

UNIVERSIDADE SÃO FRANCISCO  
*Campus de Bragança Paulista*

**JOÃO BATISTA HENRIQUE**

**IDENTIFICAÇÃO DE GENES ENDÓGENOS ADEQUADOS PARA  
ANÁLISES DE PCR EM TEMPO REAL EM *Klebsiella pneumoniae***

**Bragança Paulista - SP**

**2011**

UNIVERSIDADE SÃO FRANCISCO  
*Campus de Bragança Paulista*

**JOÃO BATISTA HENRIQUE**

**IDENTIFICAÇÃO DE GENES ENDÓGENOS ADEQUADOS PARA  
ANÁLISES DE PCR EM TEMPO REAL EM *Klebsiella pneumoniae***

**ORIENTADOR**

**Prof. Dr. Lúcio Fábio Caldas Ferraz**

Dissertação apresentada ao Curso de Pós-Graduação *Stricto Sensu* em Ciências da Saúde da Universidade São Francisco (USF) para a obtenção do título de Mestre em Ciências da Saúde.

**Bragança Paulista - SP**

**2011**

QW 138.5 K5 H448i Henrique, João Batista.  
Identificação de genes endógenos adequados para análises de PCR em tempo real em *Klebsiella pneumoniae* / João Batista Henrique. -- Bragança Paulista, 2011.  
55 p.

Dissertação (mestrado) – Programa de Pós-Graduação *Stricto Sensu* em Ciências da Saúde da Universidade São Francisco.  
Orientação de: Lúcio Fábio Caldas Ferraz.

1. *Klebsiella pneumoniae*. 2. PCR em tempo real.  
3. Genes endógenos. I. Ferraz, Lúcio Fábio Caldas.  
II. Título.

Ficha catalográfica elaborada pelas bibliotecárias do Setor de  
Processamento Técnico da Universidade São Francisco.

## **BANCA EXAMINADORA DA DISSERTAÇÃO DE MESTRADO**

**Orientador:** Prof. Dr. Lúcio Fábio Caldas Ferraz

### **Membros Titulares:**

1. Prof. Dr. Lúcio Fábio Caldas Ferraz (orientador).
2. Profa. Dra. Fabiana Alexandrino.
3. Profa. Dra. Michelle Darrieux Sampaio Bertoncini.

### **Membros Suplentes:**

1. Prof. Dr. Marcelo Lima Ribeiro.
2. Profa. Dra. Natália Reiko Sato Miyasaka.

**Programa de Pós-Graduação *Stricto Sensu* em Ciências da Saúde da  
Universidade São Francisco.**

**Data:** 30/08/2011.

HENRIQUE, João Batista “IDENTIFICAÇÃO DE GENES ENDÓGENOS ADEQUADOS PARA ANÁLISES DE PCR EM TEMPO REAL EM *Kebsiella pneumoniae*” Dissertação defendida e aprovada no programa de Pós Graduação *Stricto Sensu* em Ciências da Saúde da Universidade São Francisco em trinta de Agosto de 2011 pela Banca examinadora constituída pelos professores:



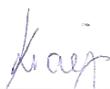
---

Prof. Dr. Lúcio Fábio Caldas Ferraz  
Universidade São Francisco



---

Prof<sup>a</sup>. Dra. Fabiana Alexandrino  
Universidade Estadual de Campinas



---

Prof<sup>a</sup>. Dra. Michelle Darrieux Sampaio Bertoncini  
Universidade São Francisco

## AGRADECIMENTOS

- Agradeço primeiramente a DEUS, que vem me acompanhando em todos os passos e momentos de minha vida, pela saúde, paz, e sabedoria.
- Ao meu orientador, Prof. Dr. Lúcio F. C. Ferraz, que com sua sabedoria, eficiência e muita paciência, esteve ao meu lado durante todo este processo de estudo e pesquisa e sem o qual este não se realizaria.
- Aos membros da banca, Profas. Dras. Michelle D. S. Bertoncini e Fabiana Alexandrino, que dispensaram seus afazeres para estar juntas a nós nesse momento tão esperado.
- Aos membros suplentes Prof. Dr. Marcelo Lima Ribeiro e Profa. Dra. Natália Reiko Sato Miyasaka, que ajudaram com suas opiniões e sugestões quanto a este trabalho.
- Aos colegas e técnicos do laboratório, em especial a Ana Érika, Rodrigo, Leonardo, Leandro, Júlio e Fabiano.
- À Universidade São Francisco (USF) e a todos os professores participantes do programa, diretores e às secretárias do programa de pós-graduação da USF, em especial a Rita e a Vilma.
- Minha esposa Ligia que sempre esteve pacientemente me incentivando nesta nova etapa em minha vida, a ela agradeço mui carinhosamente por cuidar de mim com muita dedicação amor e carinho.
- Minha filha Carolline Thabatta, que sempre torceu por mim e , no decorrer deste trabalho, me deu um netinho que é minha paixão, com um nome simples de um futuro grande homem, Murilo Henrique.

- Minha filha Francinne Haghatta, que tanto me incentivou e tenho a felicidade de ela estar na mesma área que a minha e esta se revelando minha sucessora.
- Meu filho Jean Carlos, sua esposa Andréia, minha neta Isabelle e meu neto Jean Miguel.
- Aos meus genros Felipe e Éderson, que sempre ficam debochando de mim, por ainda estar estudando, porém isso é uma forma de amor e carinho de se expressar e amor nunca é demais.
- Meus pais, Antonia Fernandes e Sebastião Henrique, que sempre torceram para que seus filhos conseguissem estudar e com orgulho, ofereço também a eles mais esta conquista em minha vida.
- Meus queridos irmãos, José, Maria, Eva, Paulo, André, Alexandre e Vanessa e meus cunhados e cunhadas.
- O sucesso, os fracassos e as vitórias de nossas vidas não dependem de terceiros e sim de nós.

## RESUMO

*Klebsiella pneumoniae* é uma bactéria gram-negativa que, sendo um patógeno oportunista, causa uma variedade de infecções em pacientes imunologicamente deprimidos, tais como pneumonia, meningite e septicemia. Um importante aspecto do estudo da patogenicidade de *Klebsiella pneumoniae* é a caracterização, por PCR em tempo real, do padrão de expressão dos genes potencialmente associados à virulência nesta bactéria. Uma etapa crítica na técnica de PCR em tempo real é a normalização das amostras investigadas por meio de um gene endógeno, ou controle interno. Um gene endógeno adequado deve, preferencialmente, apresentar um padrão de expressão constitutivo; isto é, com poucas variações, independente das condições de cultivo da bactéria. Entretanto, ao contrário de outras bactérias patogênicas, para *K. pneumoniae* não há relatos na literatura de trabalhos descrevendo a identificação de genes endógenos adequados. Neste sentido, o presente trabalho teve como objetivo a identificação de genes endógenos adequados para análises de PCR em tempo real em *Klebsiella pneumoniae*. Para tal, genes potencialmente candidatos a controles endógenos foram selecionados e tiveram o padrão de expressão aferido por PCR em tempo real a partir do cultivo de *Klebsiella pneumoniae* em diferentes condições. Os dados gerados foram estatisticamente analisados por meio dos programas *geNORM*, *NormFinder* e *BestKeeper* para se calcular a estabilidade da expressão dos genes candidatos. Baseados nestas análises, este trabalho sugere os genes *recA* e *rho* como controles endógenos mais adequados para análises de expressão gênica por PCR em tempo real em *Klebsiella pneumoniae*.

**Palavras-chave:** *Klebsiella pneumoniae*; PCR em tempo real; Genes endógenos.

## ABSTRACT

*Klebsiella pneumoniae* is a gram-negative opportunistic pathogen that may causes a variety of infections on immunocompromised individuals, such as pneumonia, meningitis, and septicemia. An important aspect on the studies of the *Klebsiella pneumoniae* pathogenicity is the expression analysis of virulence-related genes by real time PCR. One of the most crucial steps on real time PCR analysis is the normalization of the data by one or more endogenous control genes. An ideal endogenous control gene should have constitutive and stable expression, with minor variations during various experimental conditions. However, unlike other pathogenic bacteria, for *Klebsiella pneumoniae* there is no report on the literature describing the identification of suitable endogenous control genes. Therefore, the aims of this work were to identify and validate potential endogenous control genes suitable for real time PCR analysis on *K. pneumoniae*. For this purpose, genes potentially candidates to be endogenous control were selected and their expression pattern were analyzed by real time PCR on *Klebsiella pneumoniae* submitted to different experimental conditions. The expression data were statistically analyzed by *geNorm*, *NormFinder* and *BestKeeper* to calculate the expression stability of the candidate genes. Based on these analyzes, this work suggests *recA* and *rho* as endogenous control suitable for gene expression analyzes by real time PCR on *Klebsiella pneumoniae*.

**Keywords:** *Klebsiella pneumoniae*; Real time PCR; Endogenous control genes.

## LISTA DE FIGURAS

- Figura 1.** Curva de crescimento de *Klebsiella pneumoniae* em meio de cultura LB..... Página 37.
- Figura 2.** Eletroforese em gel de agarose mostrando a integridade dos RNAs totais extraídos de *Klebsiella pneumoniae*..... Página 38.
- Figura 3.** Eletroforese em gel de agarose mostrando os amplicons gerados por PCR a partir dos pares de *primers*..... Página 39.
- Figura 4.** Gráfico com o número  $C_T$  de cada gene obtido por PCR em tempo real nas diferentes condições de cultivo testadas..... Página 41.
- Figura 5.** Curva de *melting* indicando a especificidade de amplificação dos *primers* utilizados na PCR em tempo real..... Página 44.

## LISTA DE TABELAS

<b>Tabela 1.</b> Seqüência nucleotídica dos pares de <i>primers</i> utilizados nas reações de PCR.....	Página 30.
<b>Tabela 2.</b> Genes candidatos a controle endógeno selecionados para análise.....	Página 36.
<b>Tabela 3.</b> Cálculo da eficiência de amplificação dos pares de <i>primers</i> .....	Página 40.
<b>Tabela 4.</b> Variação dos números $C_{Ts}$ de cada gene nas oito condições de cultivo testadas.....	Página 42.
<b>Tabela 5.</b> Dados estatísticos gerados pelo programa <i>Bestkeeper</i> para os 11 genes candidatos .....	Página 45.
<b>Tabela 6.</b> Cálculos estatísticos gerados pelos programas <i>BestKeeper</i> , <i>NormFinder</i> e <i>geNorm</i> .....	Página 46.

## LISTA DE SIGLAS E ABREVIATURAS

<b>DEPC</b>	Dietilpirocarbonato
<b>dNTPs</b>	Desoxinucleosídeos trifosfatados
<b>DTT</b>	Ditiotreitol
<b>EDTA</b>	Ácido etilenodiaminotetraacético
<b>LB</b>	Luria-Bertani
<b>MOPS</b>	Ácido 3-[ <i>N</i> -morfolino] propanossulfônico
<b>r.p.m.</b>	Rotações por minuto
<b>TBE</b>	Tris-Borato-EDTA

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	14.
1.1 <i>Klebsiella pneumoniae</i> .....	14.
1.2 Captação e homeostase de ferro.....	15.
1.3 Fatores de virulência influenciados por Ferro.....	17.
1.4 Reação PCR e o estudo de expressão gênica.....	19.
1.5 Identificação de genes endógenos adequados.....	21.
<b>2. OBJETIVOS</b> .....	27.
2.1 Objetivo geral.....	27.
2.2. Objetivos específicos.....	27.
<b>3. MATERIAIS E MÉTODOS</b> .....	28.
3.1 Seleção de potenciais genes endógenos em <i>Klebsiella pneumoniae</i> .....	28.
3.2 Condições de cultivo de <i>Klebsiella pneumoniae</i> .....	28.
3.3 Extração de DNA genômico e teste dos <i>primers</i> por PCR convencional.....	29.
3.4 Extração de RNA total e Síntese de cDNA.....	30.
3.5 Análises por PCR em Tempo Real.....	31.
3.6 Análises Estatísticas.....	33.
<b>4. RESULTADOS E DISCUSSÕES</b> .....	35.
4.1 Seleção de candidatos a genes endógenos.....	35.
4.2 Condições de cultivo e extração de RNA total para síntese de cDNA.....	37.
4.3 Teste de eficiência dos pares de <i>primers</i> por PCR convencional e por PCR em tempo real.....	38.
4.4 Análises por PCR em tempo real.....	40.
4.5 Análises Estatísticas.....	43.
<b>5. CONCLUSÃO</b> .....	49.
<b>6. REFERÊNCIAS</b> .....	51.

## 1. INTRODUÇÃO

### 1.1 *Klebsiella pneumoniae*

*Klebsiella pneumoniae* é uma bactéria gram-negativa que pertence à família Enterobacteriaceae. Esta família é formada por um grupo grande e heterogêneo de bactérias gram-negativas, sendo composta principalmente dos gêneros *Klebsiella*, *Enterobacter*, *Serratia* e *Hafnia*). Essas bactérias representam importante causa de septicemia, abscessos abdominais, pneumonias nosocomiais e infecções dos tratos genito-urinário e respiratório (PATERSON, 2006).

O gênero *Klebsiella* apresenta características bioquímicas que permitem a sua identificação. Possui reação de oxidase negativa, fermenta glicose, reduz nitrato, é lisina positiva, citrato e indol negativos, tríplice açúcar ferro (TSI) produtora de gás, ornitina negativa, metaboliza a lactose, utiliza o citrato como fonte de carbono e também hidrolisa a uréia, formando gás ou não. A maioria das amostras é capaz de produzir o butilenoglicol como produto final da fermentação da glicose (KONEMAN et al., 2001).

Os microorganismos do gênero *Klebsiella* podem ser encontrados em quase todos os ambientes naturais (solo, água e plantas), sendo amplamente encontrados na natureza. No entanto, *Klebsiella pneumoniae* é uma das principais causadoras de infecções hospitalares, chegando a ser responsável por até 29% das infecções que, muitas vezes, estão associadas a uma alta morbimortalidade (PEREIRA et al., 2003). *K. pneumoniae* é um bastonete com tamanho que varia de 0,3 a 1  $\mu$  de diâmetro e 0,6 a 6  $\mu$  de comprimento, sem motilidade, encapsulada, não esporulado, que apresenta melhor crescimento em condições aeróbias, e que produz colônias grandes e gomosas quando cultivadas em placas com nutrientes.

## 1.2 Captação e homeostase de ferro

O ferro é um micronutriente essencial para a maioria dos microorganismos, tendo importante papel no metabolismo celular. Bactérias possuem muitos mecanismos relacionados com a importação de ferro. Uma das estratégias utilizadas pelas bactérias para a obtenção de ferro é ativar os sistemas de captação de ferro, como a produção de sideróforos e receptores específicos.

As bactérias produzem e secretam para o meio externo uma variedade de moléculas de baixo peso molecular, conhecidas como sideróforos, que possuem alta afinidade a íons férricos (SCHALK et al., 2004). O complexo sideróforo+íon férrico é captado pelas bactérias por meio de receptores específicos localizados na membrana externa. O transporte destes complexos para dentro das células depende da energia próton motora fornecida pelo sistema TonB, que é formado pelas proteínas TonB, ExbB e ExbD ancoradas na membrana citoplasmática. No citoplasma, o íon férrico é reduzido a íon ferroso, e como este apresenta baixa afinidade por sideróforo, o complexo é dissociado e o íon ferroso é então utilizado no metabolismo celular (CLARKE et al., 2001; ANDREWS et al., 2003).

Algumas espécies de bactérias obtêm o ferro diretamente do hospedeiro para seu crescimento, sem produzir sideróforos. Espécies do gênero *Neisseria* possuem receptores específicos que removem o ferro de transferrinas na superfície celular. *Yersinia pestis* utilizam heme como sua única fonte de ferro para seu crescimento, assim como *Haemophilus influenzae* que obtém ferro diretamente de hemoglobina e de heme complexada a albumina (LITWIN et al., 1993). Outro mecanismo de aquisição de ferro que não envolve sideróforos é aquele mediado por hemolisinas. Hemolisinas são exotoxinas que provocam a lise de eritrócitos e promovem a liberação de ferro intracelularmente complexado ao heme em hemoglobinas. Em muitas bactérias patogênicas a expressão de hemolisinas é frequentemente regulada de acordo com a disponibilidade de ferro no meio, de modo que sua produção é aumentada quando a quantidade de ferro é limitante (LITWIN et al., 1993).

Apesar de sua importância no metabolismo celular, o ferro pode ser extremamente tóxico quando em grande quantidade devido à sua capacidade de

gerar espécies reativas de oxigênio. Por este motivo, as bactérias têm desenvolvido um sistema homeostático que mantém os níveis intracelulares de ferro em um limite suficiente para suas necessidades essenciais, porém evitando danos oxidativos às células.

Deste modo, uma das formas de controlar a homeostase intracelular de ferro é a regulação dos mecanismos de importação deste nutriente através da membrana. Em muitas bactérias, o sistema de homeostase de ferro é regulado pelo repressor transcricional Fur (do inglês, *ferric uptake regulator*). A proteína Fur exerce uma regulação transcricional sobre quase todos os genes implicados no transporte e incorporação deste metal em função da concentração intracelular existente e também sobre outros sistemas não vinculados diretamente ao ferro (ESCOLAR et al., 1999).

O tamanho da proteína Fur oscila nos diferentes microorganismos, variando de 128 a 160 aminoácidos (ESCOLAR et al., 1999). Nesta proteína é possível identificar dois domínios perfeitamente diferenciados. A extremidade C-terminal é responsável pela ligação com íons metálicos assim como pelas interações proteína-proteína que levam à dimerização. E a extremidade N-terminal é encarregada do reconhecimento e ancoragem aos sítios específicos na região promotora dos genes alvos (STOJILJKOVIC et al., 1995).

O regulador Fur atua, na maioria das vezes, como repressor da transcrição dos genes alvos e esta repressão requer o íon ferroso ( $\text{Fe}^{2+}$ ) como co-repressor. Em ambientes ricos em ferro, o regulador Fur forma um complexo com o íon ferroso, e este complexo é capaz de se ligar a seqüências nucleotídicas específicas, chamadas boxes Fur, localizada na região promotora dos genes alvos, levando à repressão da transcrição destes genes. Por outro lado, quando a concentração intracelular de ferro é relativamente baixa, o complexo  $\text{Fe}^{2+}$ -Fur não é formado e, portanto, não se ligará aos boxes Fur dos genes sob seu controle. Como conseqüência disso, haverá a expressão das unidades transcricionais encarregadas da importação de ferro, o que permitirá o restabelecimento de sua concentração intracelular (ESCOLAR et al., 1999).

A seqüência nucleotídica dos boxes Fur é um palíndromo degenerado rico em adeninas e timinas de aproximadamente 19 pares de bases (ESCOLAR et al.,

1999). Análises bioinformáticas de genomas completos, associadas a estudos de transcriptomas e de proteomas em condições de carência de ferro possibilitaram a identificação e comparação de seqüências de boxes Fur existentes nos promotores de genes de diversas bactérias (PANINA et al., 2001).

Apesar de inicialmente descrito como um regulador específico de genes envolvidos na captação de ferro, Fur é considerado um regulador global que controla a expressão de genes envolvidos nos mais diversos processos celulares, tais como no metabolismo energético (QUATRINI *et al.*, 2005), na tolerância a acidez (HALL e FOSTER, 1996) e na resposta ao estresse oxidativo (COOKSLEY *et al.*, 2003). Assim como na maioria das bactérias patogênicas, a capacidade de *Klebsiella pneumoniae* de captar ferro constitui-se em um importante fator determinante de patogenicidade nesta bactéria. E o regulador Fur controla a expressão não apenas dos genes da homeostase de ferro, mas também de genes que determinam a patogenicidade nesta bactéria, tais como genes do mecanismo de captação de ferro, formação de biofilme, resposta ao estresse oxidativo e resistência a drogas (ACHENBACH e YANG, 1997).

### **1.3 Fatores de virulência influenciados por Ferro**

Em bactérias patogênicas, um dos sinais que desencadeia a infecção e colonização é a privação de ferro que o patógeno enfrenta quando em contato com o hospedeiro. De fato, o ferro desempenha um papel crucial na regulação da expressão de inúmeros fatores de virulência, como formação de biofilme mediado por *quorum-sensing*, expressão de genes de virulência, toxinas, proteases, e sistemas de assimilação de ferro (LITWIN et al., 1993). Muitos dos efeitos do ferro na regulação dos fatores de virulência são mediados pelo regulador Fur.

Um importante fator de virulência bacteriano regulado pelos níveis de ferro é a formação de biofilme (SINGH et al., 2004). Nestas comunidades organizadas, as células encontram-se encapsuladas em uma matriz de exopolissacarídeos, podendo se desenvolver em superfícies bióticas e abióticas (O'TOOLE et al., 2000). Em função da estrutura encapsulada, o biofilme tem sido associado a um aumento na resistência a agentes antimicrobianos em bactérias patogênicas

(MAH E O'TOOLE, 2001). Alguns trabalhos têm demonstrado que o ferro desempenha um papel crítico no processo de formação de biofilme (O'TOOLE E KOLTER, 1998; SINGH et al., 2004; BANIN et al., 2005). O'TOOLE E KOLTER (1998) demonstraram que o ferro restabelece a capacidade de aderência em bactérias que são deficientes nesta capacidade, e SINGH et al. (2004) mostraram que níveis baixos de ferro estimulam a motilidade das bactérias e bloqueiam a formação de microcolônias.

Além da influência do ferro na formação do biofilme, este processo parece estar associado ao mecanismo de sinalização denominado *quorum-sensing*. O *quorum-sensing* é um mecanismo com o qual as bactérias se comunicam entre si por meio da secreção e detecção de pequenas moléculas sinalizadoras denominadas auto-indutoras. Em resposta às moléculas sinalizadoras do *quorum-sensing* as bactérias migram e aderem a superfícies, expandindo-se em múltiplas camadas celulares que formarão o biofilme (COLE et al., 2004). Além de associado à formação de biofilme, o mecanismo de *quorum-sensing* regula outros processos celulares relacionados à patogenicidade, incluindo motilidade, expressão de genes de virulência, regulação dos sistemas de secreção, e de assimilação de ferro (FERDELE e BASSLER, 2002).

O ferro também desempenha um importante papel na regulação da expressão de inúmeros fatores de virulência, tais como toxinas. Condições de privação de ferro levam a um aumento na produção da toxina Shiga por *Shigella dysenteriae*, o mesmo ocorrendo com a produção de toxinas SLTs (do inglês, *Shiga-like toxins*) por algumas linhagens de *E. coli* enterohemorrágica (LITWIN et al., 1993). Além da toxina Shiga, a toxina diftérica, produzida por *Corynebacterium diphtheriae*, e exotoxina A, produzida por *Pseudomonas aeruginosa*, também são reguladas de acordo com a concentração de ferro no meio externo (LITWIN et al., 1993).

A importância do ferro também tem sido descrita em infecções causadas por outras bactérias patogênicas, tais como em *Neisseria gonorrhoeae* e *Neisseria meningitidis*, *Vibrio cholerae* e *Vibrio anguillarum*, espécies do gênero *Yersinia*, e etc (LITWIN et al., 1993).

## 1.4 Reação PCR e o estudo de expressão gênica

O padrão de expressão gênica reflete, em parte, os mecanismos envolvidos na adaptação às condições ambientais. Todas as respostas adaptativas implicam em vários interruptores genéticos que controlam as mudanças metabólicas que ocorrem. Assim, a capacidade de quantificar os níveis de transcrição de genes específicos pode ser útil para caracterizar a resposta do estado e estresse fisiológico de organismos (DESROCHE et al., 2005).

Os métodos usados para estudar a expressão gênica, como hibridização de *Northern Blot*, a reação PCR (reação em cadeia da polimerase, tradução livre do inglês *Polymerase Chain Reaction*) quantitativo, e ensaios de proteção RNase, podem apresentar falta de sensibilidade (VANDECASTEELE et al., 2001; DESROCHE et al., 2005). Dentre as técnicas, a PCR em tempo real é uma técnica extremamente sensível que permite a medição precisa da expressão gênica.

A PCR em tempo real é muitas vezes considerada a técnica ideal para quantificar a expressão do gene, e é comumente empregada para validar as técnicas com maior rendimento, mas menor sensibilidade em geral, como a análise por *microarray* (WALKER et al., 2009).

A técnica de PCR em tempo real depende da utilização de corantes fluorescentes para quantificar a transcrição do gene alvo, com o número do ciclo de amplificação em que estes corantes/transcrições são detectados, apresentando uma indicação quanto à abundância relativa das moléculas-alvo. A sensibilidade da técnica de PCR em tempo real constitui uma poderosa ferramenta para medição da expressão gênica, sobretudo quando as quantidades de amostra são limitadas ou a transcrição é expressa em um nível baixo. No entanto, essa sensibilidade também significa que uma grande dose de cuidado é necessária com relação ao design experimental e aplicação do procedimento (WALKER et al., 2009).

Em procaríotos a PCR também é uma técnica importante, que permite a quantificação dos níveis de RNA mensageiro (RNAm) em bactérias. Duas

quantificações poderiam ser usadas: quantificação relativa, ou quantificação absoluta (DESROCHE et al., 2005).

A quantificação relativa é o método analítico mais indicado para muitos estudos de PCR em tempo real e é mais bem aplicado quando existem muitos genes a serem testados em condições múltiplas. Este método requer o uso de controle interno para a normalização dos dados. De fato, as variações, como a quantidade de reagentes, a eficiência enzimática, e as diferenças na atividade transcricional entre diferentes culturas podem interferir na quantificação final dos níveis de transcritos (DESROCHE et al., 2005).

Os genes endógenos atuam como controles internos em reações de PCR em tempo real, tendo a finalidade de remover ou reduzir as diferenças da amostragem que muitas vezes estão relacionadas à quantidade e qualidade do RNA extraído (OLSVIK et al., 2005). A expressão desses genes é considerada pouco flutuante em comparação com outros genes. No entanto, em determinadas condições, sua expressão pode variar consideravelmente.

Hays (2009) afirma que, por razões práticas, o uso de dois genes endógenos deve ser suficiente para produzir dados de expressão gênica relativamente precisos, sem que haja aumento significativo de custos e tempo de laboratório em comparação com o uso de um único gene endógeno.

O gene endógeno ideal deve ser expresso em um nível constante nos diferentes tecidos de um organismo, em todas as fases do desenvolvimento e não deve ser afetado pelo tratamento experimental. Também deve ser expresso em aproximadamente o mesmo nível que o RNAm dos genes em estudo (OLSVIK et al., 2005).

No entanto, a normalização de dados em PCR em tempo real continua sendo um problema, sobretudo para a quantificação relativa. Numerosos estudos têm revelado que não há um único gene universal que possua um nível de expressão constante em todas as situações experimentais ou de desenvolvimento (DHEDA et al., 2004; RADONIC et al., 2004 OLSVIK et al., 2005).

Em eucariotos, a melhor escolha de um gene de referência para uso como controle endógeno varia, dependendo do tecido de interesse a ser estudado. Por essa razão, um grande número de genes tem sido selecionado para a

normalização de dados da expressão de RNAm. Se o gene de referência selecionado flutua aleatoriamente entre as amostras, pequenas diferenças na expressão entre os genes de interesse podem não ser percebidas. Neste sentido, o coeficiente de variação da expressão gênica entre diferentes experimentos deve idealmente ser tão baixo quanto possível. Em geral, a estabilidade dos genes de referência em potencial deve ser testada em todos os tecidos examinados ou células, e sob um delineamento experimental diferente. Um número crescente de trabalhos tem discutido a seleção de genes de referência na análise de PCR em tempo real (OLSVIK et al., 2005).

Em bactérias, não existe um controle interno que seja consenso. A estratégia mais freqüente implica a determinação de um controle interno específico para cada estudo para a normalização dos níveis de RNAm (DESROCHE et al., 2005).

### **1.5 Identificação de genes endógenos adequados**

Na literatura científica existem diversos estudos que têm apresentado a importância da caracterização e identificação de genes endógenos adequados para análises de PCR em tempo real (VANDESOMPELE et al., 2002; DESROCHE et al., 2005; OLSVIK et al., 2005; HEINRICH et al., 2007; TAKLE et al., 2007; HAYS, 2009; MALOUKH et al., 2009; NIETO et al., 2009; SILVEIRA et al., 2009; WALKER et al., 2009; KHEIRELSEID et al., 2010; MAROUFI et al., 2010).

Vandesompele et al. (2002) delinearão uma estratégia para identificar os genes de controle, em sua maioria de forma estável, expressos em um determinado conjunto de tecidos, e para determinar um número mínimo de genes necessários para calcular um fator de normalização confiável. Os autores concluíram que a estratégia de normalização apresentada é um pré-requisito para o perfil de expressão gênica de PCR em tempo real, que, entre outras coisas, abre a possibilidade de estudar a relevância biológica das pequenas diferenças de expressão.

Heinrich et al. (2007) investigaram se os genes endógenos comumente usados apresentam estabilidade em vários intervalos *post-mortem* em humanos. Os dados obtidos no estudo revelam que a degradação do RNAm *post-mortem* é um processo complexo, e que o uso de um único controle endógeno em estudos de expressão gênica do tecido *post-mortem* levaria a uma interpretação errônea dos dados. Deste modo, mais estudos sobre o assunto devem ser realizados no futuro, incluindo um aumento do número de amostras bem documentadas.

Kheirelseid et al. (2010) afirmam que a análise da expressão gênica tem muitas aplicações no diagnóstico de câncer, no prognóstico e nos cuidados terapêuticos. A quantificação relativa é a abordagem mais amplamente adotada, na qual a quantificação da expressão gênica é normalizada em relação a um gene endógeno. O objetivo do estudo foi avaliar um grupo de genes candidatos a gene endógeno para identificar o gene mais estável expresso para normalizar a PCR em tempo real, em dados derivados de tecido de câncer colo-retal primário. O referido estudo identificou que a combinação dos genes endógenos *B2M* (que codifica a beta-2 microglobulina) e *PPIA* (codifica a enzima peptidil prolil isomerase A) pode ser empregada de forma mais precisa para a normalização do PCR em tempo real de dados de tecidos de câncer colo-retal primário. Embora esses genes endógenos possam não ser ideais para o uso em estudos de outros tipos de câncer, a abordagem descrita pode servir como um modelo para a identificação de genes endógenos válida para outros tipos de câncer.

Walker et al. (2009) investigaram 15 genes endógenos candidatos a serem utilizados como controle interno em análises de expressão no endométrio bovino. O estudo estabeleceu o gene *SUZ12* (do inglês, *suppressor of zeste 12*) como o controle interno mais adequado para análises de comparação dos níveis de expressão nos tecidos do endométrio bovino, e demonstrou a utilidade da análise de *microarray* como um método para a identificação de genes endógenos candidatos.

Olsvik et al. (2005) afirmam que peixes salmonídeos estão entre as espécies de peixes mais estudados, mas havia uma carência de estudos que envolvessem a identificação de genes de referência para análises de PCR em tempo real. De modo geral, este trabalho sugere que os genes *EF1A<sub>A</sub>* (codifica o fator de alongação 1A<sub>A</sub>) e *EF1A<sub>B</sub>* (codifica o fator de alongação 1A<sub>B</sub>) podem ser

úteis como genes de referência em análise de PCR em tempo real da expressão gênica no salmão do Atlântico.

Maloukh et al. (2009) avaliaram a estabilidade da transcrição de seis genes endógenos candidatos a controle interno em *Humulus lupulus L* (lúpulo). A combinação dos genes *DRH1* (do inglês, *DEAD-box RNA helicase 1*), *GAPDH* (codifica a enzima gliceraldeído 3-fosfato-desidrogenase) e *7SL-RNA* (codifica o componente de RNA, subunidade 7SL, da partícula de reconhecimento de sinal) apresentaram resultados confiáveis na análise de PCR em tempo real.

Silveira et al. (2009) descreveram a identificação de genes endógenos adequados para estudos de expressão gênica em *Brachiaria brizantha*, uma gramínea forrageira de grande importância comercial e econômica para a pecuária bovina. Segundo os autores, a análise quantitativa por PCR em tempo real tem emergido como uma importante ferramenta para comparar o perfil de expressão de genes de interesse naquela gramínea. No entanto, para a obtenção de resultados confiáveis é necessário a utilização de genes de referência adequados. Os resultados obtidos no estudo de Silveira et al. (2009) forneceram informações importantes para a análise transcricional da *Brachiaria ssp*, ajudando a melhorar a qualidade dos dados de expressão gênica nessas espécies.

Maroufi et al. (2010) afirmam que em chicória a análise quantitativa por PCR em tempo real é uma técnica sensível para quantificar os níveis de expressão gênica. Um ou mais genes de referência devem ser selecionados para comparar com precisão as transcrições de RNAm em diferentes amostras e tecidos. Até o momento, apenas o gene *ACT*, que codifica actina, tem sido usado como controle interno para análises de PCR em tempo real em chicória. Os autores relatam uma comparação completa de vários genes de referência candidatos na chicória. Softwares de análise como *geNORM*, *NormFinder*, *BestKeeper*, foram empregados e indicaram que os genes *ACT*, *EF* (codifica o fator de alongação alfa-1) e *RNAr 18S* (codifica o RNA ribossômico subunidade 18S) apresentam estabilidade de expressão mais elevada em tecidos da folha e da raiz da chicória, enquanto os genes *GAPDH* (codifica a enzima gliceraldeído-3-fosfato-desidrogenase) e *NADHD* (codifica a enzima nicotinamida adenina dinucleotídeo desidrogenase) apresentaram estabilidade relativamente baixa. Os resultados do estudo enfatizam a importância de validar genes de referência para

PCR em tempo real em chicória. O uso de genes de referência mais estáveis, tais como *ACT* e *EF*, permite a normalização precisa da expressão gênica em tecidos de folhas e raiz de chicória.

Desroche et al. (2005) apresentaram um método para definir um controle interno para a quantificação relativa dos níveis de transcrição na bactéria *Oenococcus oeni* em condições de estresse e em condições miméticas de vinho no ácido láctico. Os autores concluíram que a PCR em tempo real constitui um método rápido e sensível que pode ser empregado para determinar o nível de transcrição relativa do *Oenococcus oeni* durante a fermentação no vinho sintético e poder ser usado para estudos em outras bactérias do ácido láctico com interesses industriais, tais como iniciadores de lácteos.

Takle et al. (2007) apresentaram um estudo onde a manipulação das condições de crescimento cultural em experimentos de plantas têm sido usados para validar a estabilidade da expressão gênica para genes endógenos, de um patógeno da planta denominado *Pectobacterium atrosepticum*, pertencente à bactéria da família *Enterobacteriaceae*. Os autores sugerem que os genes *recA* (codifica a enzima recombinase A) e *ffh* (codifica a proteína da partícula de reconhecimento de sinal) são candidatos adequados para a normalização precisa da PCR em tempo real de dados para experimentos que investiguem o patógeno *Pectobacterium atrosepticum*, e outros patógenos potencialmente relacionados.

Hays (2009) ressalta que a combinação de uma gama de genes endógenos putativos seja a melhor alternativa para produzir dados precisos sobre os níveis de expressão gênica. Analisando potenciais genes endógenos em *Moraxella catarrhalis*, o autor conclui que a combinação dos genes *RNAr 16S* (codifica o RNA ribossômico subunidade 16S) e *copB* (codifica proteína de captação de ferro) seria a escolha adequada para estudos de expressão gênica naquela bactéria. Contudo, o autor ressalta que diferentes combinações de genes podem ser necessárias dependendo da linhagem genética presente na amostra.

Nieto et al. (2009) afirmam que a normalização é um pré-requisito para a análise por PCR em tempo real precisa e para a validação de dados de perfis de *microarray* em sistemas microbianos. A escolha e o uso de genes de referência que são expressos de forma estável em amostras, condições experimentais e

modelos é uma questão fundamental para a interpretação precisa de dados de expressão gênica. O estudo fornece um conjunto de genes de referência validados para estudar a expressão gênica na bactéria *Acidithiobacillus ferrooxidans* sob condições típicas biológicas e um ponto de partida para explorar novas configurações experimentais nesses microorganismos e, eventualmente, em outros intimamente relacionados à *Acidithiobacillus ferrooxidans*. A informação também pode ser de valor para futuras experiências em outros sistemas de transcrição bacteriana.

Os estudos supracitados ressaltam a importância de se identificar e validar genes endógenos que sejam adequados para uso em PCR em tempo real. Na ausência tais estudos, muitos trabalhos têm empregado genes que codificam RNAs ribossômicos como controle endógeno. No entanto, há restrições quanto à utilização de um RNA ribossômico (RNAr) para a normalização do RNA mensageiro dos genes alvos.

Diversos trabalhos têm mostrado que o gene que codifica RNA ribossômico é inadequado para experimentos de PCR em tempo real, tanto em eucariotos (gene *RNAr 18S*) quanto em procariotos (gene *RNAr 16S*).

Em estudos que abordaram a identificação de genes candidatos a controle endógeno em peixes, tanto Frost e Nilsen (2003) quanto Olsvik et al. (2005) verificaram que o gene *RNAr 18S* apresentava expressão instável e com elevada variação, tornando-o inadequado como controle endógeno.

Dentre todos os genes endógenos candidatos analisados por Takle et al. (2007) no patógeno de plantas *Pectobacterium atrosepticum*, o gene *RNAr 16S* foi o que apresentou expressão mais instável nas diferentes condições testadas. Considerando que o gene *RNAr 16S* é altamente conservado entre diferentes espécies bacterianas, os autores ainda alertam para o surgimento de amplificações inespecíficas, e sugerem que este gene deve ser usado como referência apenas em culturas puras, e não em estudos que abordam a interação patógenos-hospedeiro.

Analisando possíveis genes de referência para a bactéria acidofílica *Acidithiobacillus ferrooxidans*, Nieto et al. (2009) chegaram à conclusão que, apesar de apresentarem expressão estável, o gene *RNAr 16S* deve ser

descartado como gene endógeno porque a elevada abundância de RNAs ribossômicos pode prejudicar a quantificação de transcritos raros e não tão abundantes.

Uma das principais desvantagens de se utilizar o gene *RNAr 16S* em normalizações no PCR em tempo real é a elevada quantidade do RNA ribossômico 16S dentro da célula em comparação com os RNAs mensageiros. De fato, o RNA ribossômico constitui em torno de 80 a 90% de todo o RNA presente na célula, uma concentração bem além do alcance da maioria dos métodos de calibração (THEIS et al., 2007). Segundo Vandesompele et al. (2002) esta proporção desbalanceada entre RNAs mensageiros e ribossômicos torna os genes que codificam para RNA ribossômicos inadequados como controles endógenos.

Outro fato que reforça que a impossibilidade de se comparar RNAs ribossômicos e mensageiros é que cada uma destas moléculas apresentam tempo distintos de duração dentro da célula. Enquanto os RNAs mensageiros apresentam rápido *turn-over* de acordo com as condições fisiológicas de uma bactéria, o RNA ribossômico somente é degradado em determinadas condições de estresse (DEUTSCHER, 2006).

Além disso, alguns estudos têm sugerido que a expressão dos genes que codificam RNAs ribossômicos pode estar sob algum controle regulatório (VANDECASTEELE et al., 2001; WAGNER, 2002; GRALLA, 2005). E estando sob controle, a expressão do gene deixa de ter o caráter constitutivo que é imprescindível para todo controle endógeno.

O gene endógeno atualmente utilizado em *K. pneumoniae* é o gene *rrsH*, que codifica o RNA ribossômico 16S. Ao contrário de outras bactérias patogênicas, para *Klebsiella pneumoniae* não há relatos na literatura de trabalhos descrevendo a identificação e validação de genes endógenos adequados para uso em PCR em tempo real.

## 2. OBJETIVOS

### 2.1 Objetivo geral

O presente trabalho tem como objetivo a identificação de potenciais genes endógenos adequados para análises de PCR em tempo real em *Klebsiella pneumoniae*.

### 2.2. Objetivos específicos

- Seleção de potenciais genes endógenos candidatos.
  - ✓ Busca tanto na literatura quanto no genoma anotado de *Klebsiella pneumoniae* MGH 78578, e
  - ✓ Desenho de oligonucleotídeos (pares de *primers*) específicos para cada um dos genes candidatos selecionados;
- Cultivo da bactéria *K. pneumoniae* nas condições ideais de cultura a serem testadas;
- Extração do DNA genômico de *K. pneumoniae* e teste dos pares de *primers* por meio de reação de PCR convencional;
- Extração de RNA total a partir de *K. pneumoniae* cultivada nas diferentes condições de cultura e síntese de DNA complementar (cDNA);
- Caracterização do padrão de expressão dos genes endógenos candidatos por PCR em tempo real a partir do cDNA sintetizado;
- Análises estatísticas dos dados obtidos no PCR em tempo real por meio de três programas independentes (*BestKeeper*, *geNorm* e *NormFinder*).

### 3. MATERIAIS E MÉTODOS

#### 3.1 Seleção de potenciais genes endógenos em *Klebsiella pneumoniae*

Para a seleção de potenciais candidatos a genes endógenos foi realizada uma busca tanto na literatura (banco de dados *PubMed*, disponível no *website* <http://www.ncbi.nlm.nih.gov/sites/entrez>), quanto no genoma anotado de *Klebsiella pneumoniae* MGH 78578 (disponível no banco de dados *Genome Project – NCBI*, disponível no *website* <http://www.ncbi.nlm.nih.gov/bioproject/31>).

A busca na literatura envolveu a pesquisa de estudos que abordassem a identificação e caracterização de genes endógenos adequados para análises de PCR em tempo real em bactérias gram-negativas.

Uma vez selecionados os candidatos a genes endógenos, os pares de oligonucleotídeos (*primers*) específicos para cada gene selecionado foram desenhados a partir da seqüência nucleotídica desses genes, obtida no genoma anotado de *Klebsiella pneumoniae* MGH 78578, e utilizando-se o programa *primer3* (<http://frodo.wi.mit.edu/primer3/>, ROZEN e SKALETSKY, 2000).

#### 3.2 Condições de cultivo de *Klebsiella pneumoniae*

Para este estudo foi utilizada a linhagem *Klebsiella pneumoniae* ATCC 10031. As bactérias foram crescidas em meio de cultura LB líquido, sob agitação a 250 r.p.m. e a 37°C. O padrão de crescimento da bactéria foi monitorado por meio de leitura em espectrofotômetro em densidade óptica (D.O.) de 600 nm.

O possível padrão constitutivo de expressão dos genes selecionados foi aferido a partir do cultivo de *K. pneumoniae* em diferentes condições. Com relação às fases de crescimento, foram isoladas células na fase inicial (fase *log*), fase intermediária (*mid-log*) e fase final de crescimento (estacionária). No caso dos suplementos, células foram isoladas após incubação na presença de uma fonte de ferro (FeSO<sub>4</sub>, sulfato ferroso), na presença de um quelante de ferro (2,2'-dipiridil) ou na presença de uma fonte de manganês (MnSO<sub>4</sub>, sulfato de

manganês). No caso das células submetidas ao choque térmico, estas foram isoladas após incubação a 45°C ou a 29°C. Desta forma, as oito condições testadas estão representadas como descrito abaixo:

- **Condição 1** - cultivo da bactéria até o início da fase *log* (equivalente a uma D.O.<sub>600nm</sub> de 0.3);
- **Condição 2** - cultivo da bactéria até a fase *mid-log* (D.O.<sub>600nm</sub> de 0.6);
- **Condição 3** - cultivo da bactéria até a fase estacionária (D.O.<sub>600nm</sub> de 1.5);
- **Condição 4** - após atingirem a fase *mid-log*, células foram incubadas por 30 minutos na presença de 100 µM de FeSO<sub>4</sub> (concentração final);
- **Condição 5** - após atingirem a fase *mid-log*, células incubadas por 30 minutos na presença de 100 µM de 2,2'-dipiridil (concentração final);
- **Condição 6** - após atingirem fase *mid-log*, células incubadas por 30 minutos na presença de 100 µM de MnSO<sub>4</sub> (concentração final);
- **Condição 7** - após atingirem a fase *mid-log*, células foram incubadas por 30 minutos a 29°C;
- **Condição 8** - após atingirem a fase *mid-log*, células foram incubadas por 30 minutos a 45°C;

### 3.3 Extração de DNA genômico e teste dos *primers* por PCR convencional

O DNA genômico de *Klebsiella pneumoniae* foi extraído por meio do kit Wizard<sup>®</sup> Genomic DNA purification Kit (Promega). Ao final do processo de extração, o DNA isolado foi checado em eletroforese em gel de agarose 0,8%, quantificado em espectrofotômetro e posteriormente utilizado em experimentos de PCR convencional.

Os pares de *primers* (Tabela 1) foram testados por uma reação de PCR convencional que continha os seguintes reagentes: 2,5 micro litros (µL) de tampão 10x, 0,2 mili molar (mM) de cada dNTPs (concentração final), 1,5 mM de MgCl<sub>2</sub> (concentração final), 0,2 micro molar (µM) de cada *primer* (concentração final), 1 µL de DNA genômico extraído de *K. pneumoniae*, 0,5 µL de Taq DNA polimerase (5U/µL, Invitrogen) e 12,75 µL de água autoclavada. As reações foram

submetidas aos seguintes ciclos: etapa inicial de desnaturação a 95°C por 5 minutos, 33 ciclos de 95°C por 1 minuto, 60°C por 45 segundos e 72°C por 1 minuto, e uma etapa final de extensão a 72°C por 10 minutos. O resultado obtido na reação de PCR foi checado após aplicação das amostras em gel de agarose 1,5% e corrida de eletroforese em tampão de corrida TBE 1x.

**TABELA 1.** Seqüência nucleotídica dos pares de *primers* usados na reação de PCR convencional e PCR em tempo real. Em “Seqüências”, **S** refere-se à seqüência sense dos *primers*, e **AS**, seqüência anti-sense.

<b>Gene</b>	<b>Seqüências (direção 5' &gt; 3')</b>	<b>Amplicon (pares de bases)</b>
<b><i>aat</i></b>	S - CTGGATAACCAGCAGTATCGTTC	106
	AS - GTACATTCCACCTACCAGCGTATT	
<b><i>ffh</i></b>	S - GCTAAGCCGGAAATCATCAA	104
	AS - ATGTCGTCGAACTGCTTGAG	
<b><i>glnA</i></b>	S - GAAGGCGGTAACAAAGGTCA	97
	AS - TACACATGGTGGAACGGATG	
<b><i>gyrA</i></b>	S - GTGACCCGTCGTACGATTTT	99
	AS - GATAATCGGGTCGATGTTGG	
<b><i>proC</i></b>	S - GATTGCCGATATCGTCTTCG	99
	AS - GAGACCACCAGCGACTCTTT	
<b><i>recA</i></b>	S - TTAAACAGGCCGAATTCAG	99
	AS - CCGCTTTCTCAATCAGCTTC	
<b><i>rho</i></b>	S - AACTACGACAAGCCGGAAAA	99
	AS - ACCGTTACCACGCTCCATAC	
<b><i>rpoC</i></b>	S - TATTCTGGTTCCACGCAACA	97
	AS - GGATACAACGGAACGCACTT	
<b><i>rpoD</i></b>	S - TCCGGTGCATATGATTGAGA	105
	AS - ATACGCTCAGCCAGCTCTTC	
<b><i>rrsH</i></b>	S - GACGATCCCTAGCTGGTCTG	95
	AS - GTGCAATATTCCCCACTGCT	
<b><i>trpS</i></b>	S - GCCACTGTAAGGCGCTACTC	100
	AS - GCCGATAACGTCAGCGTATT	

### 3.4 Extração de RNA total e Síntese de cDNA

RNA total foi extraído das células de *K. pneumoniae* cultivadas nas diferentes condições de cultivo descritas no item 3.2. Para a estabilização e preservação do RNA, as células foram isoladas e armazenadas em *RNAprotect*<sup>®</sup>

*Bacteria Reagent* (Qiagen) a  $-80^{\circ}\text{C}$  até o momento da extração do RNA. A extração foi realizada utilizando-se o kit *RNeasy Mini Kit* (Qiagen), seguindo instruções do fabricante. Em uma das etapas deste kit de extração há o tratamento do RNA total com *DNase*, para que haja a remoção de DNA. Em seguida, a integridade do RNA isolado foi verificada por meio de eletroforese em gel de agarose 1% contendo 6% de formaldeído em tampão de corrida MOPS (Sigma-Aldrich). Por fim, o RNA foi quantificado em espectrofotômetro e guardado a  $-80^{\circ}\text{C}$  até o momento de sua utilização na síntese de DNA complementar (cDNA).

A síntese da primeira fita de cDNA foi realizada com o RNA total extraído conforme descrito acima e a partir do kit *ThermoScript™ RT-PCR System* (Invitrogen), de acordo com as instruções do fabricante. Para a reação de síntese, a 3  $\mu\text{g}$  de RNA total foram acrescentados 2  $\mu\text{L}$  de um mix de dNTPs (concentração inicial de 10 mM), 1  $\mu\text{L}$  de *primer* hexâmero randômico (50ng/ $\mu\text{L}$ ) e quantidade suficiente de água tratada com DEPC para um volume final de 12  $\mu\text{L}$ . Esta mistura foi submetida à desnaturação por 5 minutos a  $65^{\circ}\text{C}$ . Após este período de incubação, foram acrescentados à mistura 4 $\mu\text{L}$  de tampão de reação (5x), 1  $\mu\text{L}$  de DTT (100mM), 1  $\mu\text{L}$  de *RNaseOUT™* (40U/ $\mu\text{L}$ ), 1 $\mu\text{L}$  de água tratada com DEPC e 1 $\mu\text{L}$  da enzima *ThermoScript™ RT* (15U/ $\mu\text{L}$ ). Esta mistura foi incubada por 10 minutos a  $25^{\circ}\text{C}$ , seguida de incubação por 50 minutos a  $50^{\circ}\text{C}$  para que a reação de síntese do cDNA ocorresse. A reação foi finalizada por desnaturação após incubação por 5 minutos a  $80^{\circ}\text{C}$ . Uma vez atingido a temperatura ambiente, à reação foi acrescentado 1  $\mu\text{L}$  de RNase H (2U/ $\mu\text{L}$ ) e a reação foi novamente incubada a  $37^{\circ}\text{C}$  por 20 minutos para que houvesse a degradação de qualquer RNA remanescente. Após incubação, o cDNA sintetizado foi então quantificado em espectrofotômetro e utilizado nas reações de PCR em tempo real.

### 3.5 Análises por PCR em Tempo Real

A técnica de PCR em tempo real foi empregada para investigar o padrão de expressão dos genes endógenos candidatos em células de *K. pneumoniae*

cultivadas nas diferentes condições de cultivo, conforme descritas acima. As reações foram realizadas em duplicatas no equipamento *Real-Time PCR System 7300 Applied Biosystems* e com o kit *Platinum<sup>®</sup> SYBR<sup>®</sup> Green qPCR SuperMix-UDG* (Applied Biosystems), seguindo instruções do fabricante.

A reação de PCR em tempo real consistiu em 100ng de cDNA (sintetizado conforme descrito na seção anterior), 6,25 µL de *Platinum<sup>®</sup> SYBR<sup>®</sup> Green*, 0,25 µL de *Rox Reference Dye*, 0.4 µM de cada *primer* e água em quantidade suficiente para um volume final de 12,5 µL. Em uma etapa inicial as reações foram tratadas com UDG (Uracil-DNA Glicosilase, presente no próprio kit de reação) por 2 minutos a 50° C, seguida de desnaturação por 2 minutos a 95° C. Após este pré-tratamento, as amostras foram submetidas a 40 ciclos nas seguintes condições de termo-ciclagem: 95° C por 15 segundos, 60° C por 20 segundos e 72° C por 32 segundos. Ao final, uma etapa adicional de dissociação foi realizada (40 ciclos com um decréscimo de 1° C a cada 15 segundos, iniciando-se em 95° C) para a geração de uma curva de *melting*. A análise desta curva permitiu verificar a especificidade da reação, ou seja, se não havia a presença de bandas inespecíficas e/ou dímeros de *primers*.

A análise dos dados gerados foi feita com o auxílio do software *Real-Time System RQ Study software V1.3.1 9* (Applied Biosystems). Por meio deste software determinou-se o *threshold*, ou limiar de detecção, de cada amostra testada. O *threshold* corresponde a um ponto de referência, uma vez que neste ponto todas as amostras possuem a mesma intensidade de fluorescência. Uma vez determinado o *threshold*, foi estabelecido o número de C<sub>T</sub> (do inglês *threshold cycle*) de cada amostra. O número de C<sub>T</sub> serve como base de comparação entre as amostras, pois se refere ao número de ciclo no qual a curva de amplificação de cada amostra atinge o *threshold* estabelecido.

A eficiência de amplificação (*E*) de cada um dos pares de *primers* foi determinada por meio de uma curva padrão composta de cinco pontos e que foi gerada a partir de reações em triplicata de PCR em tempo real nas quais foram utilizados diluições seriadas de cDNA.

### 3.6 Análises Estatísticas

A estabilidade da expressão dos genes candidatos nas oito condições de cultivo foi determinada por meio de três programas estatísticos diferentes: *geNorm* (VANDESOMPELE et al., 2002), *NormFinder* (ANDERSEN, JENSEN e ORNTTOFT, 2004) e *BestKeeper* (PFAFFL et al., 2004). Em comum, estes programas são baseados em planilhas do aplicativo Microsoft Excel e requerem o valor da eficácia de amplificação dos pares de *primers* para os cálculos estatísticos (Tabela 3).

Para a análise no *software geNorm* (disponível para *download* no *website* <http://medgen.ugent.be/~jvdesomp/>) é necessário, inicialmente, transformar os valores do número  $C_T$  de todos os genes e em todas as condições em quantidades relativas. Estas quantidades relativas são exportadas para o programa *geNorm* que, então, promove a comparação, aos pares, da expressão dos genes em análise. Por meio desta comparação aos pares, o programa determina o nível de variação da expressão de cada gene em comparação com a expressão dos demais genes, nas diferentes condições testadas. Desta maneira, é gerado um índice, chamado índice *M*, que reflete a estabilidade da expressão de cada gene. Um ranking dos genes mais estáveis é determinado de acordo com o índice *M* mensurado: quanto menor o índice *M*, mais estável é a expressão do gene; ao passo que genes com os maiores índices *M* significam maior variação na expressão.

A análise no *software NormFinder* (disponível para *download* no *website* <http://www.mdl.dk/publicationsnormfinder.htm>) também requer que os valores do número  $C_T$  de todos os genes e em todas as condições testadas sejam convertidos em quantidades relativas. O programa *NormFinder* considera as variações na expressão dos genes intra- e inter-tratamentos para o cálculo do índice de estabilidade. Desta forma, genes com a menor variação na expressão apresentarão o menor índice de estabilidade.

A análise no *software BestKeeper* (disponível para *download* no *website* <http://gene-quantification.com/bestkeeper.html#download>) utiliza apenas a média do número  $C_T$  dos genes analisados em todas as condições testadas, não sendo necessário a conversão para quantidades relativas como nos programas

anteriores. O programa *BestKeeper* baseia-se no coeficiente de variação do número  $C_T$  (CV [%  $C_T$ ]) e no desvio padrão das médias do número  $C_T$  para estimar a estabilidade de expressão dos genes analisados. Além disso, o programa cria um coeficiente de correlação, simbolizado por  $r$ , entre cada gene analisado e a média geométrica do valor de  $C_T$  de todos os genes candidatos juntos. Desta forma, quanto maior o coeficiente de correlação ( $r$ ), maior a estabilidade de expressão do gene.

## 4. RESULTADOS E DISCUSSÃO

### 4.1 Seleção de candidatos a genes endógenos

Para a seleção dos genes candidatos a controle endógeno, foi realizada uma pesquisa na literatura científica. Para esta pesquisa priorizou-se artigos científicos que abordassem a identificação e validação de possíveis genes endógenos para uso em análises de PCR em tempo real quantitativo em bactérias gram-negativas ou gram-positivas.

Um dos critérios utilizados para a seleção dos genes foi a de que estes genes codificassem para proteínas envolvidas nas mais diversas categorias funcionais, para garantir uma maior reprodutibilidade e segurança nos resultados finais. Neste sentido, foram selecionados genes envolvidos em diferentes atividades celulares, tais como no metabolismo celular basal, no metabolismo de ácidos nucleicos, e genes relacionados aos processos de transcrição do DNA e de síntese de proteínas. A Tabela 2 resume os 11 genes candidatos a controle endógeno que foram selecionados, de acordo com a categoria funcional em que se enquadram. Também estão representados na Tabela 2 os *loci* gênicos dos genes selecionados.

Uma vez selecionados os genes endógenos candidatos, as seqüências nucleotídicas desses genes foram obtidas a partir do banco de dados *Klebsiella pneumoniae* MGH 78578 Genome Project do NCBI (disponível no *website* <http://www.ncbi.nlm.nih.gov/bioproject/31>).

Em seguida, os pares de *primers* específicos para cada gene foram desenhados a partir das seqüências nucleotídicas dos genes e utilizando-se o programa *primer3*. Na medida do possível, os *primers* foram desenhados seguindo os seguintes parâmetros: que tivessem tamanho em torno de 20 pares de bases, com temperatura de anelamento de 60°C, que amplificassem uma região interna do gene e que o tamanho do amplicon ficasse entre 95 a 105 pares de bases. Os pares de *primers* utilizados nos experimentos de PCR convencional e de PCR em tempo real são mostrados na Tabela 2.

**TABELA 2.** Genes candidatos a controle endógeno selecionados para análise, de acordo com a categoria funcional, a proteína codificada e a função da proteína.

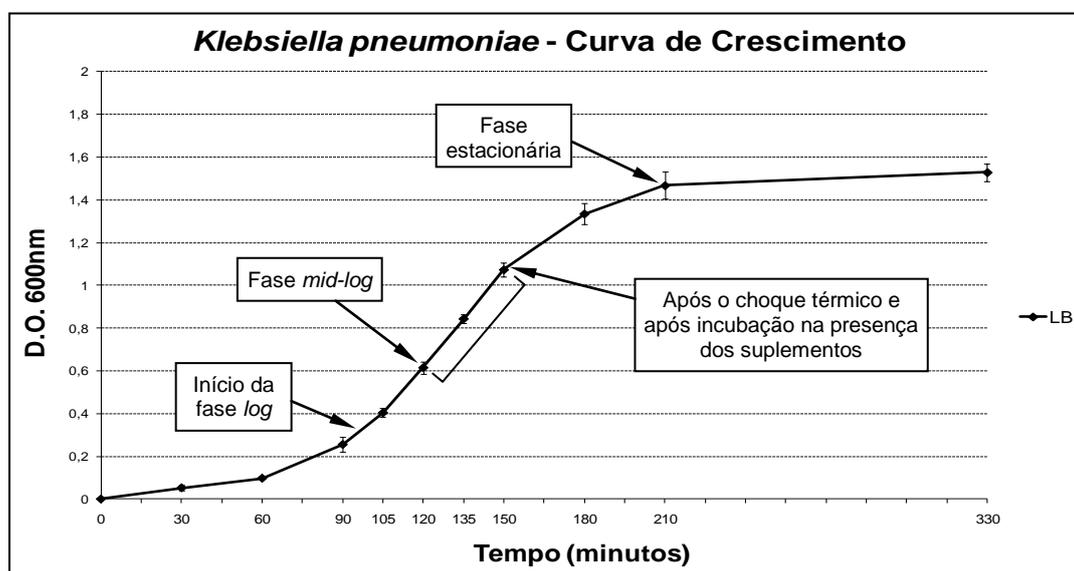
<b>Categoria funcional</b>	<b>Gene/locus</b>	<b>Produto</b>	<b>Função<sup>1</sup></b>
<b>Metabolismo celular</b>	<b><i>glnA</i></b> (KPN_04178)	Glutamina sintetase.	Enzima que catalisa a condensação de glutamato e amônia para formar glutamina (metabolismo do nitrogênio).
	<b><i>proC</i></b> (KPN_00329)	Pirrolina-5-Carboxilato Redutase.	Enzima que converte pirrolina-5-carboxilato em prolina (metabolismo do aminoácido prolina).
<b>Metabolismo do DNA</b>	<b><i>gyrA</i></b> (KPN_02640)	DNA girase.	Enzima que participa do processo de desenrolamento do DNA durante a replicação, ao quebrar uma das fitas do DNA (replicação do DNA).
	<b><i>recA</i></b> (KPN_03031)	recombinase A.	Enzima envolvida com o processo de recombinação do DNA.
<b>Transcrição do DNA</b>	<b><i>rho</i></b> (KPN_04283)	Fator de terminação da transcrição.	Fator envolvido com a terminação da transcrição.
	<b><i>rpoC</i></b> (KPN_04366)	RNA polimerase, subunidade beta.	Enzima que catalisa a polimerização de RNA dependente de DNA.
	<b><i>rpoD</i></b> (KPN_03474)	Fator sigma RpoD.	Fator de iniciação que auxilia o posicionamento da RNA polimerase em sítios específicos na região promotora dos genes alvos.
<b>Síntese Protéica</b>	<b><i>aat</i></b> (KPN_00917)	Leucil/fenilalanil-tRNA transferase.	Enzima que catalisa a transferência dos aminoácidos leucina, metionina fenilalanina e triptofano dos aminoacil-tRNAs para a extremidade amino-terminal da proteína acceptora.
	<b><i>ffh</i></b> (KPN_02933)	Proteína do complexo de reconhecimento do peptídeo sinal.	Chaperonina envolvida na translocação de proteínas e na biogênese da membrana celular.
	<b><i>rrsH</i></b> (KPN_00215)	RNA ribossômico, subunidade 16S.	Componente dos ribossomos.
	<b><i>trpS</i></b> (KPN_04752)	triptofanil-tRNA sintetase II.	Enzima que catalisa a ligação do aminoácido triptofano ao seu RNA transportador específico.

1. Função de acordo com o disposto em *Gene Ontology* (<http://www.geneontology.org/>)

## 4.2 Condições de cultivo e extração de RNA total para a síntese de cDNA

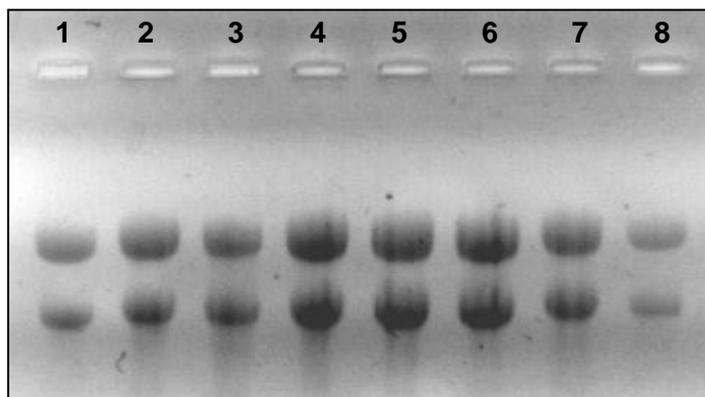
O possível padrão constitutivo de expressão dos genes selecionados foi aferido a partir do cultivo de *Klebsiella pneumoniae* em diferentes condições. Como o intuito deste trabalho foi determinar genes endógenos adequados para PCR em tempo real em qualquer que fosse a condição de cultura, procurou-se determinar o possível padrão constitutivo de expressão dos genes nas mais diferentes condições de cultivo possíveis.

Neste sentido, três situações distintas de cultivo foram abordadas neste estudo: **1)** células isoladas em diferentes fases de crescimento (fases log, mid-log e estacionária); **2)** células incubadas na presença dos suplementos FeSO<sub>4</sub>, 2,2'-dipiridil e MnSO<sub>4</sub>, e **3)** células submetidas ao choque térmico, em temperatura abaixo (29°C) e acima (45°C) da temperatura ótima de crescimento de *K. pneumoniae*. A Figura 1 mostra a curva de crescimento de *Klebsiella pneumoniae* em meio LB.



**FIGURA 1.** Curva de crescimento de *Klebsiella pneumoniae* em meio de cultura LB. As setas e boxes indicam os pontos em que as células de *K. pneumoniae* foram isoladas para extração do RNA total. O colchete indica o período (30 minutos) em que as células foram incubadas na presença dos suplementos (FeSO<sub>4</sub>, 2,2'-dipiridil e MnSO<sub>4</sub>) e submetidas ao choque térmico (29°C e 45°C).

A Figura 2 mostra o RNA total extraído de células de *K. pneumoniae* cultivadas nas condições acima. Como pode ser observado na Figura 2, a extração resultou em RNA total íntegro e apropriado para as reações subseqüentes de síntese de cDNA.

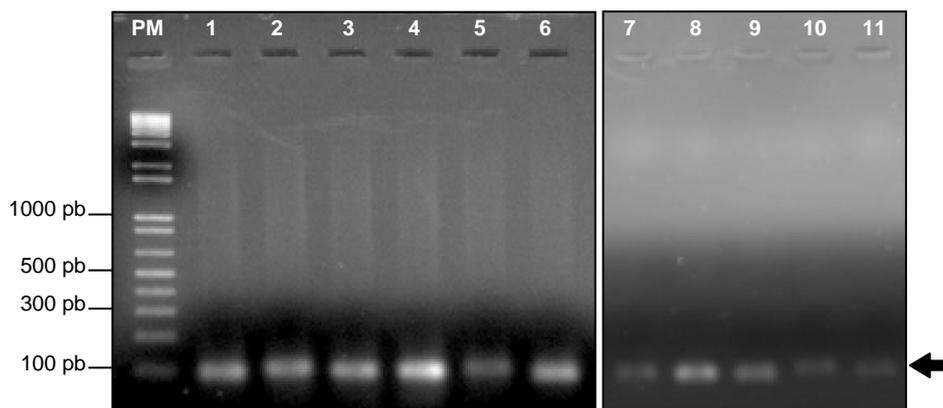


**FIGURA 2.** Eletroforese em gel de agarose mostrando a integridade dos RNAs totais extraídos de *Klebsiella pneumoniae*. As bandas referem-se aos RNAs ribossômicos subunidades 23S e 16S. O RNA total foi extraído das seguintes condições de cultivo: **1.** Cultivo até início da fase *log*; **2.** Cultivo até fase *mid-log*; **3.** Cultivo até fase estacionária; **4.** Células incubadas por 30 minutos na presença de  $\text{FeSO}_4$ ; **5.** Incubadas por 30 minutos na presença de 2,2'-dipiridil; **6.** Incubadas por 30 minutos na presença de  $\text{MnSO}_4$ ; **7.** Células incubadas por 30 minutos a 29°C, e **8.** Incubadas por 30 minutos a 45°C.

#### 4.3 Teste de eficiência dos pares de *primers* por PCR convencional e por PCR em tempo real

A eficiência de amplificação dos pares de *primers* desenhados foi checada por meio de PCR convencional e por meio de PCR em tempo real.

Inicialmente os *primers* foram testados por PCR convencional, a partir de DNA genômico isolado de *K. pneumoniae*. Conforme pode ser observado na Figura 3, todos os pares de *primers* renderam amplicons com tamanhos que correspondem ao esperado (conforme mostrado na Tabela 1).



**FIGURA 3.** Eletroforese em gel de agarose mostrando os amplicons (seta) gerados por PCR convencional a partir dos pares de *primers* desenhados para os genes em análise. **PM.** Marcador de peso molecular de 100 pares de bases (pb); **1.** Gene *rrsH*; **2.** Gene *ffh*; **3.** Gene *glnA*; **4.** Gene *gyrA*; **5.** Gene *aat*; **6.** Gene *proC*; **7.** Gene *recA*; **8.** Gene *rho*; **9.** Gene *rpoC*; **10.** Gene *rpoD*; **11.** Gene *trpS*.

Em seguida, os pares de *primers* tiveram sua eficiência de amplificação determinada por meio de uma curva padrão gerada por reações de PCR em tempo real nas quais foram utilizadas diluições seriadas de cDNA. A curva padrão gerada para cada um dos genes é representada pelo o valor  $C_T$  correspondente versus o log da quantidade de cDNA utilizada na reação. A partir do *slope* desta curva foi possível calcular a eficiência de amplificação ( $E$ ) dos *primers*, por meio da fórmula:

$$E = [10^{(-1/Slope)} - 1] \times 100$$

Para uma quantificação confiável em PCR em tempo real, o valor do *slope* deve estar num intervalo entre  $-3,3$  a  $-3,8$ , o que corresponde a uma eficiência de amplificação de 83 a 100%. Conforme pode ser observado na Tabela 3, somente o par de *primers* para o gene *trpS* encontra-se fora deste intervalo, resultado que já indica para uma exclusão deste gene nas análises futuras. Os pares de *primers* dos demais genes encontram-se dentro do intervalo, o que valida a sua utilização em experimentos de PCR em tempo real.

**TABELA 3.** Cálculo da eficiência de amplificação dos pares de *primers* utilizados nas reações de PCR em tempo real.

<b>Genes</b>	<b>Slope</b>	<b>R<sup>2</sup> *</b>	<b>Eficácia **</b>	<b>Eficiência (%)</b>
<i>aat</i>	-3,77	0,99	1,84	84,1
<i>ffh</i>	-3,39	0,98	1,97	97,2
<i>glnA</i>	-3,36	0,98	1,98	98,1
<i>gyrA</i>	-3,40	0,98	1,96	96,6
<i>proC</i>	-3,79	0,98	1,83	83,4
<i>recA</i>	-3,52	0,98	1,92	92,1
<i>rho</i>	-3,51	0,99	1,92	92,4
<i>rpoC</i>	-3,55	0,97	1,91	91,2
<i>rpoD</i>	-3,66	0,98	1,87	87,5
<i>rrsH</i>	-3,47	0,95	1,94	94,2
<i>trpS</i>	-3,13	0,99	2,08	108,5

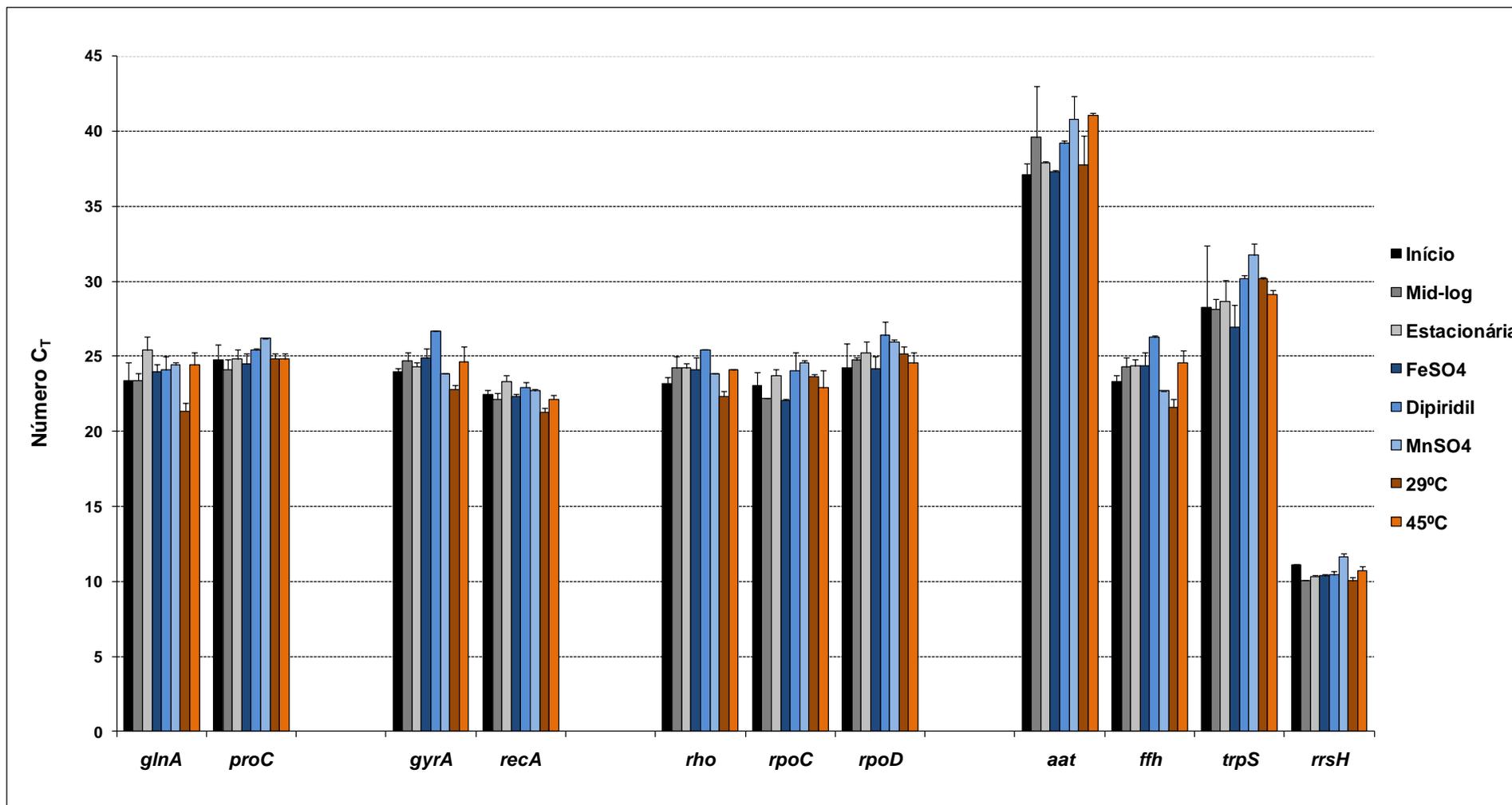
\* **R<sup>2</sup>**, coeficiente de correlação do *slope* da curva padrão.

\*\* **Eficácia**, calculado de acordo com a fórmula  $10^{(-1/slope)}$ . Um valor de eficácia de 2,0 indica uma reação de PCR 100% eficiente.

Os cálculos de eficiência de amplificação são importantes porque o valor da eficácia dos pares de *primers* é necessário para as análises estatísticas, como será descrito mais adiante.

#### 4.4 Análises por PCR em tempo real

Após os pares de *primers* terem sido testados e a eficiência da amplificação mensurada, procedeu-se à investigação do padrão de expressão dos genes candidatos nas condições testadas por meio de reações de PCR em tempo real. A Figura 4 mostra a média do número C<sub>T</sub> de cada gene em cada uma das oito diferentes condições de cultivo.



**FIGURA 4.** Gráfico com o número  $C_T$  de cada gene obtido por PCR em tempo real em cada uma das oito diferentes condições de cultivo testadas. Cada barra representa a média do número  $C_T$ , e as barras de erros representam os desvios padrões de cada uma das médias.

Como pode ser observado na Figura 4, alguns genes apresentaram uma variação pronunciada do número  $C_T$  dentre as condições testadas, ao passo que outros genes apresentaram uma variação menor.

Estas variações são melhores visualizadas na Tabela 4, que mostra os valores do número  $C_T$  mínimo e máximo observado para cada gene nas condições testadas, além da média geral do número  $C_T$  em todas estas condições e o respectivo desvio padrão da média. De acordo com a Tabela 4, os genes que apresentaram maior variação no número  $C_T$  foram os genes *aat* e *trpS*, com desvios padrões de 1,55 e 1,50 respectivamente. Por outro lado, os genes *recA* e *rrsH* apresentaram a menor variação do número  $C_T$ , com desvios padrões de 0,61 e 0,55 respectivamente.

**TABELA 4.** Variação dos números  $C_{Ts}$  de cada gene nas oito condições de cultivo testadas, com a média geral dos  $C_{Ts}$  e respectivos desvios padrões.

<b>Genes</b>	<b>Variação do <math>C_T</math> (Mínimo – máximo)</b>	<b>Média dos <math>C_{Ts}</math></b>	<b>Desvio Padrão</b>
<i>aat</i>	37,10 - 41,05	38,83	1,55
<i>ffh</i>	21,61 - 26,25	23,94	1,39
<i>glnA</i>	21,34 - 25,42	23,79	1,19
<i>gyrA</i>	22,81 - 26,66	24,48	1,10
<i>proC</i>	24,12 - 26,18	24,92	0,62
<i>recA</i>	21,27 - 23,33	22,40	0,61
<i>rho</i>	22,32 - 25,45	23,93	0,90
<i>rpoC</i>	22,09 - 24,57	23,28	0,88
<i>rpoD</i>	24,15 - 26,40	25,06	0,80
<i>rrsH</i>	10,04 - 11,67	10,60	0,55
<i>trpS</i>	26,91 - 31,73	29,15	1,50

Após a reação de PCR em tempo real na qual foram gerados os  $C_{Ts}$  mostrados na Figura 4, foi realizada uma etapa adicional de dissociação dos amplicons para a geração de uma curva de *melting*. A análise desta curva é

importante, pois permite verificar a especificidade de amplificação dos *primers* e da reação, na medida em que detecta a presença de possíveis bandas inespecíficas e/ou dímeros de *primers*.

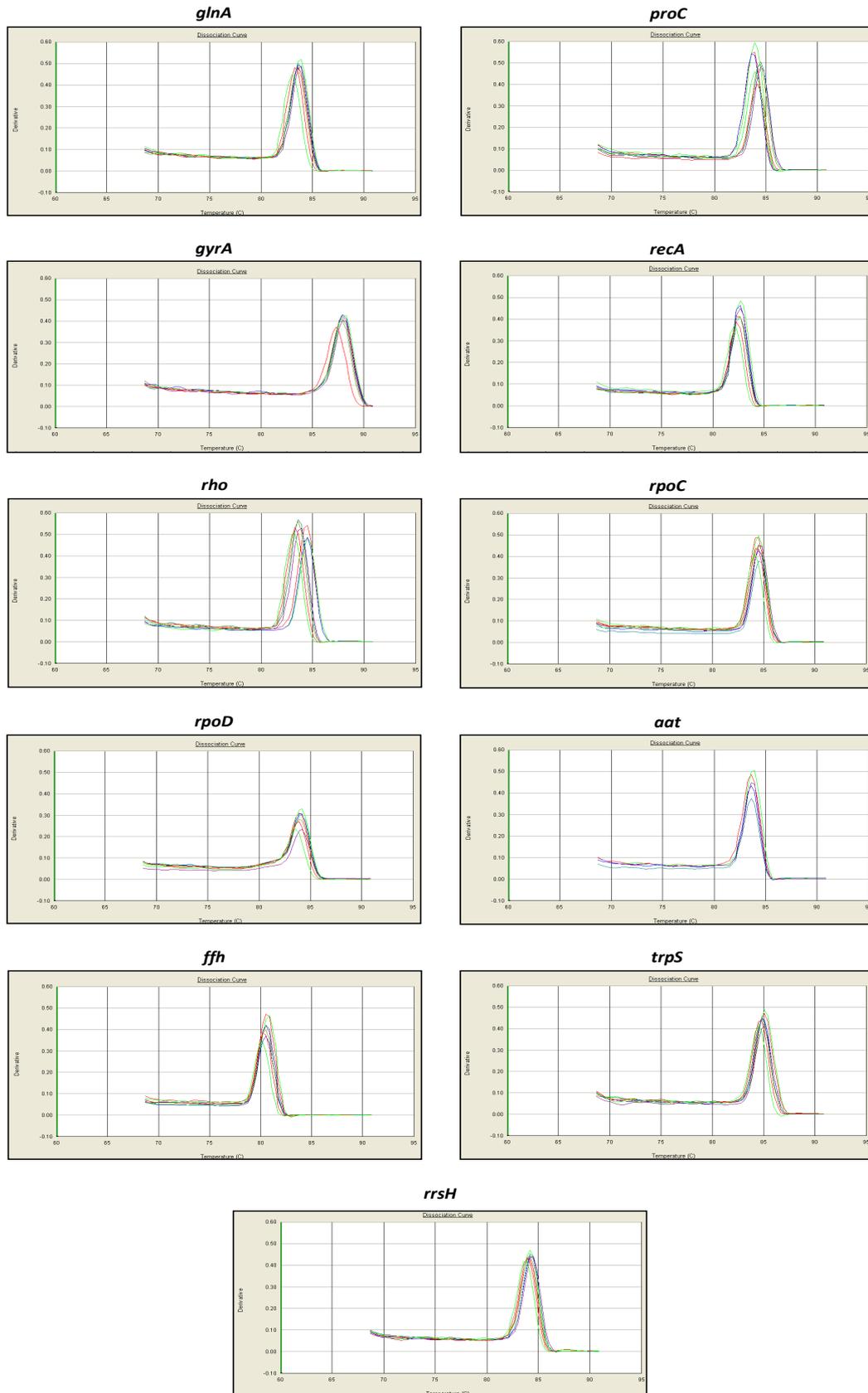
A Figura 5 mostra os gráficos com as curvas de *melting*, ou de dissociação, de cada um dos genes analisados. A presença de apenas um pico nas curvas de dissociação revela a ausência de bandas inespecíficas, o que confirma a especificidade de amplificação dos *primers* utilizados no PCR em tempo real.

#### 4.5 Análises Estatísticas

Três programas estatísticos diferentes foram utilizados para se calcular a estabilidade da expressão dos genes candidatos nas diferentes condições testadas: *geNorm* (VANDESOMPELE et al., 2002), *NormFinder* (ANDERSEN, JENSEN e ORNTOFT, 2004) e *BestKeeper* (PFAFFL et al., 2004).

Inicialmente a estabilidade de expressão dos onze genes candidatos foi aferida por meio do programa *BestKeeper*. O programa *BestKeeper* baseia-se no coeficiente de correlação ( $r$ ), no coeficiente de variação do número  $C_T$  (CV - %  $C_T$ ) e no desvio padrão das médias do número  $C_T$  para estimar a estabilidade de expressão dos genes analisados. Estes dados estão resumidos na Tabela 5.

De acordo com os dados estatísticos gerados pelo *BestKeeper* e mostrados na Tabela 5, os genes *aat* e *trpS* apresentam elevado coeficiente de variação (4,03 e 5,14 respectivamente) e os maiores desvios padrões tanto da média dos  $C_{Ts}$  ( $\pm C_T$ , 1,56 e 1,50, respectivamente) quanto dos coeficientes absolutos de regulação ( $\pm x\text{-fold}$ , 2,82 e 2,70, respectivamente). De fato, conforme pode ser observado na Figura 4 e Tabela 4, *aat* e *trpS* foram os genes que apresentaram maior variação no número  $C_T$  nas condições testadas dentre todos os genes analisados. Por estas razões, e considerando que genes com desvio padrão ( $\pm C_T$ ) maior que 1 são considerados inconsistentes pelo *BestKeeper*, os genes *aat* e *trpS* foram excluídos das análises estatísticas com os programas *geNorm* e *NormFinder*.



**FIGURA 5.** Curva de *melting* (dissociação) indicando a especificidade de amplificação dos pares de *primers* utilizados na PCR em tempo real.

**TABELA 5.** Dados estatísticos gerados pelo *Bestkeeper* para os 11 genes candidatos.

<b>Genes</b>	<b>Média geométrica</b> [dos C <sub>Ts</sub> ]	<b>Desvio padrão</b> [± C <sub>T</sub> ]	<b>Coefficiente de Variação</b> <sup>1</sup> [% C <sub>T</sub> ]	<b>Desvio Padrão</b> <sup>2</sup> [± x-fold]
<i>aat</i> *	38,79	1,56	4,03	2,82
<i>ffh</i>	23,90	1,00	4,49	2,04
<i>glnA</i>	23,76	0,95	4,01	1,88
<i>gyrA</i>	24,45	0,84	3,45	1,75
<i>proC</i>	24,91	0,59	2,37	1,48
<i>recA</i>	22,40	0,69	2,10	1,37
<i>rho</i>	23,91	0,47	2,90	1,58
<i>rpoC</i>	23,26	0,83	3,59	1,74
<i>rpoD</i>	25,04	0,74	2,97	1,64
<i>rrsH</i>	10,58	0,43	4,10	1,33
<i>trpS</i> *	29,09	1,50	5,14	2,70

1. Coeficiente de variação, expresso como porcentagem do número do C<sub>T</sub>.

2. Desvio padrão dos coeficientes absolutos de regulação.

\* Genes excluídos das análises pelos programas *NormFinder* e *geNorm*.

A Tabela 6 resume todos os cálculos estatísticos gerados pelos programas *BestKeeper*, *NormFinder* e *geNorm* e mostra um ranking dos genes que tiveram a expressão mais estável nas diferentes condições testadas.

De acordo com o programa *BestKeeper*, quanto maior o coeficiente de correlação ( $r$ ) e menor o desvio padrão do gene, maior a estabilidade de expressão do gene, o que aumenta a chance deste gene ser selecionado como controle endógeno. Conforme mostrado na Tabela 6, *BestKeeper* indica os genes *recA* ( $r = 0,849$ ), *rho* ( $r = 0,816$ ) e *glnA* ( $r = 0,814$ ) com os maiores coeficiente de correlação e  $p$ -value de 0,001. Assim, de acordo com *BestKeeper* estes genes apresentam a expressão mais estável, sendo, portanto, os mais indicados para controle endógeno.

**Tabela 6.** Cálculos estatísticos gerados pelos programas *BestKeeper*, *NormFinder* e *geNorm* e ranking dos genes que tiveram a expressão mais estável nas diferentes condições testadas. Em negrito, os genes com os melhores escores nas diferentes análises estatísticas.

	<i>BestKeeper</i>			<i>Normfinder</i>			<i>geNorm</i>	
	<i>r</i>	<i>p-value</i>	Ranking	Índice de estabilidade	Erro padrão	Ranking	Índice <i>M</i>	Ranking
<i>ffh</i>	0,712	0,002	5	0,760	0,148	8	0,99	9
<i>glnA</i>	0,814	0,001	3	0,776	0,150	9	0,92	8
<i>gyrA</i>	0,739	0,001	4	0,535	0,112	5	0,86	7
<i>proC</i>	0,66	0,005	7	0,423	0,096	3	<b>0,44</b>	<b>1/2</b>
<i>rho</i>	<b>0,816</b>	<b>0,001</b>	<b>2</b>	<b>0,328</b>	<b>0,085</b>	<b>2</b>	0,78	6
<i>recA</i>	<b>0,849</b>	<b>0,001</b>	<b>1</b>	<b>0,296</b>	<b>0,082</b>	<b>1</b>	0,67	5
<i>rpoC</i>	0,541	0,03	8	0,614	0,124	6	<b>0,44</b>	<b>1/2</b>
<i>rpoD</i>	0,687	0,003	6	0,674	0,134	7	0,49	3
<i>rrsH</i>	0,42	0,10	9	0,515	0,109	4	0,60	4

Apesar de apresentar expressão estável nas oito condições de cultivo testadas (Figura 4 e Tabela 4), o uso do gene *rrsH*, que codifica RNA ribossômico da subunidade 16S, não é recomendado pelo programa *BestKeeper* para análises de PCR em tempo real. De fato, o gene *rrsH* apresenta elevado coeficiente de variação (4,10, Tabela 5) e o menor coeficiente de correlação ( $r = 0,42$ , nono lugar no ranking da Tabela 6). Em adição a isso, a utilização do gene *rrsH* não é recomendado para análises de PCR em tempo real devido à elevada abundância dos transcritos deste gene, como pode ser observado pelo número  $C_T$  baixo (média de 10,60, Tabela 4 e Figura 4). Esta elevada abundância inviabiliza a quantificação de transcritos mais raros e, portanto, menos abundantes.

O programa *NormFinder* analisa a estabilidade de expressão dos genes candidatos de maneira individual e independente. Segundo este programa, genes com os menores índices de estabilidade apresentarão a menor variação na expressão. Conforme mostrado na Tabela 6, *NormFinder* indica os genes, *recA*, *rho* e *proC* com os menores índices de estabilidade (Tabela 6). Os genes *rrsH*, *gyrA*, *rpoC* e *rpoD* apresentam índices de estabilidade intermediários. Ao passo que os genes *ffh* e *glnA* possuem os maiores índices de estabilidade e, portanto, níveis de expressão não tão estáveis quanto os outros genes.

Ao contrário do *NormFinder*, que analisa a estabilidade dos genes individualmente, o programa *geNorm* promove a comparação aos pares da expressão dos genes candidatos. Desta forma, um índice  $M$  é gerado para cada gene, e este índice é utilizado para se determinar o gene que apresenta expressão mais estável. O programa *geNorm* recomenda índice  $M$  abaixo do limite de 1.5, e quanto menor o índice  $M$ , mais estável é a expressão do gene. Neste sentido, o ranking dos genes de acordo com o índice  $M$  é mostrado na Tabela 6.

Conforme pode ser observado na Tabela 6, todos os 9 genes analisados pelo *geNorm* apresentam índice  $M$  abaixo do limite de 1.5. Os genes *rpoC*, *proC* e *rpoD* tiveram a maior estabilidade na expressão (menor índice  $M$ ), ao passo que os genes *rrsH*, *recA* e *rho* tiveram estabilidade e índices  $M$  intermediários. Os genes *gyrA*, *glnA* e *ffh* tiveram os maiores índices  $M$  e, portanto, apresentam níveis de expressão não tão estáveis quanto os outros genes. No entanto,

segundo o *geNorm* estes genes apresentam índices *M* abaixo do limite de 1.5 e, portanto, ainda assim possuem expressão com estabilidade confiável.

A análise em conjunto dos três programas estatísticos utilizados revela algumas diferenças com relação ao ranking dos genes com expressão mais e menos estável. Esta aparente discrepância, que já foi relatada por alguns estudos (HIBBELER et al., 2008; NIETO et al., 2009; PAOLACCI et al., 2009), é de certa maneira esperada na medida em que cada programa realiza os cálculos estatísticos por meio de metodologias distintas.

Tanto *geNorm* quanto *NormFinder* indicaram os genes *ffh* e *glnA* como os genes com menor estabilidade de expressão (8<sup>a</sup> e 9<sup>a</sup> posição no ranking). Porém este resultado não foi detectado pelo programa *BestKeeper*, que indicou os genes *ffh* e *glnA* na 5<sup>a</sup> e 3<sup>a</sup> posição do ranking, respectivamente.

Com relação aos genes com maior estabilidade de expressão, de maneira geral os três programas sugerem os genes *recA*, *rho*, *proC* e *gyrA* como os mais adequados para análises em PCR em tempo real em *Klebsiella pneumoniae*. Ambos os programas *NormFinder* e *Bestkeeper* indicaram os genes *rho* e *recA* como os genes com maior estabilidade de expressão e, portanto, os melhores candidatos a controle endógeno. Ainda que não estejam entre os primeiros colocados no ranking do *geNorm*, os genes *recA* e *rho* tiveram índices *M* abaixo do limite de 1,5 ( $M = 0,67$  e  $M = 0,78$ , respectivamente) e, portanto, apresentam expressão com estabilidade confiável.

## 5. CONCLUSÃO

A análise da expressão gênica por PCR em tempo real é cada vez mais importante na pesquisa biológica, tornando-se o método de escolha para o estudo do perfil de expressão dos genes de interesse. Em microbiologia, a PCR em tempo real é considerada uma ferramenta poderosa para analisar a expressão de genes de virulência de bactérias patogênicas nos tecidos do hospedeiro infectado.

A avaliação correta dos dados de PCR em tempo real exige normalizações precisas e confiáveis, e isto requer o emprego de um ou mais genes endógenos que atuam como controle interno das reações. O gene endógeno ideal deve apresentar expressão estável e que não muda durante o curso do experimento. Com o emprego de um controle interno, as amostras podem ser comparadas entre si, independentemente das condições experimentais ou dos tratamentos analisados.

Diversos estudos em bactérias têm abordado a identificação de genes endógenos adequados para análises por PCR em tempo real. Contudo, não há relatos na literatura de estudos que abordam este tema em *Klebsiella pneumoniae*, o que ressalta o ineditismo deste trabalho. Atualmente, o gene endógeno freqüentemente empregado para *K. pneumoniae* é o gene *rrsH*, que codifica o RNA ribossômico 16S. Entretanto, existem restrições quanto à utilização de um RNA ribossômico para a normalização de RNAs mensageiros.

O presente trabalho teve como objetivo a identificação e validação de genes endógenos adequados para uso em análises de PCR em tempo real em *Klebsiella pneumoniae*. Genes potencialmente candidatos a controle endógeno foram selecionados e tiveram o padrão de expressão aferido por PCR em tempo real a partir do cultivo de *Klebsiella pneumoniae* em diferentes condições. Os dados gerados foram estatisticamente analisados por meio dos programas *geNORM*, *NormFinder* e *BestKeeper* para se calcular a estabilidade da expressão dos genes selecionados.

A partir dos resultados aqui apresentados é possível concluir que o gene *rrsH* deve ser descartado como controle endógeno, pois apresentou elevado coeficiente de variação e o menor coeficiente de correlação, conforme atestam as análises pelo programa *BestKeeper* (última posição no ranking deste programa).

Além disso, a elevada abundância do RNA ribossômico 16S, constatada pelo número  $C_T$  baixo, inviabiliza a quantificação de transcritos mais raros e, portanto, menos abundantes. Em contrapartida, *recA*, *rho*, *proC* e *gyrA* são os genes mais adequados para atuarem como controle endógeno em análises de PCR em tempo real em *K. pneumoniae*, pois foram os que apresentaram a maior estabilidade de expressão nas condições testadas.

Em suma, o presente trabalho sugere os genes *recA* e *rho* como os melhores candidatos a controle endógeno para análises de expressão gênica por PCR em tempo real em *Klebsiella pneumoniae*, uma vez que foram os genes com a maior estabilidade de expressão e melhores colocados nos rankings dos programas *NormFinder* e *Bestkeeper*.

## 6. REFERÊNCIAS

- ACHENBACH, L.A., e YANG, W. *The fur gene from Klebsiella pneumoniae: characterization, genomic organization and phylogenetic analysis.* **Gene**, 185(2): 201-207, 1997.
- ANDERSEN, C.L.; JENSEN, J.L., e ORNTOFT, T.F. *Normalisation of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalisation, applied to bladder and colon cancer data sets.* **Cancer Res.**, 64: 5245-5250, 2004.
- ANDREWS, S.C.; ROBINSON, A.K., e RODRÍGUEZ-QUIÑONES, F. *Bacterial iron homeostasis.* **FEMS Microbiol. Rev.**, 27: 215-237, 2003.
- BANIN, E.; VASIL, M.L., e GREENBERG, E.P. *Iron and Pseudomonas aeruginosa biofilm formation.* **Proc. Nat. Acad. Sci. USA**, 102: 11076-11081, 2005.
- CLARKE, T.E.; TARI, L.W., e VOGEL, H.J. *Structural biology of bacterial iron uptake systems.* **Curr. Top. Med. Chem.**, 1: 7-30, 2001.
- COLE, S.P.; HARWOOD, J.; LEE, R.; SHE, R., e GUINEY, D.G. *Characterization of monospecies biofilm formation by Helicobacter pylori.* **J. Bacteriol.**, 186: 3124-3132, 2004.
- COOKSLEY, C.; JENKS, P.J.; GREEN, A.; COCKAYNE, A.; LOGAN, R.P., e HARDIE, K.R. *NapA protects Helicobacter pylori from oxidative stress damage, and its production is influenced by the ferric uptake regulator.* **J. Med. Microbiol.**, 52: 461-469, 2003.
- DESROCHE, N.; BELTRAMO, C., e GUZZO, J. *Determination of an internal control to apply reverse transcription quantitative PCR to study stress response in the lactic acid bacterium Oenococcus oeni.* **J. Microbiol. Methods**, 60: 325-333, 2005.
- DEUTSCHER, M.P. *Degradation of RNA in bacteria: comparison of mRNA and stable RNA.* **Nucleic Acids Res.**, 34: 659-666, 2006.
- DHEDA, K.; HUGGETT, J.F.; BUSTIN, S.A.; JOHNSON, M.A.; ROOK, G., e ZUMLA, A. *Validation of housekeeping genes for normalizing RNA expression in real-time PCR.* **Biotechniques**, 37(1): 112-119, 2004.

- ESCOLAR, L.; PEREZ-MARTIN, J., e LORENZO, V. *Opening the iron box: transcriptional metalloregulation by the Fur protein.* **J. Bacteriol.**, 181(2): 6223-6229, 1999.
- FERDELE, M.J., e BASSLER, B.L. *Interspecies communication in bacteria.* **J. Clin. Investig.**, 121: 1291-1299, 2002.
- FROST, P., e NILSEN, F. *Validation of reference genes for transcription profiling in the salmon louse, Lepeophtheirus salmonis, by quantitative real-time PCR.* **Vet. Parasitol.**, 118: 169-174, 2003.
- GRALLA, J.D. *Escherichia coli ribosomal RNA transcription: regulatory roles for ppGpp, NTPs, architectural proteins and a polymerase-binding protein.* **Mol. Microbiol.**, 55: 973-977, 2005.
- HALL, H.K., e FOSTER, J.W. *The role of Fur in acid tolerance response of Salmonella typhimurium is physically and genetically separable from its role in iron acquisition.* **J. Bacteriol.**, 178: 5683-5691, 1996.
- HAYS, J.P. *The evolution of putative endogenous control housekeeping genes for real-time polymerase chain reaction expression studies in Moraxella catarrhalis.* **Diagn. Microbiol. Infect Dis.**, 65: 323-326, 2009.
- HEINRICH, M.; LUTZ-BONENGEL, S.; MATT, K., e SCHMIDT, U. *Real-time PCR detection of five different "endogenous control gene" transcripts in forensic autopsy material.* **Forensic Sci. Int. Genet.**, 1(2): 163-169, 2007.
- HIBBELER, S.; SCHARSACK, J.P., e BECKER, S. *Housekeeping genes for quantitative expression studies in the three-spined stickleback Gasterosteus aculeatus.* **BMC Mol Biol.**, 9: 18, 2008.
- KHEIRELSEID, E.A.H.; CHANG, K.H.; NEWELL, J.; KERIN, M.J., e MILLER, N. *Identification of endogenous control genes for normalization of real-time quantitative PCR data in colorectal cancer.* **BMC Mol. Biol.**, 11: 12, 2010.
- KONEMAN, E.W.; ALLEN, S.D.; DOWEL JR., V.R., e SOMMER, H.M. **Diagnóstico microbiológico: texto e atlas colorido. Enterobacteriaceae.** 5ª Ed. Rio de Janeiro: MDSI, 2001.
- LITWIN, C.M., e CALDERWOOD, S.B. *Role of iron in regulation of virulence genes.* **Clin. Microbiol. Rev.**, 6: 137-149, 1993.

- MAH, T.F., e O'TOOLE, G.A. *Mechanisms of biofilm resistance to antimicrobial agents*. **Trends Microbiol.**, 9: 34-39, 2001.
- MALOUKH, L.; MATOUSEK, J.; BOCKSTAELE, E.V., e ROLDÁN-RUIZ, I. *Housekeeping gene selection for real time PCR normalization in female HOP (Humulus lupulus L) tissues*. **J. Plant Biochem. Biotechnol.**, 18(1): 53-58, 2009.
- MAROUFI, A.; BOCKSTAELE, E.V., e LOOSE, M. *Validation of reference genes for gene expression analysis in chicory (Cichorium intybus) using quantitative real-time PCR*. **BMC Mol. Biol.**, 11: 15, 2010.
- NIETO, P.A.; COVARRUBIAS, P.C.; JEDLICKI, E.; HOLMES, D.S., e QUATRINI, R. *Selection and evaluation of reference genes for improved interrogation of microbial transcriptomes: case study with the extremophile Acidithiobacillus ferrooxidans*. **BMC Mol. Biol.**, 10: 63, 2009.
- OLSVIK, P.A.; LIE, K.K.; JORDAL, A.E.O.; NILSEN, R.O., e HORDVIK, I. *Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon*. **BMC Mol. Biol.**, 6: 21, 2005.
- O'TOOLE, G.A., e KOLTER, R. *Initiation of biofilm formation in Pseudomonas fluorescens WCS365 proceeds via multiple, convergent signaling pathways: a genetic analysis*. **Mol. Microbiol.**, 28: 449-461, 1998.
- O'TOOLE, G., KAPLAN, H.B., e KOLTER, R. *Biofilm formation as microbial development*. **Annu. Rev. Microbiol.**, 54: 49-79, 2000.
- PANINA, E.M.; MIRONOV, A.A., e GELFAND, M.S. *Comparative analysis of Fur regulons in gamma-proteobacteria*. **Nucleic Acids Res.**, 29(24): 5195-5206, 2001.
- PAOLACCI, A.R.; TANZARELLA, O.A.; PORCEDDU, E., e CIAFFI, M. *Identification and validation of reference genes for quantitative RT-PCR normalization in wheat*. **BMC Mol. Biol.**, 10: 11, 2009.
- PATERSON, D.L. *Resistance in gram-negative bacteria: Enterobacteriaceae*. **Am. J. Infect. Control.**, 34(5): S20-S28, 2006.
- PEREIRA, A.S.; CARMO FILHO, J.R.; TOGNIM, M.C.B., e SADER, H.S. *Avaliação da acurácia de testes laboratoriais para detecção de amostra de*

- Klebsiella pneumoniae* produtora de betalactamase de espectro estendido. **J. Bras. Patol. Medic. Laborat.**, 39(4): 301-308, 2003.
- PFAFFL, M.W.; TICHOPAD, A.; PRGOMET, C., e NEUVIANS, T.P. *Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pairwise correlations.* **Biotechnol. Lett.**, 26: 509-515, 2004.
- QUATRINI, R.; LEFIMIL, C.; HOLMES, D.S., e JEDLICH, E. *The ferric iron uptake regulator (Fur) from the extreme acidophile Acidithiobacillus ferrooxidans.* **Microbiology**, 151: 2005-2015, 2005.
- RADONIC, A.; THULKE, S.; MACKAY, I.M.; LANDT, O.; SIEGERT, W., e NITSCHKE, A. *Guideline to reference gene selection for quantitative realtime PCR.* **Biochem. Biophys. Res. Comm.**, 313(4): 856-862, 2004.
- ROZEN, S., e SKALETSKY, H. *Primer3 on the WWW for general users and for biologist programmers.* In: Krawetz S, Misener S (eds) **Bioinformatics methods and protocols: methods in molecular biology.** Humana Press, Totowa, pp 365–386, 2000.
- SCHALK, I.J.; YUE, W.W., e BUCHANAN, S.K. *Recognition of iron-free siderophores by TonB-dependent iron transporters.* **Mol. Microbiol.**, 54: 14-22, 2004.
- SINGH, P.K. *Iron sequestration by human lactoferrin stimulates P. aeruginosa surface motility and blocks biofilm formation.* **Biomaterials**, 17: 267-270, 2004.
- SILVEIRA, E.D.; ALVES-FERREIRA, M.; GUIMARÃES, L.A.; SILVA, F.R., e CARNEIRO, V.T.C. *Selection of reference genes for quantitative real-time PCR expression studies in the apomictic and sexual Grass Brachiaria brizantha.* **BMC Plant Biol.**, 9: 84, 2009.
- STOJILJKOVIC, J., e HANTKE, K. *Functional domains of the Escherichia coli ferric uptake regulator protein (Fur).* **Mol. Gen. Genet.**, 247: 199-205, 1995.
- TAKLE, G.W.; TOTH, I.K., e BRUGER, M.B. *Evaluation of reference genes for real-time RT-PCR expression studies in the plant pathogen Pectobacterium atrosepticum.* **BMC Mol. Biol.**, 7: 50, 2007.

- THEIS, T., SKURRAY, R.A., e BROWN, M.H. *Identification of suitable internal controls to study expression of a Staphylococcus aureus multidrug resistance system by quantitative real-time PCR.* **J. Microbiol. Methods**, 70: 355-362, 2007.
- VANDECASTEELE, S.J.; PEETERMANS, W.E.; MERCKX, R., e VAN ELDERE, J. *Quantification of expression of Staphylococcus epidermidis housekeeping genes with Taqman quantitative PCR during in vitro growth and under different conditions.* **J. Bacteriol.**, 183, 7094-7101, 2001.
- VANDESOMPELE, J.; DE PRETER, K.; PATTYN, F.; POPPE, B.; VAN ROY, N.; DE PAEPE, A., e SPELEMAN, F. *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.* **Genome Biol.**, 3(7): research0034.1-0034.11, 2002.
- WAGNER, R. *Regulation of ribosomal RNA synthesis in E. coli: effects of the global regulator guanosine tetraphosphate (ppGpp).* **J. Mol. Microbiol. Biotechnol.**, 4: 331-340, 2002.
- WALKER, C.G.; MEIER, S.; MITCHELL, M.D.; ROCHE, J.R., e LITTLEJOHN, M. *Evaluation of real-time PCR endogenous control genes for analysis of gene expression in bovine endometrium.* **BMC Mol. Biol.**, 10: 100, 2009.