

## Documento Complementar - CGAL

Delineamentos experimentais para validação e verificação de métodos para detecção de sequência específica por PCR em tempo real

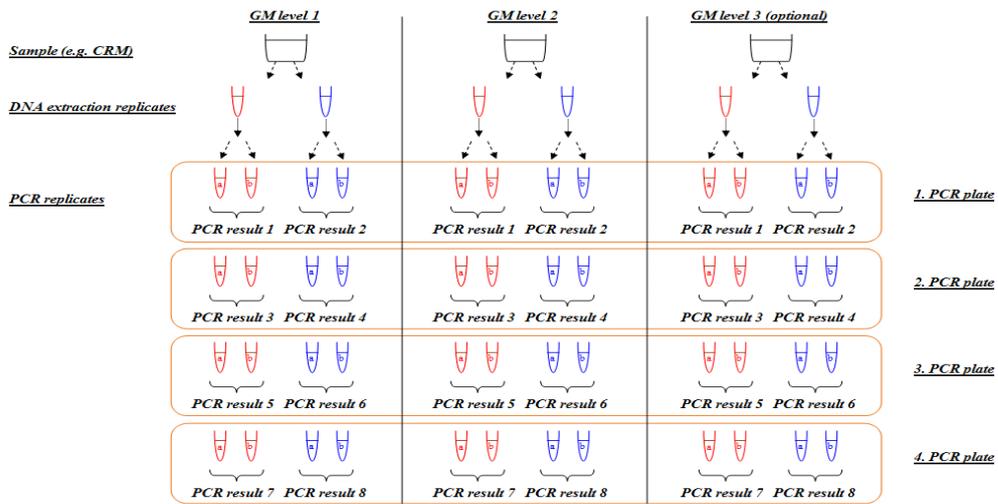
**Tabela 1 – Exemplo de configurações práticas para a validação e verificação de um método quantitativo em PCR em tempo real.**

<b>1. Opcional: Teste preliminar que define a concentração apropriada de DNA</b>	Testar pelo menos três concentrações alvo na faixa de 200 ng a 0,1 ng (dependendo da espécie), por exemplo, 200 ng de DNA de soja corresponde a aproximadamente 176.991 cópias do gene endógeno uma vez que uma cópia corresponde a 1,33 pc de DNA genômico.
<b>2. Faixa dinâmica, R<sup>2</sup> e Eficiência de amplificação</b>	<p><b>Exemplo 1:</b> Requerimentos mínimos para 2 curvas de calibração, 5 pontos em triplicatas. Todas as inclinações (slope) devem estar na faixa de -3,6 a -3,1 e todos os valores de R<sup>2</sup> devem ser maiores que 0,98.</p> <p><b>Exemplo 2:</b> Requerimentos mínimos para 4 curvas de calibração; 5 pontos em duplicatas. A média de 4 inclinações (slope) e R<sup>2</sup> é utilizada para verificar os critérios de aceitação.</p> <p><b>Exemplo 3:</b> Requerimentos mínimos para 2 curvas de calibração; 8 pontos de calibração em 5 repetições cobrindo baixas concentrações para o LOD e LOQ. A média da parte acima do LOQ para inclinação (slope) e R<sup>2</sup> é utilizada para verificação dos critérios de aceitação.</p>
<b>3. Exatidão, precisão e Desvio padrão de repetibilidade (RSDr)</b>	<p>Pelo menos dois níveis (um próximo do limite de rotulagem e outro próximo do LOQ, recomenda-se um terceiro nível próximo ao ponto superior da faixa dinâmica)</p> <p><b>Exemplo 1:</b> 2 repetições de extração de DNA por nível, 2 repetições de PCR por extração/placa, 4 placas resultando em 16 resultados e 8 estimativas por nível* (Fig. 2)</p> <p><b>Exemplo 2:</b> 2 repetições de extração de DNA por nível, 4 repetições de PCR por extração/placa, duas placas resultando em 16 resultados e 4 estimativas por nível* (Fig. 3)</p> <p>Para estimação da precisão intermediária as reações de PCR devem ser conduzidas em pelo menos dois dias diferentes pelo mesmo operador ou, se possível por um operador adicional.</p>
<b>4. LOQ, LOD</b>	<p>LOQ: 10 repetições de PCR em baixa concentração (p. ex. 80, 60, 40, 20, 10, 5, e 1 cópia). O LOQ é a menor concentração da série onde o desvio padrão de repetibilidade for menor que 25% e o ponto deve estar coberto pela curva padrão.</p> <p>LOD: 10 repetições de PCR em baixa concentração (p. ex. 20, 10, 5, e 1 cópia). O LOD é então a menor concentração da série onde todas as repetições forem positivas.</p>

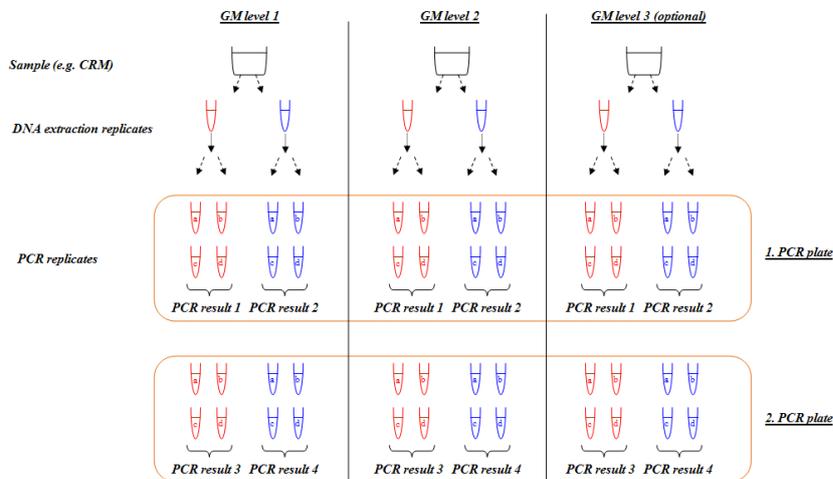
\* Se baseado na experiência, o laboratório conseguir provar que a repetibilidade entre dois analistas experientes é semelhante à obtida por um único analista, não é necessário incluir a repetição feita por outro analista. Do mesmo modo, quando se atender aos critérios pré-estabelecidos após a análise de dados nos quais houve a adição de dados obtidos por um segundo analista, ou em diferentes condições de análises, não se justifica a execução de novo delineamento.

**Tabela 2 – Exemplo de configurações práticas para a validação e verificação de um método qualitativo em PCR em tempo real.**

<p><b>1. Faixa dinâmica, R2, coeficiente e Eficiência de amplificação</b></p>	<p>Opcional.</p>
<p><b>2. LOD</b></p>	<p>Exemplo 1: Teste 10 repetições de PCR em torno do LODabs esperado (Ex. diluição serial com 20, 10, 5, 3 cópias e 1 cópia/reação) *. O LOD é o menor nível da diluição serial onde todas as repetições forem positivas. Exemplo 2: Testar seis níveis de diluição (Ex. 20, 10, 5, 2, 1 e 0,1 cópias/reação) e 12 repetições de PCR por nível. Com base na curva de modelagem, calcula-se o LODabs no intervalo de 95 % de confiança (ver <a href="https://quodata.de/content/validation-qualitative-pcr-methods-singlelaboratory">https://quodata.de/content/validation-qualitative-pcr-methods-singlelaboratory</a>). Exemplo 3: Testar 60 repetições na concentração do LODabs esperado. Nesta concentração o LOD é verificado em pelo menos 59 repetições positivas. Este teste depende do conhecimento da concentração correta de DNA (número de cópias da sequência alvo por volume de solução medido).</p>
<p><b>3. Especificidade</b></p>	<p>Opcional se já foi acessada no estudo de validação. Para ensaios OGM verificar a especificidade do método utilizando a interface JRC GMO-Matrix. Selecionar os taxon(s) ou GMO(s) específicos e o método de PCR a ser testado e realizar uma simulação <i>in silico</i>. O resultado da simulação é uma lista de eventos GM para os quais a amplificação pelo método pode ser predita ou não. Exemplo para avaliação experimental: Materiais de referência disponíveis para novos OGMs que contém a sequência alvo são testados em duplicata usando pelo menos 100 cópias do DNA alvo por reação por PCR. Para outros métodos, testar a amplificação das sequências alvos em matrizes próximas do ponto de vista taxonômico.</p>
<p><b>4. LODasym (Apenas para métodos multiplex)</b></p>	<p>Pelo menos uma das mais críticas combinações de acordo com os dados da validação deve ser testada. 10 repetições de PCR com baixa concentração da sequência alvo (correspondendo ou próximo ao LOD absoluto) na presença de alta quantidade de outro alvo (por exemplo 25 cópias por reação do gene alvo na presença de <i>background</i> de outro alvo ao nível de 20.000 cópias/reação).</p>



**Figura 1: Delineamento Experimental para Exatidão/Precisão (exemplo 1)**



**Figura 2: Delineamento Experimental para Exatidão/Precisão (exemplo 2)**

Onde se lê sample, entenda-se amostra; extraction: extração, level: nível, plate: placa, result: resultado e optional: opcional.

## Efeito da concentração de DNA no LOD prático

Como demonstrado na Tabela 3, em uma amostra 0,1% há 1000 vezes mais cópias da sequência endógena do que da sequência alvo. Isto implica que para um LOD absoluto do método de 10 cópias, é necessário rodar em uma reação de PCR 10.000 cópias da sequência endógena para ter um LOD prático de 0,1%. Se o LOD absoluto é 10 cópias e roda-se 100.000 cópias na reação de PCR então o LOD prático é 0,01% (ver Tabela 2). O LOD prático deve ser calculado para cada amostra individualmente.

**Tabela 3: Exemplo do efeito da concentração de DNA no LOD prático**

Cópias do gene endógeno	LOD absoluto (cópias do OGM alvo)	LOD prático (%)
100.000	10	0,01
10.000	10	0,1
1.000	10	1

## 1. Teste de Inibição para avaliação do método de extração de DNA

### Avaliação do método de extração de DNA (teste de inibição)

Substâncias conhecidas como componentes inibidores de PCR afetam a eficiência de amplificação do DNA alvo por interação com o molde de DNA ou por interferência com a atividade da DNA polimerase ou ainda por diminuir a eficiência enzimática de co-fatores ( $Mg^{+2}$ ). Os procedimentos de extração de DNA eliminam ou reduzem consideravelmente a quantidade de substâncias inibidoras de PCR. No entanto, a quantidade final de inibidores em uma amostra depende muito da natureza da amostra.

Quanto mais processadas são as amostras, maior é a probabilidade de o DNA estar fragmentado. DNA de plantas também podem ser afetados pela presença de metabólitos secundários tais como polifenóis, óleos e polissacarídeos, os quais podem complexar com as fitas de DNA.

Inibidores podem também ser adicionados ao DNA no processo de isolamento do mesmo, tais como: KCl e NaCl, detergentes iônicos, etanol, isopropanol, fenol, entre outros. Diferentes estratégias podem ser adotadas para testar as preparações de DNA para detectar a presença de compostos inibitórios. O anexo ilustra a aplicação do critério de aceitação da ENGL para avaliar a eficiência da reação (inclinação e  $R^2$ ) de amostras diluídas serialmente de uma amostra não diluída, para testar a presença de compostos inibidores de PCR na amostra não diluída a ser utilizada nas análises de PCR. Basicamente a inibição depende da concentração dos inibidores. Quando o DNA é diluído, o efeito de inibição é reduzido ou eliminado nas

menores concentrações de DNA. A avaliação da eficiência de reação na série diluída e a comparação do Ct teórico de uma amostra não diluída e sem inibição com o Ct medido, fornece informações para verificação da qualidade do DNA para as análises de PCR. No caso, apenas concentrações maiores de DNA demonstra inibição nos exemplos a seguir.

Concentrações menores de DNA podem ser usadas para quantificação, mas isso afeta os LOD e LOQ práticos. No entanto, em certos casos os compostos inibidores aderidos aos fragmentos de DNA podem não ser eliminados pela diluição da amostra, resultando então em um menor número de cópias de DNA disponíveis para amplificação do que o esperado pela concentração nominal de DNA na amostra.

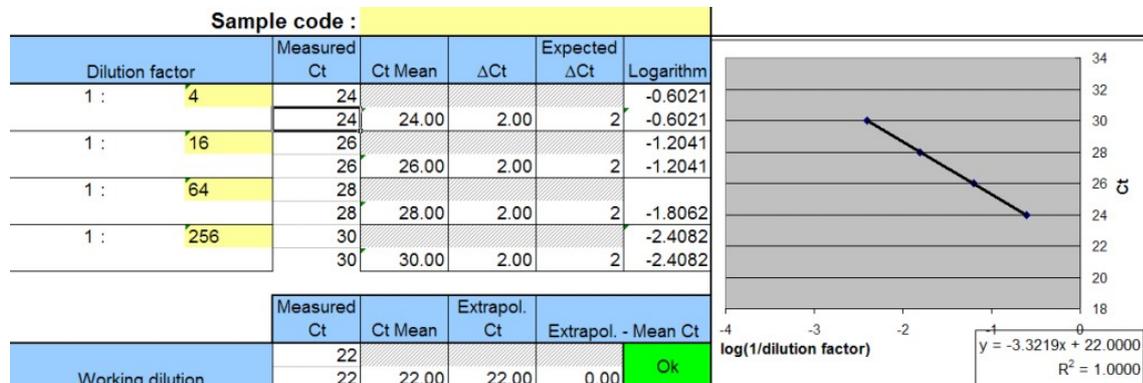
### **Procedimento**

A qualidade do DNA (relativa à ausência de compostos inibidores de PCR) pode ser demonstrada por análise de duas repetições de PCR usando uma diluição serial de quatro pontos (por exemplo 1:4, 1:16, 1:64 e 1:256) de cada repetição de extração de DNA (corrida de inibição) usando o sistema de referência da espécie (gene endógeno). A solução de DNA é ajustada para o nível correspondente à maior concentração a ser utilizada na reação de PCR subsequente, chamada amostra não diluída (diluição de trabalho). A partir desta primeira amostra, prepara-se a diluição serial na base quatro (de 1:4 a 1:256). Para verificar a presença de inibidores, os valores de Ct das quatro amostras diluídas serialmente são “plotados” contra o logaritmo do fator de diluição e calcula-se uma equação por regressão linear. O valor de Ct da amostra não diluída calculado a partir da regressão linear é comparado com o valor de Ct medido para a mesma amostra. Para ser considerada pura deve-se considerar os seguintes critérios de aceitabilidade: a inclinação da regressão (slope) deve estar entre -3,6 a -3,1 o coeficiente de determinação ( $R^2$ ) maior que 0,98 e a diferença entre o Ct medido e o Ct calculado menor que 0,5.

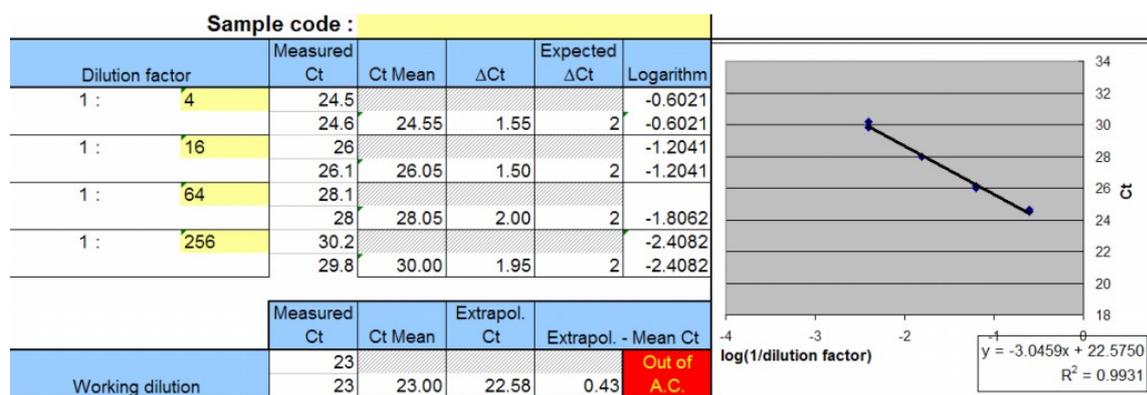
São necessários os valores de Ct de duas repetições de PCR para cada diluição para os cálculos ilustrados a seguir.

## Avaliação da qualidade do DNA

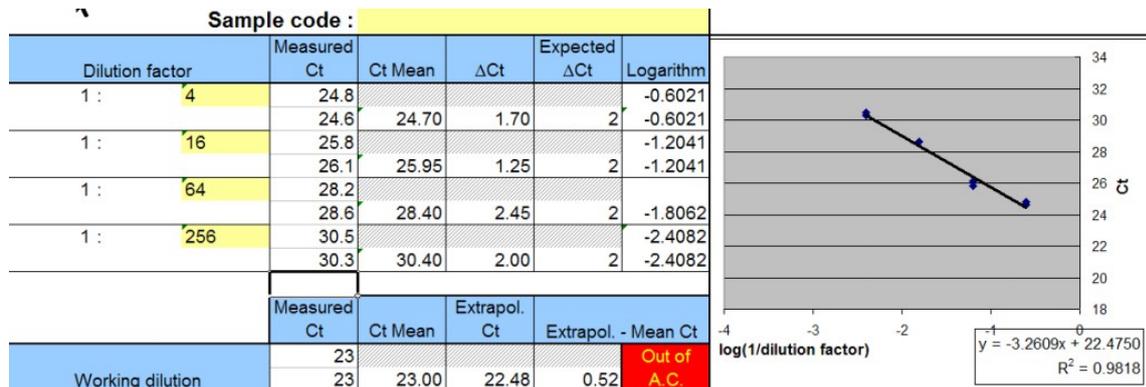
Nas figuras abaixo a expressão “Diluição de trabalho” (*working dilution*) significa “amostra não diluída”



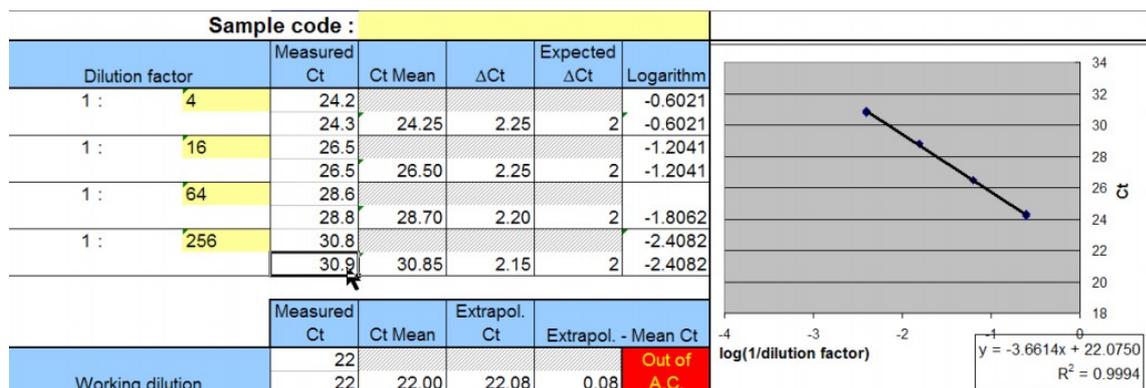
Exemplo A: Quando a qualidade do DNA é aceitável: todos os critérios são atendidos



Exemplo B: DNA inibido. Embora o  $\Delta$ Ct não tenha excedido o limite de 0,5 (mesmo que tenha ficado próximo deste valor, a baixa qualidade do DNA pode ser demonstrada pelo atraso no início da reação (Cts) para a amostra não diluída e para a diluída 1:4. O atraso afeta a inclinação (slope) expressa pela diluição serial, a qual aparece mais plana que o aceitável (-3.0).



Exemplo C: DNA inibido. Esta é uma outra ocorrência de baixa qualidade associada à solução de DNA. A inclinação da diluição serial está dentro da variação aceitável, no entanto, o Ct calculado da amostra não diluída (22,48), baseado nos quatro pontos deveria ser menor que o Ct medido (23). Isto indica um atraso no início da reação para a amostra não diluída o qual é menos evidente na amostra subsequente 1:4. Portanto, enquanto a inclinação da regressão linear se adequa à faixa aceitável de -3.6 to -3.1, o  $\Delta$ Ct demonstrou a presença de compostos que inibem a amplificação do DNA.



Exemplo D: Neste exemplo, a inclinação da reta de regressão não atende aos critérios de aceitabilidade (-3.66). Entretanto, ao contrário do exemplo B, não há atraso no início da reação para a amostra não diluída e a primeira amostra da diluição serial. A coluna de ' $\Delta$ Ct' mostra a diferença no Ct medido entre as amostras subsequentes da diluição serial. Estes valores são sempre maiores que o valor esperado de 2 para uma reação com 100% de eficiência. **Basicamente, a série de diluição comporta-se como se tivesse menos DNA do que a quantidade calculada.** Portanto, neste caso as amostras não devem ser classificadas como afetadas pela presença de compostos inibidores de PCR. Deve-se considerar, no entanto, que se não forem identificados erros técnicos (como erro de pipetagem) ou outras razões não forem claramente identificadas, a possibilidade de inibidores ligados aos alvos de DNA não deve ser descartada *a priori*.

## 2. Orientações para produção de concentrações intermediárias de material positivo

### 3.1. Produção de concentrações intermediárias de material positivo

Alguns materiais de referência somente estão disponíveis em uma ou poucas concentrações da sequência específica. Nestes casos pode ser necessário produzir material alvo em outras concentrações pela mistura de material positivo com material negativo, por exemplo, para determinar o LOQ e LOD relativos. Isto pode ser feito medindo-se o conteúdo do gene de referência nas soluções de DNA dos materiais positivo e negativo na mesma placa e com a mesma curva padrão. Seguindo estes passos, o fator de diluição para as duas soluções de DNA pode ser calculado usando a seguinte fórmula:

$$X = (A/B) * (Y-1) + 1$$

X = Fator prático de diluição (Quanto do material alvo deve ser diluído para compensar a diferença na concentração).

A = Número de cópias do gene de referência para a solução de DNA alvo positiva

B = Número de cópias do gene de referência para a solução de DNA alvo negativa

Y = Fator de diluição teórico, por exemplo, de 10 % de DNA alvo até 0.1 % de DNA alvo = 100x

Exemplo:

DNA A = 100% DNA alvo,

DNA B = 0%

5 µL DNA é adicionado por poço na PCR para A e B

Resultado da quantificação como amostra desconhecida (unknown sample) na curva de calibração do gene de referência:

A (do DNA A): 10 cópias/5 µL

B (do DNA B): 8 cópias/5 µL

Para preparar 10 % de DNA alvo de um material 100 % de DNA alvo, o que corresponde a uma diluição de 10 vezes (teoricamente Y = 10).

X deve ser usado como Y no cálculo do volume a ser misturado. Se X=12,25, então o fator de diluição prático X:  $(10/8) * (10-1) + 1 = ((10/8) * 9) + 1 = 90/8 + 1 = 11,25 + 1 = 12,25$ .

Então, 1 µl A deve ser misturado a 11,25 µl B.

Depois de misturadas as duas soluções de DNA, agitar cuidadosamente até a completa homogeneização. A exatidão das misturas pode ser analisada usando uma mistura 100% para a curva padrão e analisando três amostras em triplicata em três dias diferentes.

Para preparar futuras diluições: uma solução 1% pode ser feita diluindo a solução 10% na proporção 1:10 da solução 10% com uma solução 0%. Uma solução 0,1% pode ser preparada pela diluição da solução 1% com uma solução 0% na proporção de 1:10. A concentração de DNA de uma solução diluída 10 vezes deve estar próxima da concentração da solução de DNA 0% e esta concentração deve ser utilizada nos cálculos posteriores em diluições futuras. A exatidão da mistura pode ser analisada usando 100% da mistura da curva padrão e analisando três amostras em triplicadas 3 vezes.

### **3. Estimativa da média, desvio padrão e desvio padrão relativo de repetibilidade do conteúdo de DNA alvo por PCR em tempo real**

#### **4.1. Introdução**

O cálculo correto da concentração de DNA alvo e seu desvio padrão de análises de PCR são feitos, na maioria dos casos, por um processo em duas etapas, combinando valores médios calculados de duas formas diferentes. O procedimento detalhado abaixo inicia pelos valores medidos do número de cópias do DNA alvo e do gene endógeno. Destes resultados, calcula-se uma estimativa do conteúdo de DNA alvo e desvio padrão. A maioria dos procedimentos experimentais proporciona vários desses valores (por exemplo, de várias corridas na mesma ou em diferentes placas) de concentração de DNA alvo e desvio padrão. Estas estimativas podem, se necessário e apropriado, ser combinadas de forma padrão, por exemplo, tomando a média aritmética no caso de concentração de DNA alvo.

#### **4.2. Estimativa da concentração de DNA alvo a partir do número de cópias dos genes alvo e endógeno**

Duas análises são necessárias para estimar a percentagem de DNA alvo de amostras usando PCR em tempo real:

Uma análise é usada para detectar o número de cópias da sequência DNA alvo (X) e outra é usada para determinar o número de cópias da sequência de DNA do gene endógeno (Y). A estimativa da percentagem de concentração de DNA alvo é obtida usando a seguinte razão de:

$$\%GM = \frac{\text{target DNA copy number}}{\text{reference DNA copy number}} 100 = \frac{X}{Y} 100. \quad (1)$$

Onde lê-se target, leia-se alvo, copy: cópia, number: número e reference: referência.

Tanto o  $X$  quanto o  $Y$  são variáveis aleatórias. É uma prática padronizada analisar o DNA alvo e o gene endógeno em duplicatas, triplicatas, etc. Isto resulta em 2, 3, ou mais resultados para cada sequência de DNA alvo e endógena, possibilitando o cálculo da média da percentagem de DNA alvo destes dois conjuntos de resultados. Infelizmente, não há uma fórmula exata para cálculo da média e variância da razão entre variáveis aleatórias. Mas existem aproximações. A média, denominada  $E$ , e a variância da razão,  $Var$ , entre duas variáveis aleatórias e independentes são aproximadamente:

$$E\left[\frac{X}{Y}\right] \approx \frac{\bar{x}}{\bar{y}} + \frac{\bar{x}}{\bar{y}^3} Var(y) \quad (2)$$

and

$$Var\left[\frac{X}{Y}\right] \approx \left(\frac{\bar{x}}{\bar{y}}\right)^2 \left(\frac{Var(x)}{\bar{x}^2} + \frac{Var(y)}{\bar{y}^2}\right), \quad (3)$$

Onde  $x$  é a média aritmética do número de cópias do DNA alvo e  $y$  é a média aritmética do número de cópias do gene endógeno.

Estas aproximações assumem que não há correlação entre  $X$  e  $Y$ . O desvio padrão é dado por  $sd[X/Y] = \sqrt{Var[X/Y]}$ . O desvio padrão relativo de repetibilidade  $RSDr$  é calculado no final do procedimento dos componentes de desvio padrão. Os detalhes de como calcular o  $RSDr$  são mostrados nos exemplos a seguir. Todos os cálculos podem ser implementados no Excel.

### 4.3. Exemplos

Nos exemplos,  $x$  representa o número de cópias do DNA alvo e  $y$  representa o número de cópias do gene endógeno. Estes exemplos correspondem aos exemplos dados na Tabela 1 (anexo 1) e demonstram em detalhes os cálculos necessários para uma placa e descrevem também como os resultados de várias placas podem ser combinados.

### 4.3.1. Exemplo 1

**Duas extrações de DNA, para cada extração tanto o DNA alvo quanto o gene endógeno são testados em duas repetições de PCR em quatro placas**

Este delineamento proporciona duas estimativas da concentração de DNA alvo e desvios padrões para cada placa e então oito estimativas da concentração de DNA alvo (por exemplo GM1-8) e oito estimativas de desvios padrões ( $sd1-8$ ) no total. Cada uma destas oito estimativas de concentração de DNA alvo e desvios padrões são derivados das equações (1) e (2) de dois resultados para o número de cópias do DNA alvo e do gene endógeno. Se for tomada a média de todas as oito estimativas, o valor médio depende de 16 resultados, o que implica também no desvio padrão combinado.

OGM alvo		Gene endógeno	
Ct	Número de cópias	Ct	Número de cópias
24,41	16119	21,30	156758
24,61	13954	21,18	171196

Então,  $x = 15036,97$ ,  $y = 163977$ ,  $Var(x) = 2343612,5$  and  $Var(y) = 104227922$ . Colocando o valor apropriado na equação (2) fornece uma média de GM1 de 0,092 ou 9,2%; usando a equação (3) com os valores acima e tomando a raiz quadrado fornece um desvio padrão de  $sd1$  de 0,010943.

OGM alvo		Gene endógeno	
Ct	Número de cópias	Ct	Número de cópias
25,50	13405	21,10	172089
25,44	14000	21,19	160907

Aqui,  $x = 13702,5$ ,  $y = 166498$ ,  $Var(x) = 177012,5$  e  $Var(y) = 62518562$ . Com os valores apropriados, a equação (2) fornece uma média de GM2 de 0,082 ou 8,2%; usando a equação (3) e tomando a raiz quadrada fornece um desvio padrão  $sd2$  de 0,004654.

#### 4.3.1.1. Combinando as placas

Todo o procedimento é repetido em quatro diferentes placas, dando em adição às concentrações de GM1, GM2,  $sd1$  e  $sd2$  as médias GM3, GM4, GM8 e os desvios padrões  $sd3$ ,  $sd4$ , ...,  $sd8$ .

A média total da amostra  $GM$  pode ser calculada tomando-se a média aritmética de GM1 -

GM8, i.e.  $\overline{GM} = \sum_{i=1}^8 GM_i / 8$ . Usando  $n$  como número de repetições por extração (neste

exemplo  $n=2$ ) e  $k$  como número de desvios padrões a serem agrupados ( $k=8$ ), o desvio padrão

da média total  $sd_{GM} = \sqrt{\sum_i^8 (n_i - 1)sd_i^2 / (\sum_i^8 n_i - k)}$ ; o termo no denominador é 16-8=8 neste

exemplo. O desvio padrão relativo de repetibilidade  $RSD_r = \frac{sd_{GM}}{GM} 100$ .

#### 4.3.2. Exemplo 2

**Duas extrações de DNA, o DNA alvo e o gene endógeno são testados em quatro repetições de PCR em duas placas**

Este delineamento proporciona duas estimativas da concentração de DNA alvo e desvio padrão para cada placa e quatro estimativas da concentração de DNA alvo (por exemplo GM1-4) e quatro estimativas de desvios padrões ( $sd1-4$ ) no total. Cada uma destas estimativas é derivada usando as equações (1) e (2) para quatro resultados do número de cópia do DNA alvo e o gene endógeno. Se a média de todas as quatro estimativas for tomada, o valor médio depende – como no exemplo 1 acima – de 16 resultados do número de cópias do DNA alvo e do gene endógeno, o que implica no desvio padrão combinado.

Extração 1

DNA alvo		Gene endógeno	
Ct	Número de cópias	Ct	Número de cópias
24,41	16119	21,30	156758
24,61	13954	21,18	171196
25,50	13405	21,10	172089
25,44	14000	21,19	160907

Neste exemplo,  $x = 14369,5$ ,  $y = 165237,5$ ,  $Var(x) = 1433394$  e  $Var(y) = 57700642$ . Aplicando a equação (2) obtemos uma média de GM1 de 0,087 or 8,7%; usando a equação (3) e calculando a raiz quadrada fornece um desvio padrão  $sd1$  de 0,00828.

Extração 2

DNA alvo		Gene endógeno	
Ct	Número de cópias	Ct	Número de cópias
26,21	14826,97	21,09	165248
26,30	13885,92	21,09	165248
26,38	13099,69	21,20	152168
26,20	14935,39	21,25	146569

Neste exemplo,  $x = 14187$ ,  $y = 157308$ ,  $Var(x) = 747515$  e  $Var(y) = 89272181$ . Aplicando a equação (2) fornece uma média de GM2 de 0,091 or 9,1%; usando a equação (3) e calculando a raiz quadrada fornece um desvio padrão  $sd1$  de 0,0077.

#### 4.3.2.1. Combinando placas

O procedimento inteiro é repetido em duas diferentes placas, dando em adição a GM1, GM2,  $sd1$  e  $sd2$  as médias GM3 e GM4 e os desvios padrões  $sd3$  e  $sd4$ . A média total da amostra GM pode então ser calculada tomando-se a média aritmética de GM1 - GM8, i.e.,

$\overline{GM} = \sum_1^4 GM_i / 4$ . Usando  $n$  como número de repetições por extração ( $n=4$  neste exemplo) e

$k$  como número de desvios padrões separados a serem agrupados ( $k=4$  neste exemplo), o

desvio padrão da média total  $sd_{GM} = \sqrt{\sum_1^4 (n_i - 1)sd_i^2 / (\sum_1^4 n_i - k)}$ ; o termo no denominador é

$16-4=12$  neste exemplo. O desvio padrão relativo de repetibilidade  $RSDr = \frac{sd_{GM}}{GM} 100$ .