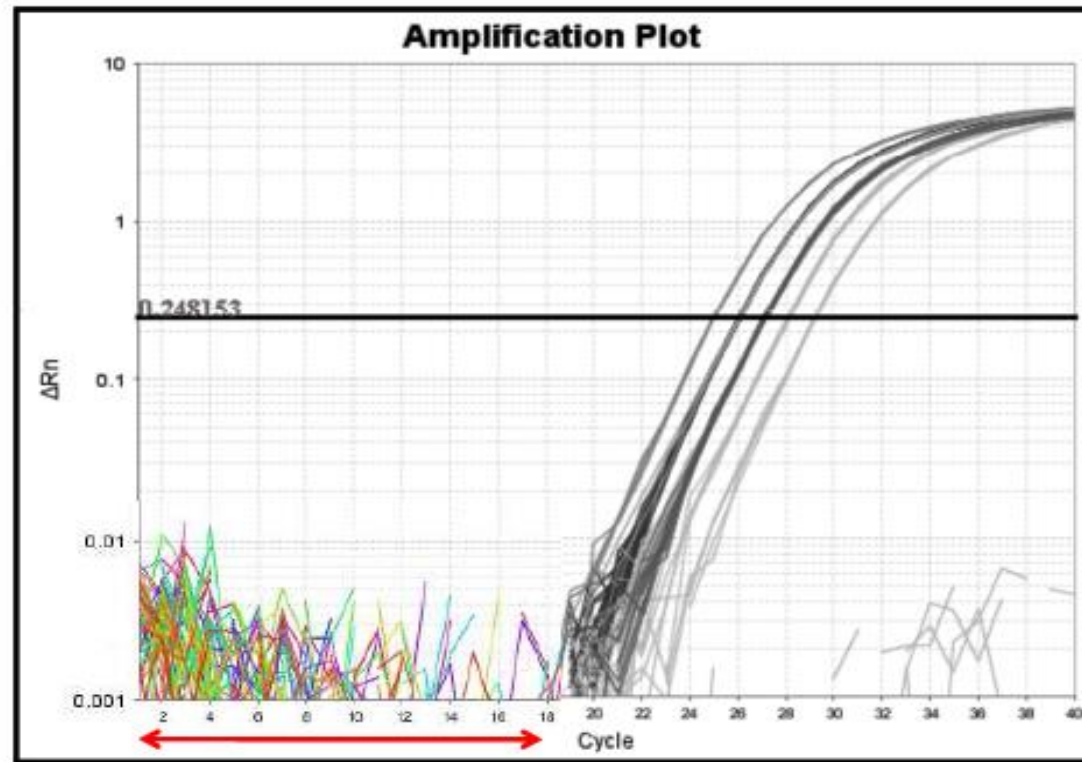
- *Baseline*

Fase onde a intensidade de sinal de produto amplificado não ultrapassa a quantidade de fluorescência presente no meio.



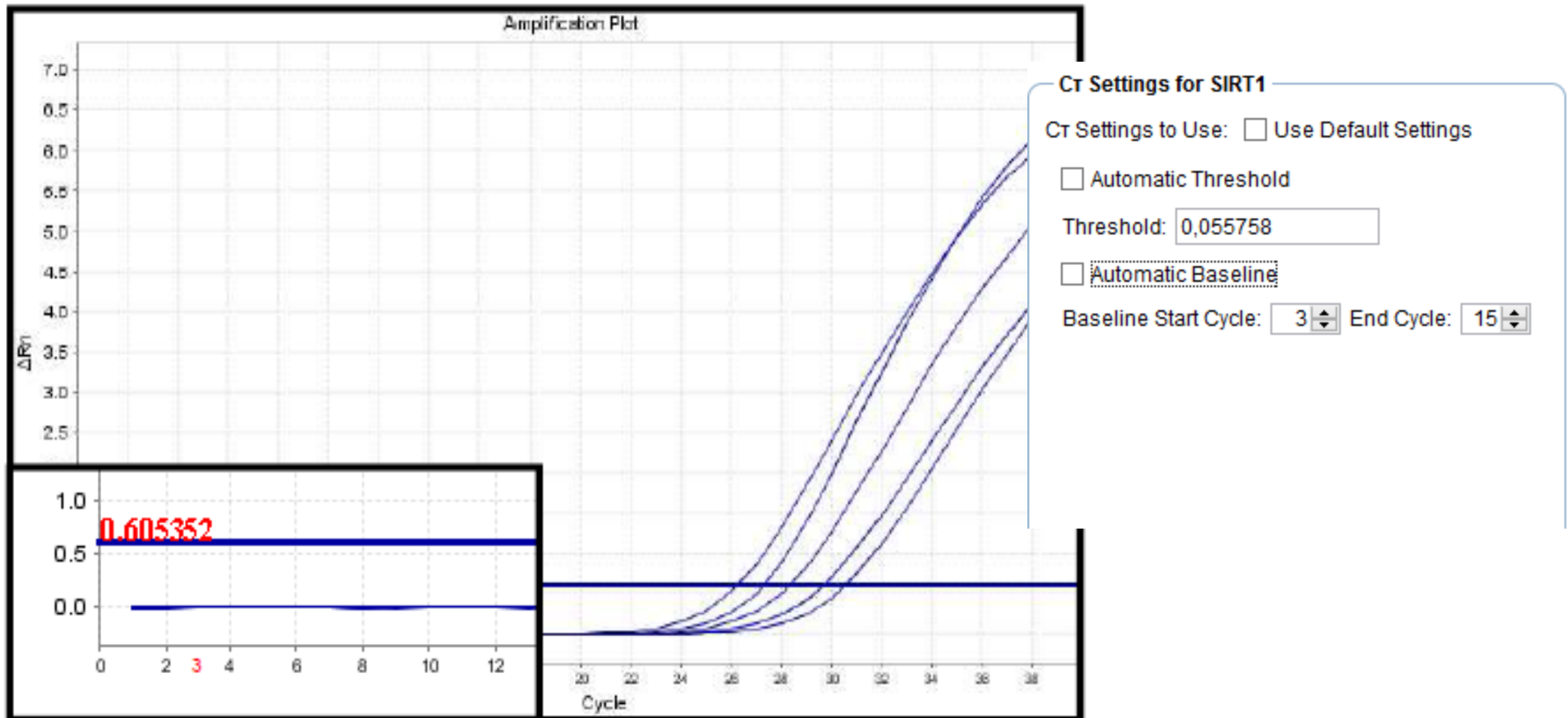
Baseline

- Background – fluorescência dos primeiros ciclos da PCR



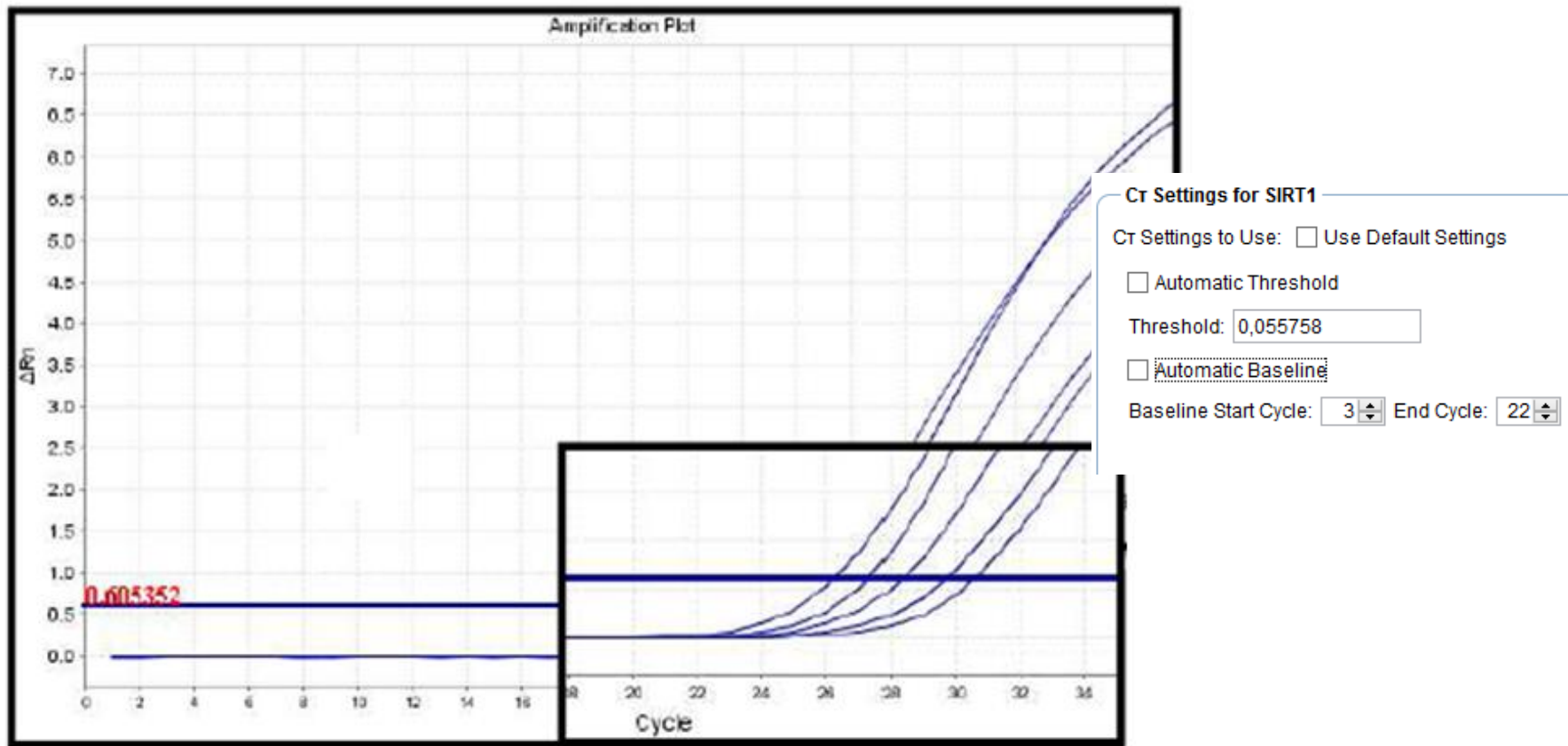
Baseline

Em que ciclo começa?

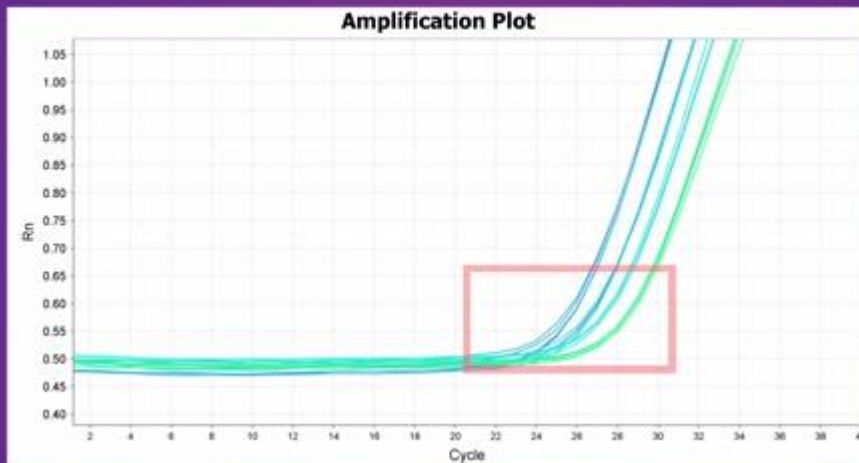


Baseline

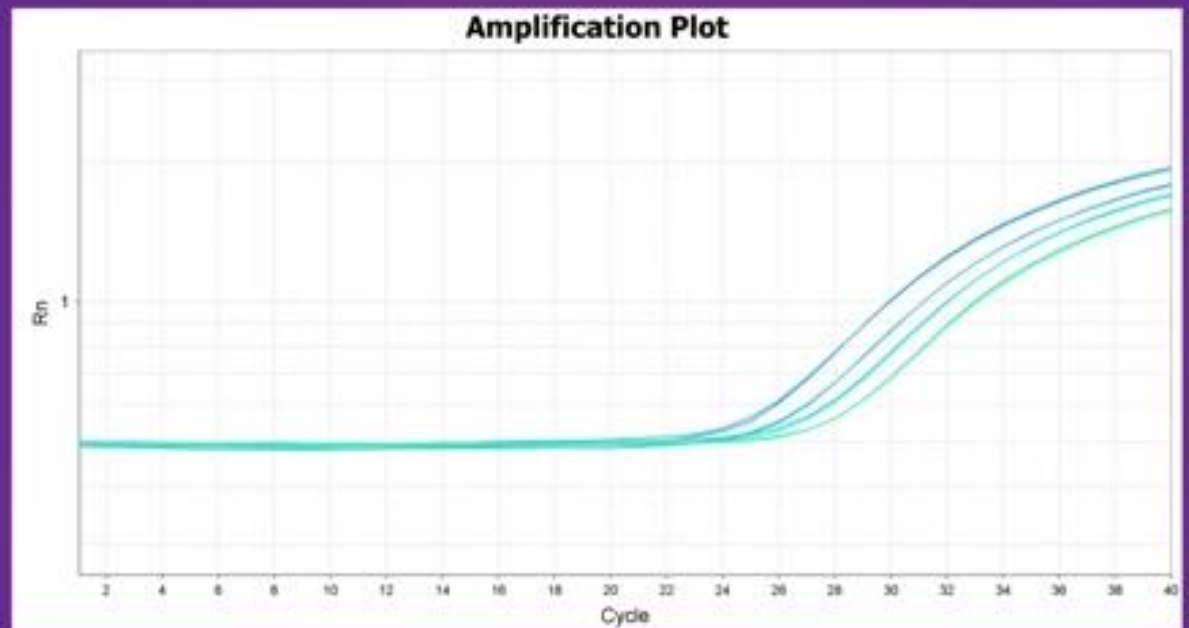
Em que ciclo termina?



Before baseline subtraction



After baseline subtraction



Baseline

O que acontece se o *baseline* for muito baixo?



Analysis Settings

Cr Settings for RNase P TAMRA

Cr Settings to Use: ☐ Use Default Settings

☐ Automatic Threshold

Threshold: 0.605352

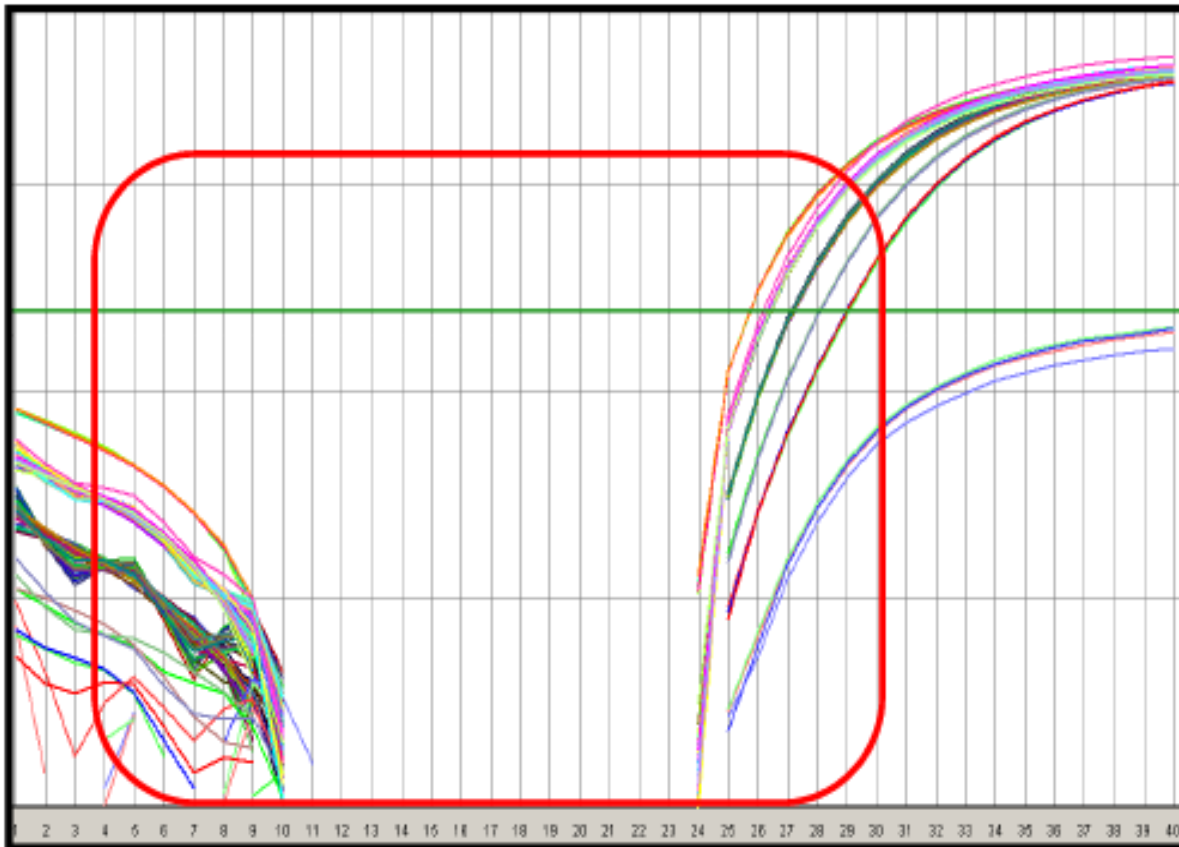
☐ Automatic Baseline

Baseline Start Cycle: 3 End Cycle: 6

Formato sigmoidal
Sinal fluorescente não
informativo

Baseline

O que acontece se o *baseline* for muito alto?



Analysis Settings

Cr Settings for RNase P TAMRA

Cr Settings to Use: ☐ Use Default Settings

☐ Automatic Threshold

Threshold: 0.605352

☐ Automatic Baseline

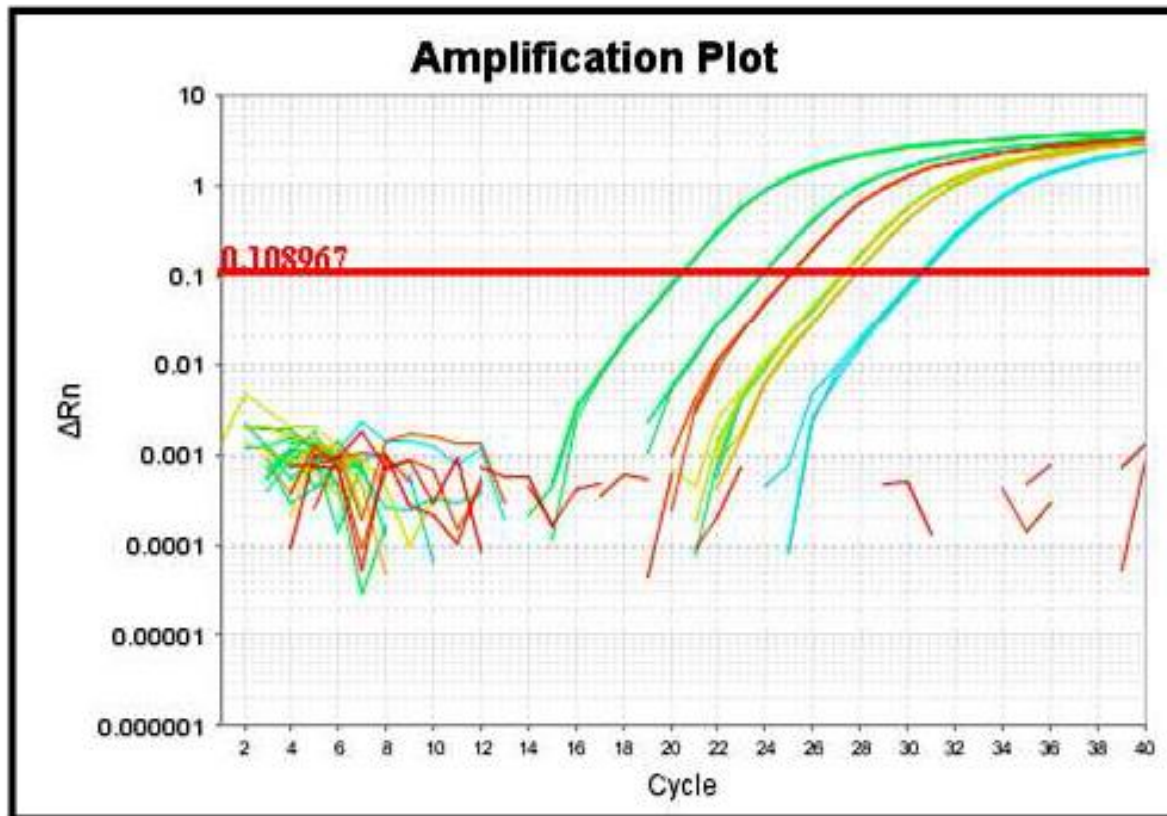
Baseline Start Cycle: 3 End Cycle: 26

Efeito dupla cascata
Perda de sinal
fluorescente informativo

Conceitos chave

- *Threshold*

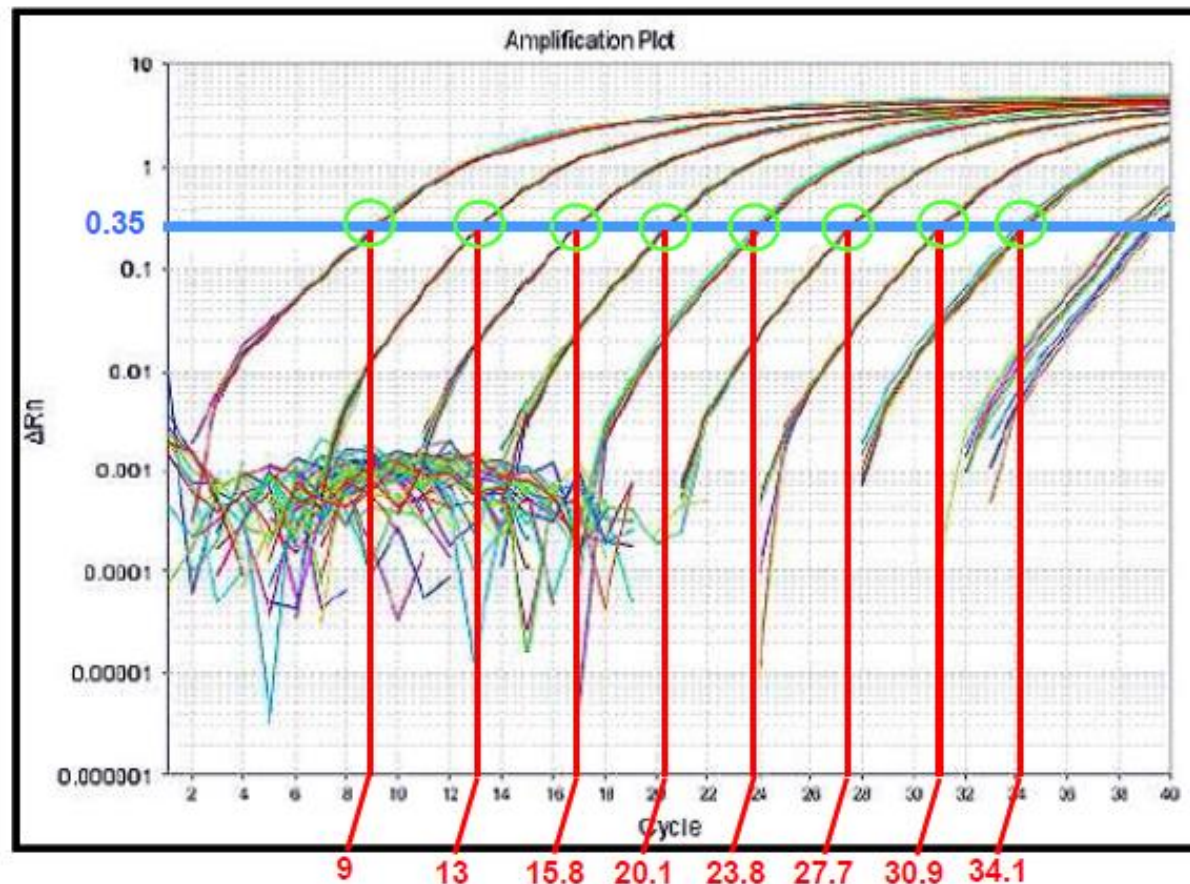
Nível arbitrário de fluorescência estabelecido em cima do *baseline* e dentro da região exponencial



Conceitos chave

- *Threshold cycle (Ct)*

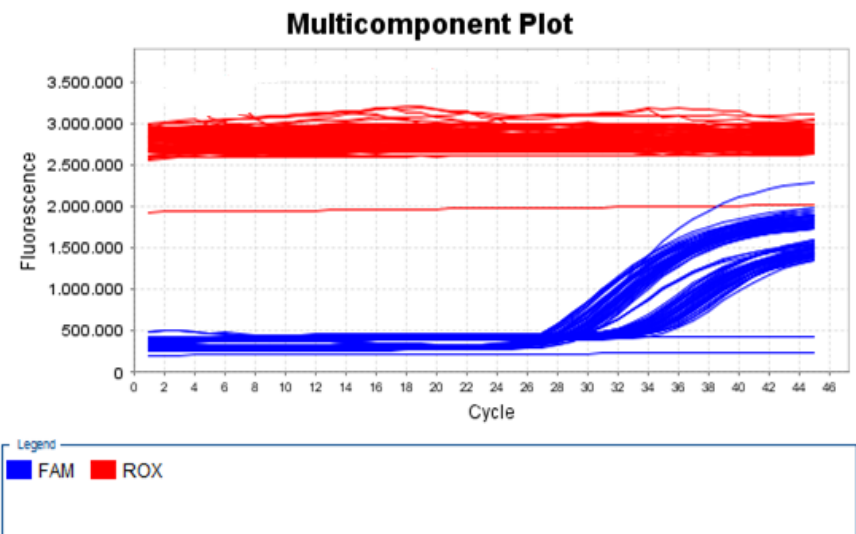
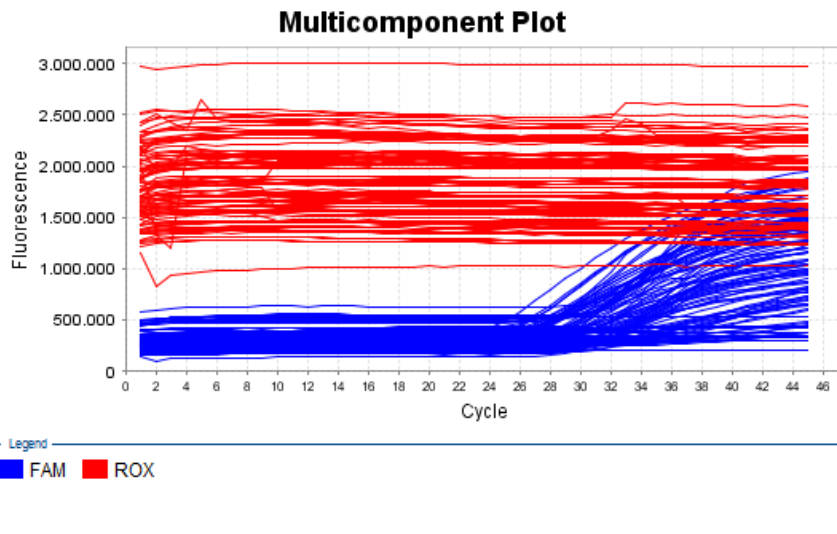
Número de ciclos da PCR no qual a fluorescência atinge o *threshold*



Conceitos chave

Normalizações:

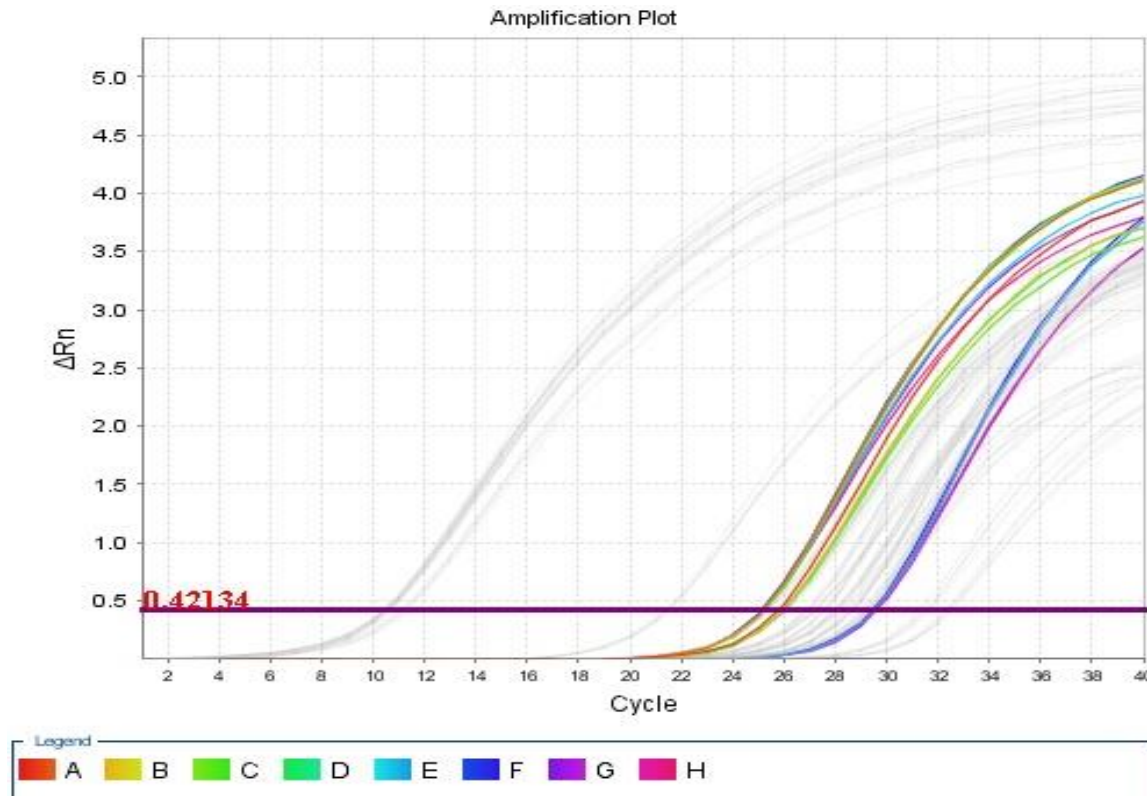
- Reporter normalizado (Rn)
 - Sinal fluorescente normalizado por uma referência passiva
 - ROX – fluorescência presente no meio (correção de volumes e diferenças de detecção na placa)
 - Fluorescência do gene alvo/fluorescência do ROX



Conceitos chave

Normalizações

- ΔRn : Reporter normalizado corrigido pelo *baseline*



Conceitos chave

Eficiência

- E= 100% tem que garantir que a amplificação dobra a cada ciclo

$$\textit{Template inicial} = \frac{\textit{Produto Final}}{(1 + \textit{Eficiência})^n \text{ de ciclos}}$$

$$P = T (1 + E)^n$$

Conceitos chave

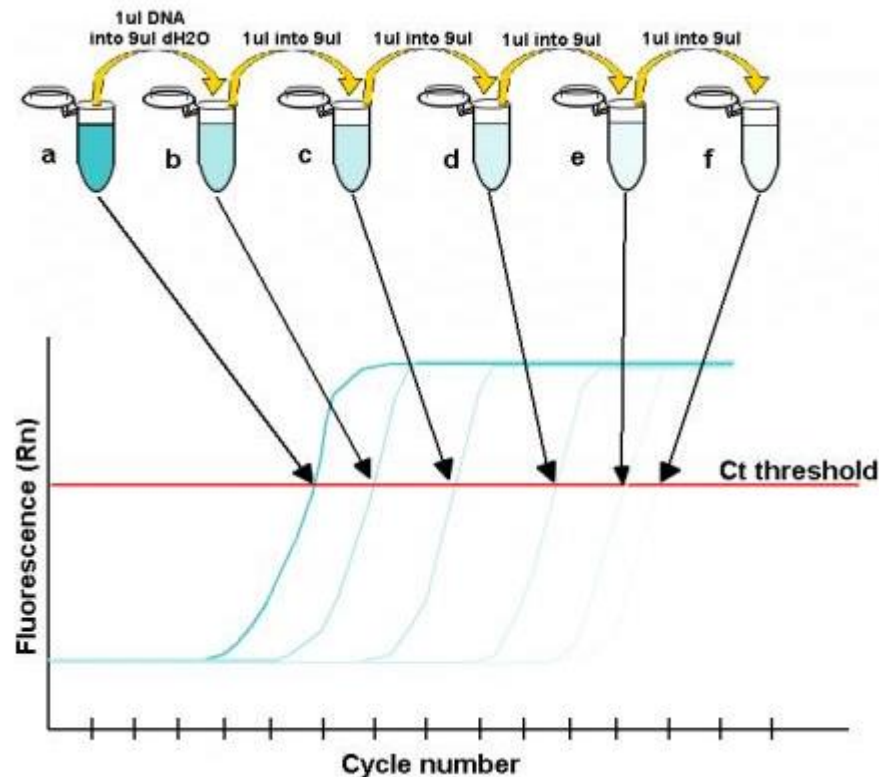
Eficiência

Pode variar de acordo com:

- O preparo da reação
- A qualidade do *templates*
- A presença de inibidores
- O desenho dos *primers*
- As condições de ciclagem
- A qualidade dos reagentes
- A concentração dos reagentes
- O tamanho do *amplicon*

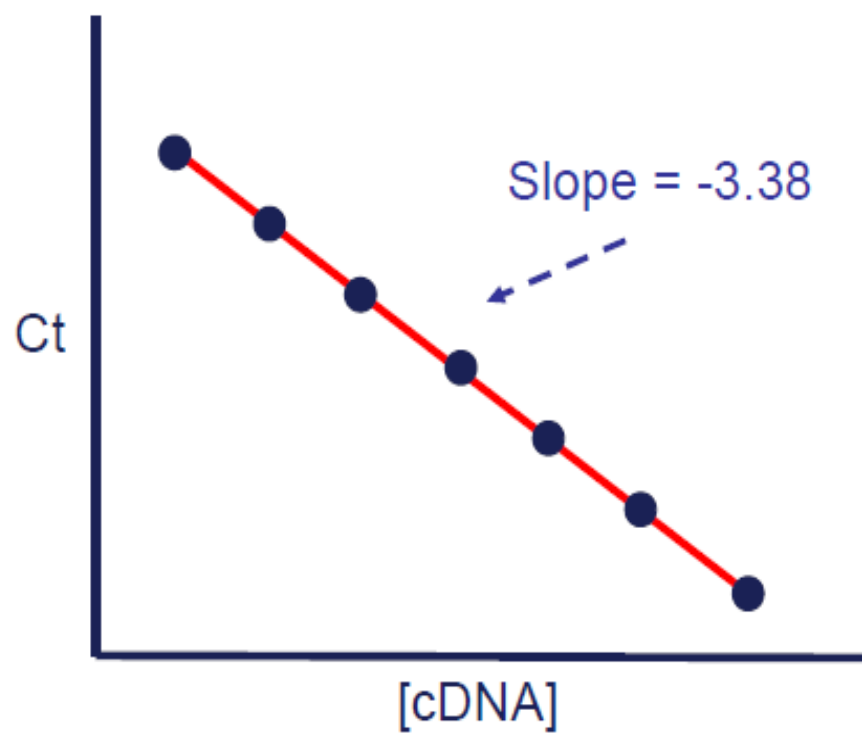
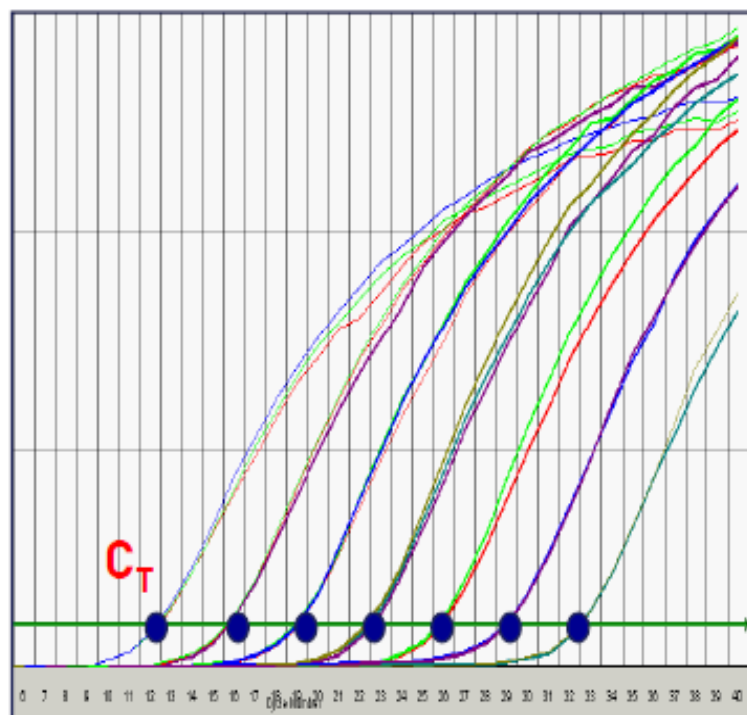
Eficiência da Reação - Curva padrão

- Curva padrão para calcular a eficiência da reação
- Diluições seriadas 1:10, 1:5, 1:2



$\Delta \Delta C_t$

Eficiência de PCR



$\Delta \Delta C_t$

Eficiência de PCR

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

- E equivale ao coeficiente de regressão R^2
- Calculado com base no coeficiente angular da reta (*slope*) de acordo com a equação da regressão linear da curva padrão
- Proximidade entre os pontos individuais de C_t e a reta
- Varia entre 0-1, quanto mais perto de 1, mais ajustado

$\Delta \Delta Ct$

Eficiência de PCR

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

$$E = 10^{(-1/-3,33)} - 1$$

$$E = 10^{0,30} - 1$$

$$E = 1,995 - 1$$

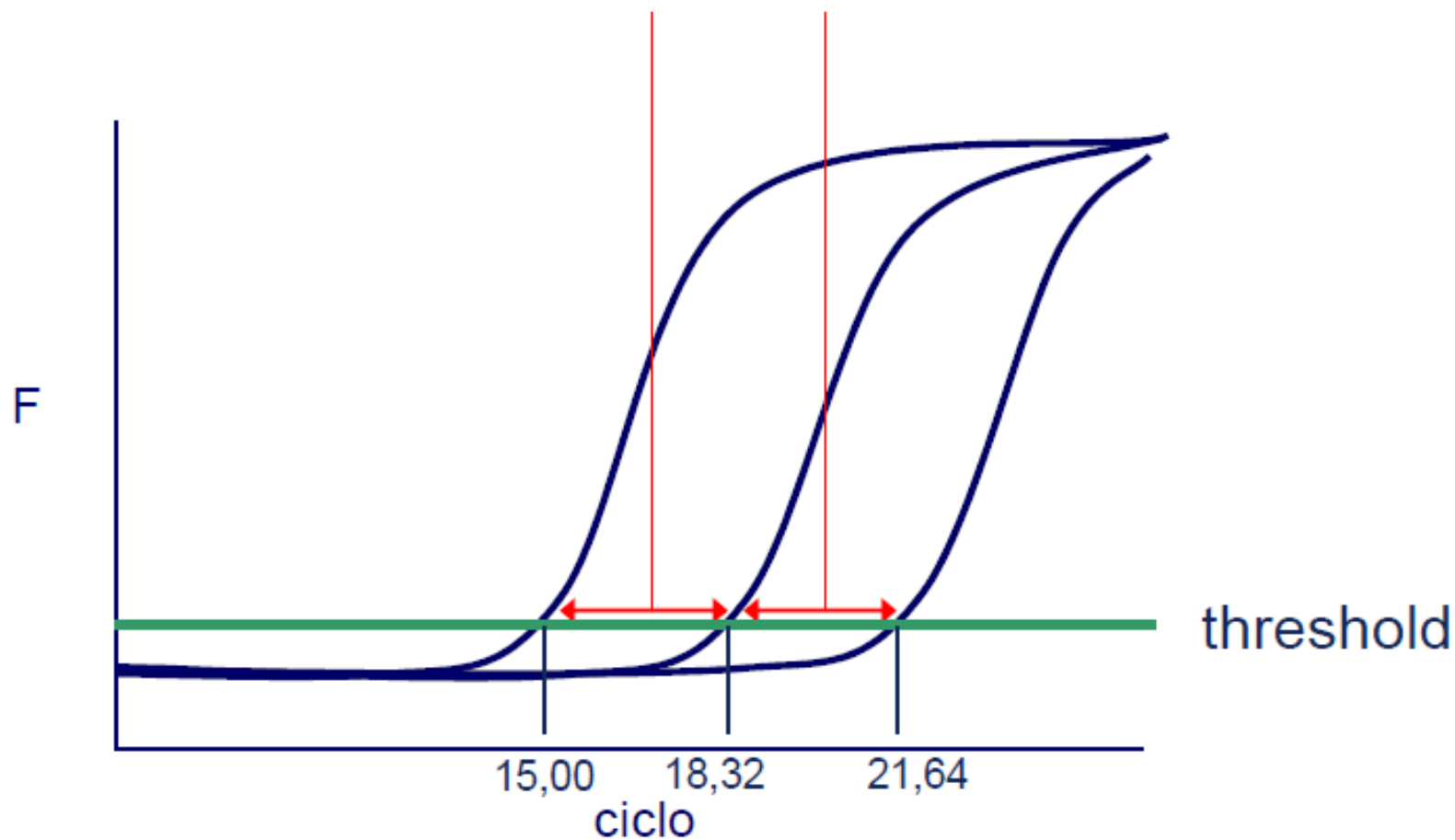
$$E = 0,995 \text{ ou } 99,5\%$$

- $E=100\%$ - slope de -3,32
- Mais negativos (Ex: -3,7) $E < 100\%$
- Mais positivos (Ex: -2,9) $E > 100\%$ (pipetagem, qualidade da amostra)

$\Delta \Delta C_t$

Eficiência de PCR

100% eficiência: diluição 1:10, 3.32 ciclos entre cada diluição



Gene de referência/endógeno

- **Utilizado para corrigir variações relacionadas à:**
 - Quantidade de tecido utilizado na extração (número de células total)
 - Eficiência de extração de RNA (por exemplo, em tecidos diferentes)
 - Eficiência da transcrição reversa (RT) e da amplificação (PCR)
 - Quantificação de RNA/cDNA inicial
 - Pipetagem de RNA/cDNA nas reações
 - Degradação de RNA/cDNA

Gene de referência

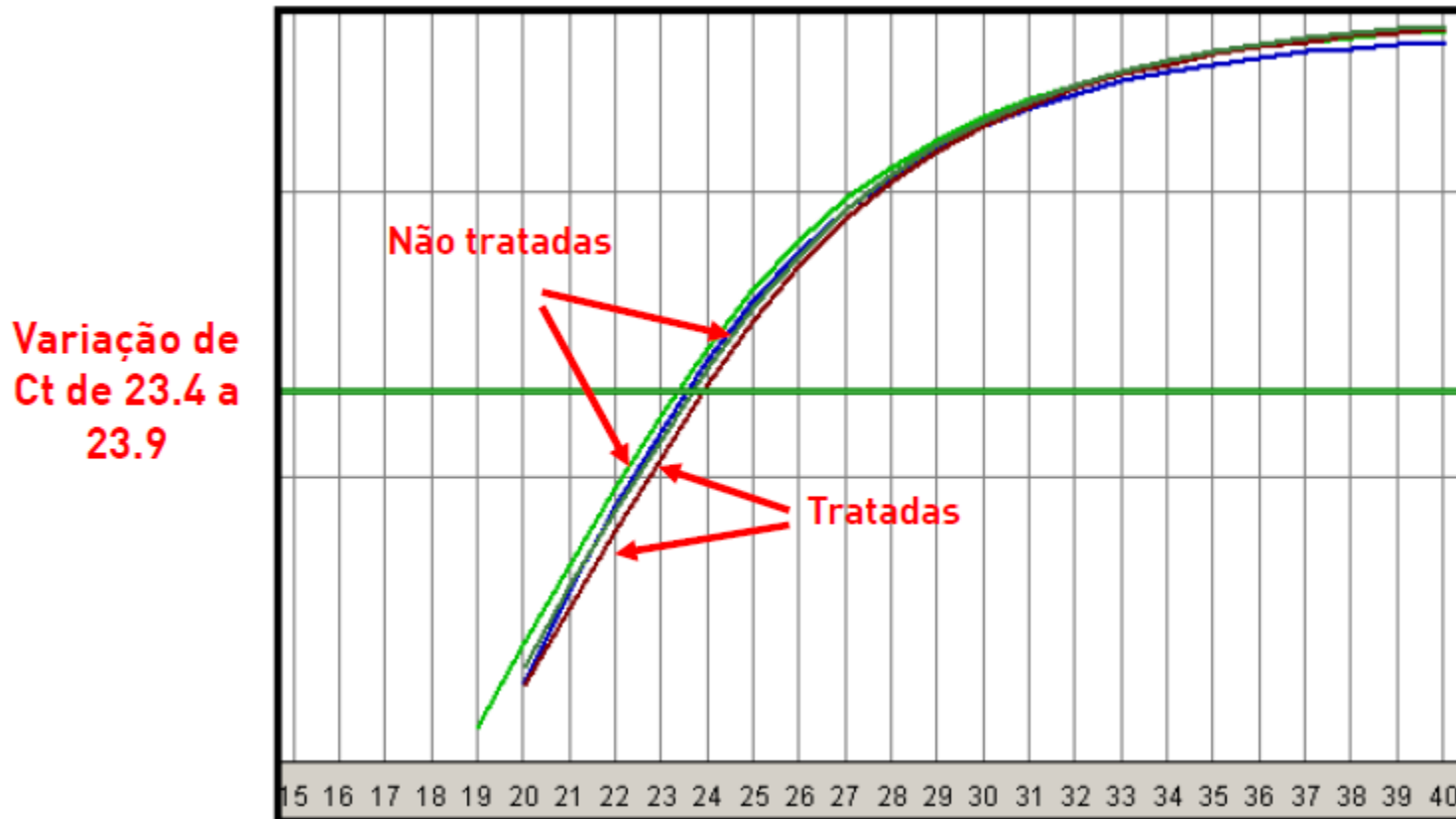
Qual é o gene de referência ideal?

$$\frac{\text{Expressão no Tratado}}{\text{Expressão no Controle}} = \sim 1$$

Aquele cuja expressão gênica não tenha grandes
variações entre as diferentes condições
experimentais!

Gene Referência

- Amostras tratadas e não tratadas: amplificação do gene referência



Métodos para Cálculo da Quantificação Relativa

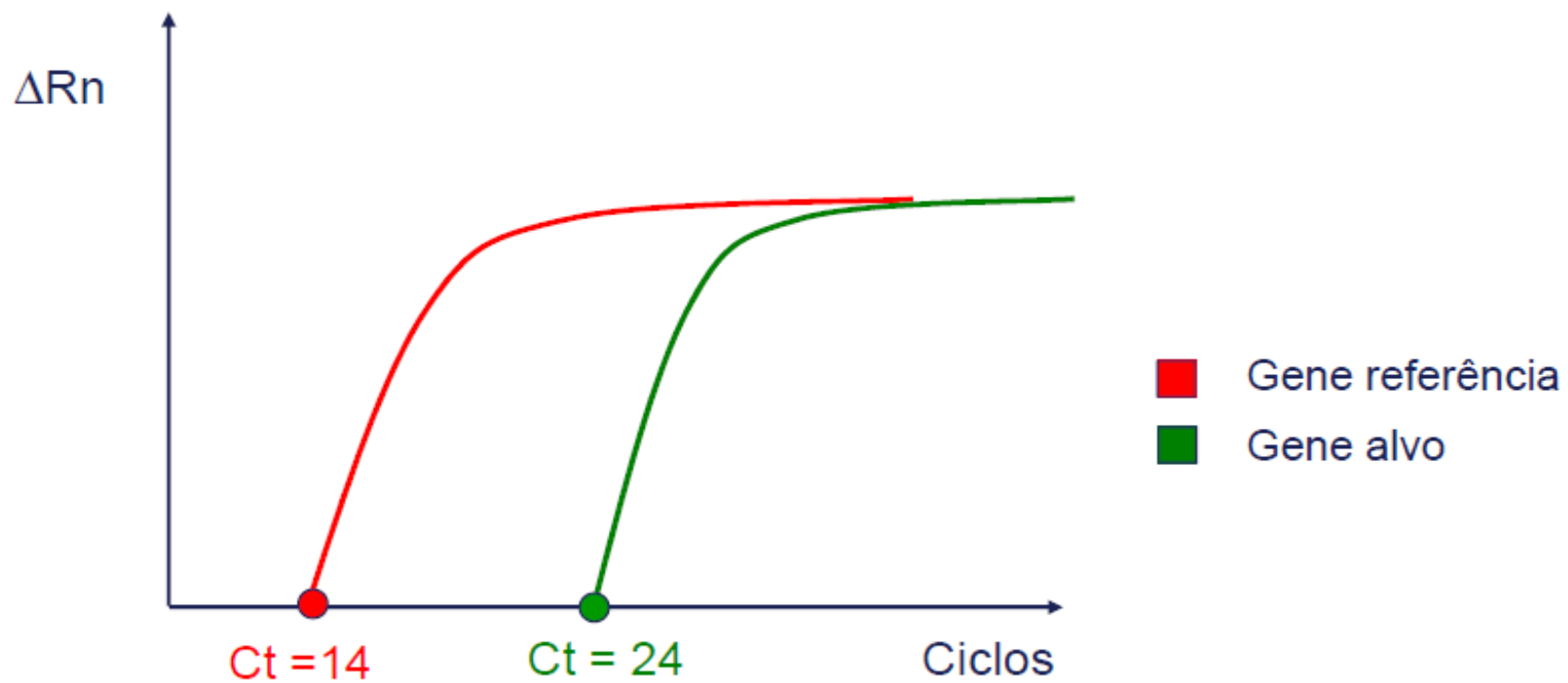
- **Ct Comparativo ($\Delta\Delta Ct$)** – Preparo mais simples pois não requer curva padrão. É necessária a validação das eficiências dos ensaios do gene alvo e endógeno, que devem ser semelhantes
- **Curva Padrão Relativa** – Não requer validação da eficiência, nem que a eficiência dos ensaios do alvo e do endógeno sejam semelhantes. Requer construção de curva padrão, portanto uso de mais reagentes e espaço na placa

$\Delta \Delta C_t$

Validação do Método do C_T Comparativo ($\Delta \Delta C_t$)

Gene Alvo - Gene Referência	Eficiência* <small>* Tolerância de +- 10%</small>	Cálculo
Alvo = GR	100 %	$2^{-\Delta \Delta C_t}$
Alvo = GR	<100%	$(1 + E)^{-\Delta \Delta C_t}$
Alvo \neq GR	NA	Curva Padrão Relativa

$\Delta \Delta Ct$



$$\Delta Ct = 24 - 14 = 10$$

$\Delta \Delta C_t$

	c-myc <i>Ct ± desvio</i>	GAPDH <i>Tecido</i> <i>Ct ± desvio</i>	ΔC_t <i>c-myc - GAPDH</i>	$\Delta \Delta C_t$ $\Delta C_t - \Delta C_t$	$2^{-\Delta \Delta C_t}$ <i>Rel. ao cérebro</i>
Cérebro	30.49 ± 0.15	23.63 ± 0.09	6.86 ± 0.17	0.00 ± 0.17	1.0 ± 0.11
Rim	27.03 ± 0.06	22.66 ± 0.08	4.37 ± 0.10	-2.50 ± 0.10	5.6 ± 0.32
Fígado	26.25 ± 0.07	24.60 ± 0.07	1.65 ± 0.10	-5.21 ± 0.10	37.0 ± 2.52
Pulmão	25.83 ± 0.07	23.01 ± 0.07	2.81 ± 0.10	-4.05 ± 0.10	16.5 ± 1.10

$$\Delta C_T = C_T (\text{alvo}) - C_T (\text{gene de referência})$$

$$\Delta \Delta C_T = \Delta C_T (\text{amostra}) - \Delta C_T (\text{calibrador})$$

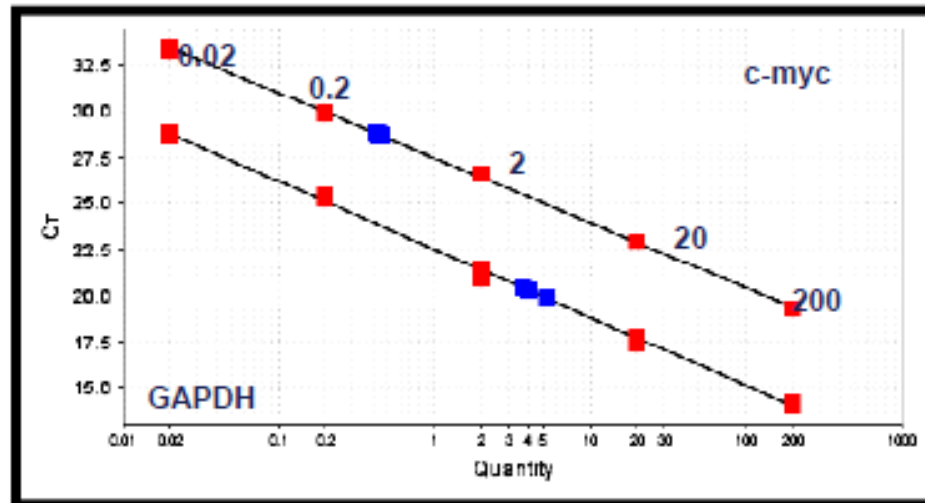
$$\text{Quantidade Relativa} = 2^{-\Delta \Delta C_t}$$

$\Delta \Delta C_t$

- Vantagem:
 - > Não há necessidade de utilizar curva padrão em toda placa.
 - > Redução na quantidade de reagentes.
- Considerações:
 - > Necessita ser validado.
 - > Eficiência de amplificação do alvo (gene de interesse) e da referência interna (gene de referência) deve ser igual (dp = $\pm 10\%$).
 - > Taqman Gene Expression Assays possuem eficiência próxima de 100%
 - > Ideal para um grande número de alvos e/ou amostras ou para validação de resultados de microarray.

Curva Padrão Relativa

- Necessita de curva padrão do gene alvo e do endógeno a cada placa
- Gasta mais reagentes, amostra e \$
- Ideal para estudos com número amostral pequeno
- Não necessita testar a eficiência dos ensaios
- Resultados mais precisos



Curva Padrão Relativa

<i>Tecido</i>	c-myc <i>Valor Arbitrário</i>	GAPDH <i>Valor Arbitrário</i>	c-myc_N <i>Norm. c/ GAPDH</i>	c-myc_N <i>Rel. ao cérebro</i>
Cérebro	0.039 ± 0.004	0.54 ± 0.034	0.07 ± 0.008	1.0 ± 0.12
Rim	0.41 ± 0.016	1.02 ± 0.052	0.40 ± 0.025	5.5 ± 0.35
Fígado	0.70 ± 0.036	0.28 ± 0.013	2.49 ± 0.173	34.2 ± 2.37
Pulmão	0.93 ± 0.044	0.81 ± 0.041	1.15 ± 0.079	15.7 ± 1.09

Gene alvo

Gene de referência

Amostra alvo normalizada

Amostra calibradora normalizada

Análise estatística dos dados

- Verificar a distribuição dos dados:
 - Eles apresentam distribuição normal?
- Existem *outliers* entre as amostras?
 - Calcular zscore e excluí-los?
- Comparação entre grupos:
 - Teste-t, Anova, Mann-Whitney...
- Qual variável usar?
 - RQ? $2\text{-}\Delta\text{CT}$? $\Delta\Delta\text{CT}$?



1: Amostra A100

	Amostra	Idade
1	A100	6
2	A103	6
3	A105	7
4	A107	8
5	A110	5
6	A115	8
7	A116	6
8	A120	6
9	A121	8
10	A127	7
11	A128	7
12	A137	7
13	A138	8
14	A139	7
15	A140	7
16	A141	7
17	A142	6
18	A143	8
19	A144	8
20	A145	7
21	A146	74
22	A147	85
23	A148	75

- Reports
- Descriptive Statistics**
- Tables
- Compare Means
- General Linear Model
- Generalized Linear Models
- Mixed Models
- Correlate
- Regression
- Loglinear
- Neural Networks
- Classify
- Dimension Reduction
- Scale
- Nonparametric Tests
- Forecasting
- Survival
- Multiple Response
- Missing Value Analysis...
- Multiple Imputation
- Complex Samples
- Quality Control
- ROC Curve...

- 123 Frequencies...
- Pd Descriptives...
- Explore...**
- Crosstabs...
- Ratio...
- P-P Plots...
- Q-Q Plots...



po	Idade	Bencao_Cerebral	Gene_Nome	Média_CTs
2		999	HSPA8	28,875
2		999	HSPA8	28,256
2		999	HSPA8	27,237
2		999	HSPA8	26,087
2		999	HSPA8	27,903
1	2	999	HSPA8	28,362
1	2	999	HSPA8	29,116
1	2	999	HSPA8	26,424
1	2	999	HSPA8	26,349
1	2	999	HSPA8	27,419
1	2	999	HSPA8	27,773
1	2	999	HSPA8	28,030
1	2	999	HSPA8	27,765
1	2	999	HSPA8	27,403
1	2	999	HSPA8	27,110
1	2	999	HSPA8	26,931
1	2	999	HSPA8	26,702
1	2	999	HSPA8	27,209
1	2	999	HSPA8	27,806
1	2	999	HSPA8	28,716
1	2	999	HSPA8	27,578
1	2	999	HSPA8	27,173
1	2	999	HSPA8	30,821



1: Amostra

A100

	Amos
1	A100
2	A103
3	A105
4	A107
5	A110
6	A115
7	A116
8	A120
9	A121
10	A127
11	A128
12	A137
13	A138
14	A139
15	A140
16	A141
17	A142
18	A143
19	A144
20	A145
21	A146
22	A147
23	A148

Explore: Plots

Boxplots

- ☒ Factor levels together
- ☐ Dependents together
- ☐ None

Descriptive

- ☒ Stem-and-leaf
- ☒ Histogram

☒ Normality plots with tests

Spread vs Level with Levene Test

- ☐ None
- ☐ Power estimation
- ☐ Transformed Power: Natural log
- ☐ Untransformed

Continue Cancel Help

decido	Gene	Tipo_Extração	Regiao_Cerebral	Gene_Nome	Média_CTs	C
1	2	2	999	HSPA8	28,8750	
1	2	2	999	HSPA8	28,2565	
1	2	2	999	HSPA8	27,2376	
					0874	
					9030	
					3628	
					1164	
					4244	
					3491	
					4196	
					7736	
					0304	
					7657	
					4031	
					1107	
					9312	
					7020	
					2099	
	2	2	999	HSPA8	27,8061	
1	2	2	999	HSPA8	28,7165	
1	2	2	999	HSPA8	27,5788	
1	2	2	999	HSPA8	27,1737	
1	2	2	999	HSPA8	30,8218	

Explore

Dependent List: dois_DCT

Factor List:

Label Cases by:

Display: ☒ Both ☐ Statistics ☐ Plots

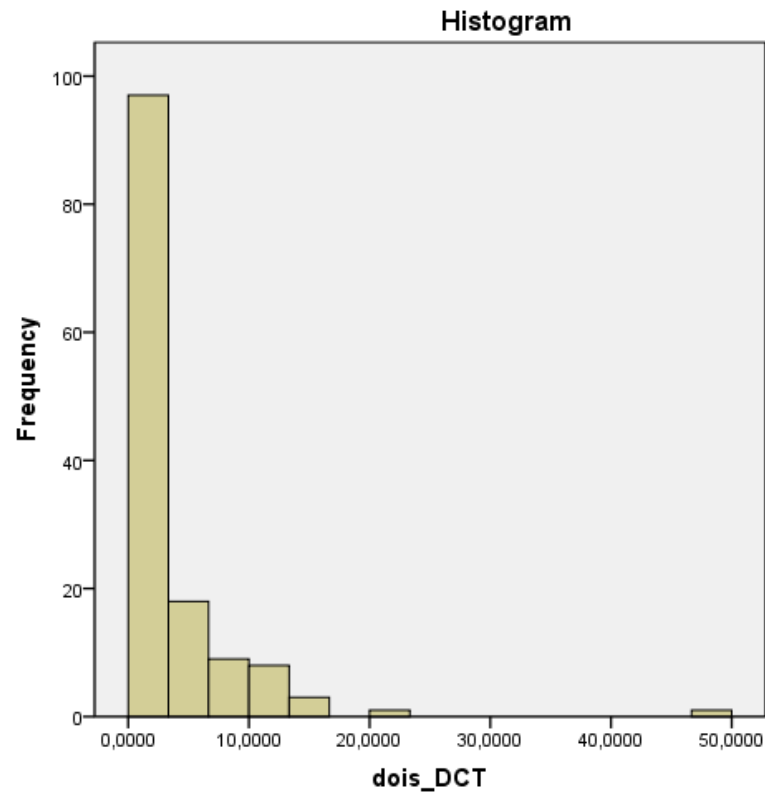
OK Paste Reset Cancel Help

Distribuição

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
dois_DCT	,274	137	,000	,559	137	,000

a. Lilliefors Significance Correction



	po	Tecido	C
1	1	1	
2	1	1	
3	1	1	
4	1	1	
5	1	1	
6	1	1	
7	1	1	
8	1	1	
9	1	1	
10	1	1	
11	1	1	
12	1	1	
13	1	1	
14	1	1	
15	1	1	
16	1	1	
17	1	1	
18	1	1	
19	1	1	
20	1	1	
21	1	1	
22	1	1	
23	1	1	

		Media_CT	Ct_B_Actina	DCT	dois_DCT	sangu
Compare Means	Explore...	28,8750	28,5367	,3383	1,2643	
General Linear Model	Crosstabs...	28,2565	27,7403	,5161	1,4301	
Generalized Linear Models	Ratio...	27,2376	27,0597	,1779	1,1313	
Mixed Models	P-P Plots...	26,0874	27,5968	-1,5094	,3512	
Correlate	Q-Q Plots...	27,9030	28,7008	-,7978	,5752	
Regression	999 HSPA8	28,3628	29,1938	-,8310	,5622	
Loglinear	999 HSPA8	29,1164	28,3979	,7185	1,6455	
Neural Networks	999 HSPA8	26,4244	25,8127	,6117	1,5281	
Classify	999 HSPA8	26,3491	27,9729	-1,6238	,3245	
Dimension Reduction	999 HSPA8	27,4196	28,3757	-,9561	,5155	
Scale	999 HSPA8	27,7736	27,1956	,5781	1,4928	
Nonparametric Tests	999 HSPA8	28,0304	26,4956	1,5348	2,8975	
Forecasting	999 HSPA8	27,7657	27,9710	-,2053	,8673	
Survival	999 HSPA8	27,4031	27,0410	,3621	1,2853	
Multiple Response	999 HSPA8	27,1107	26,6239	,4868	1,4013	
Missing Value Analysis...	999 HSPA8	26,9312	26,7134	,2177	1,1629	
Multiple Imputation	999 HSPA8	26,7020	27,9439	-1,2419	,4228	
Complex Samples	999 HSPA8	27,2099	27,9723	-,7625	,5895	
Quality Control	999 HSPA8	27,8061	27,3492	,4570	1,3726	
ROC Curve...	999 HSPA8	28,7165	28,5972	,1193	1,0862	
2	2	999 HSPA8	27,5788	25,7370	1,8417	3,5844
2	2	999 HSPA8	27,1737	27,8417	-,6680	,6294
2	2	999 HSPA8	30,8218	31,5305	-,7087	,6119

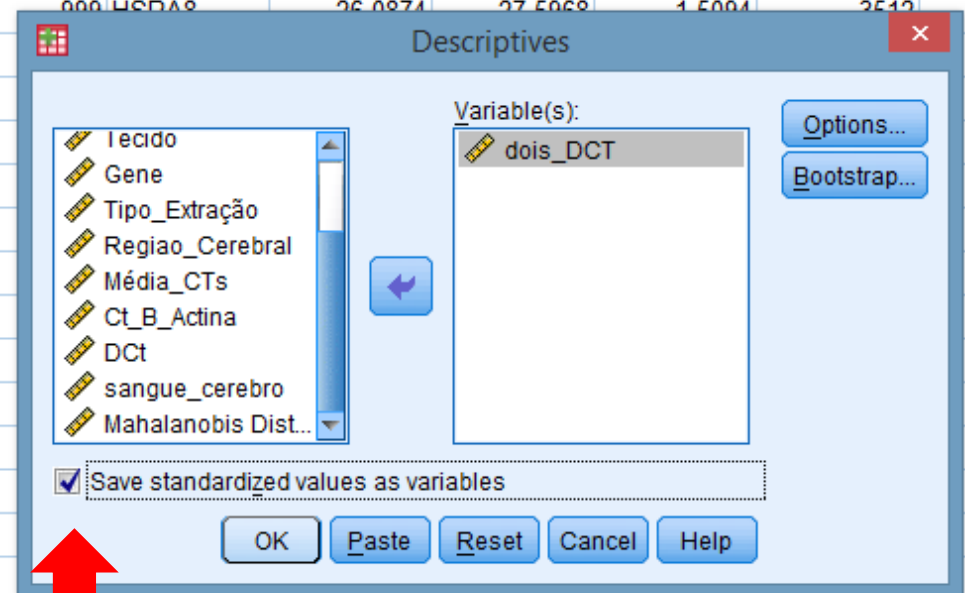


File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help



1: dois_DCT 1,2643

	po	Tecido	Gene	Tipo_Extração	Regiao_Cerebral	Gene_Nome	Média_CTs	Ct_B_Actina	DCt	dois_DCT	sangue
1	1	1	2	2	999	HSPA8	28,8750	28,5367	,3383	1,2643	
2	1	1	2	2	999	HSPA8	28,2565	27,7403	,5161	1,4301	
3	1	1	2	2	999	HSPA8	27,2376	27,0597	,1779	1,1313	
4	1	1	2	2	999	HSPA8	26,0874	27,5068	1,5004	2,512	
5	1	1	2	2							
6	1	1	2	2							
7	1	1	2	2							
8	1	1	2	2							
9	1	1	2	2							
10	1	1	2	2							
11	1	1	2	2							
12	1	1	2	2							
13	1	1	2	2							
14	1	1	2	2							
15	1	1	2	2							
16	1	1	2	2							
17	1	1	2	2							
18	1	1	2	2	999	HSPA8	27,2099	27,9723	-,7625	,5895	
19	1	1	2	2	999	HSPA8	27,8061	27,3492	,4570	1,3726	
20	1	1	2	2	999	HSPA8	28,7165	28,5972	,1193	1,0862	
21	1	1	2	2	999	HSPA8	27,5788	25,7370	1,8417	3,5844	
22	1	1	2	2	999	HSPA8	27,1737	27,8417	-,6680	,6294	
23	1	1	2	2	999	HSPA8	30,8218	31,5305	-,7087	,6119	





1: Zdois_DCT -38960439669266 Visible: 15 of 23

	Grupo	Tecido	Gene	Tipo_Extração	Regiao_Cerebral	Gene_Nome	Média_CTs	Ct_B_Actina	DCt	dois_DCT	sa	Zdois_DCT	va
1	2	1	1	2	2	999 HSPA8	28,8750	28,5367	,3383	1,2643	4	-,38960	
2	2	1	1	2	2	999 HSPA8	28,2565	27,7403	,5161	1,4301	4	-,35944	
3	1	1	1	2	2	999 HSPA8	27,2376	27,0597	,1779	1,1313	4	-,41380	
4	2	1	1	2	2	999 HSPA8	26,0874	27,5968	-1,5094	,3512	4	-,55572	
5	2	1	1	2	2	999 HSPA8	27,9030	28,7008	-,7978	,5752	4	-,51497	
6	2	1	1	2	2	999 HSPA8	28,3628	29,1938	-,8310	,5622	4	-,51733	
7	2	1	1	2	2	999 HSPA8	29,1164	28,3979	,7185	1,6455	4	-,32026	
8	2	1	1	2	2	999 HSPA8	26,4244	25,8127	,6117	1,5281	4	-,34161	
9	2	1	1	2	2	999 HSPA8	26,3491	27,9729	-1,6238	,3245	4	-,56058	
10	1	1	1	2	2	999 HSPA8	27,4196	28,3757	-,9561	,5155	4	-,52583	
11	2	1	1	2	2	999 HSPA8	27,7736	27,1956	,5781	1,4928	4	-,34804	
12	2	1	1	2	2	999 HSPA8	28,0304	26,4956	1,5348	2,8975	4	-,09249	
13	2	1	1	2	2	999 HSPA8	27,7657	27,9710	-,2053	,8673	4	-,46183	
14	1	1	1	2	2	999 HSPA8	27,4031	27,0410	,3621	1,2853	4	-,38578	
15	2	1	1	2	2	999 HSPA8	27,1107	26,6239	,4868	1,4013	4	-,36468	
16	1	1	1	2	2	999 HSPA8	26,9312	26,7134	,2177	1,1629	4	-,40805	
17	2	1	1	2	2	999 HSPA8	26,7020	27,9439	-1,2419	,4228	4	-,54269	
18	1	1	1	2	2	999 HSPA8	27,2099	27,9723	-,7625	,5895	4	-,51237	
19	2	1	1	2	2	999 HSPA8	27,8061	27,3492	,4570	1,3726	4	-,36990	
20	1	1	1	2	2	999 HSPA8	28,7165	28,5972	,1193	1,0862	4	-,42200	
21	2	1	1	2	2	999 HSPA8	27,5788	25,7370	1,8417	3,5844	4	,03247	
22	1	1	1	2	2	999 HSPA8	27,1737	27,8417	-,6680	,6294	4	-,50511	
23	1	1	1	2	2	999 HSPA8	30,8218	31,5305	-,7087	,6119	4	-,50829	

Outliers

