

ADRIANO CARLOS SOARES

**IDENTIFICAÇÃO DE LIPOCALINAS COM DOMÍNIO DE LIGAÇÃO A
HISTAMINA E SEROTONINA E VALIDAÇÃO DE GENES DE
REFERÊNCIA EM TRANSCRIPTOMAS DE GLÂNDULA SALIVAR,
INTESTINO E OVÁRIO DO CARRAPATO *Amblyomma sculptum***

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2015

**Ficha catalográfica preparada pela Biblioteca Central da
Universidade Federal de Viçosa - Câmpus Viçosa**

T

S676i
2015 Soares, Adriano Carlos, 1972-
Identificação de lipocalinas com domínio de ligação a histamina e serotonina e validação de genes de referencia em transcriptomas de glândula salivar, intestino e ovário do carrapato *Amblyomma sculptum* / Adriano Carlos Soares. - Viçosa, MG, 2015.
xvii, 113f. : il. (algumas color.) ; 29 cm.

Inclui anexos.

Orientador : Cláudio Lisias Mafra de Siqueira.
Tese (doutorado) - Universidade Federal de Viçosa.
Referências bibliográficas: f.93-103.

1. *Amblyomma sculptum* - Fisiologia. 2. Proteínas.
3. Glândulas salivares. 4. Intestinos. 5. Ovários.
I. Universidade Federal de Viçosa. Departamento de Bioquímica e Biologia Molecular. Programa de Pós-graduação em Bioquímica Aplicada. II. Título.

CDD 22. ed. 595.42

ADRIANO CARLOS SOARES

**IDENTIFICAÇÃO DE LIPOCALINAS COM DOMÍNIO DE LIGAÇÃO
A HISTAMINA E SEROTONINA E VALIDAÇÃO DE GENES DE
REFERÊNCIA EM TRANSCRIPTOMAS DE GLÂNDULA SALIVAR,
INTESTINO E OVÁRIO DO CARRAPATO *Amblyomma sculptum***

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Doctor Scientiae*.

APROVADA: 26 de Junho de 2015.

Glenda Nicioli da Silva

Camilo Elber Vital

Luciano Gomes Fietto

Ronny Francisco de Souza

Cláudio Lísias Mafra de Siqueira
(Orientador)

“Ebenezer, pois até aqui me ajudou o SENHOR DEUS...”

“A maravilhosa disposição e harmonia do universo só pode ter tido origem segundo o plano de um DEUS que tudo sabe e tudo pode. Isso fica sendo minha última e mais elevada descoberta.”

Isaac Newton

“A verdadeira medida de um homem não se vê na forma como se comporta em momentos de conforto e conveniência, mas em como se mantém em tempos de controvérsia e desafio.”

Martin Luther King

À DEUS, que me colocou como mordomo do maior tesouro desse Mundo, minha Família!

A minha amada esposa NÉIA que soube suportar, com discernimento, os momentos difíceis de minha ausência como pai presente no lar.

Aos meus filhos Rodrigo, Felipe e Ana Laura que juntamente com minha esposa me deram total apoio nessa jornada em busca do conhecimento!

Aos meus pais Pedro e Luiza pela educação que me deram e o meu amigo e orientador Cláudio Mafra que aprendi a respeitar e admirar como exemplo de integridade e ética na realização da pesquisa científica!

Aos amigos do Lapem, em especial ao Rafael, Higo, Paulo, Gabriel, Fernanda e Raphael Klein e do Nubiomol Edvaldo, Carlos e Pedro Vidigal pela acolhida e conselhos relacionados ao trabalho. Essa conquista não é minha mas de todos nós!

AGRADECIMENTOS

A DEUS, pela amizade e constante direção em todos os momentos, inclusive durante as etapas de elaboração, desenvolvimento e conclusão deste trabalho. Por ter me concedido a ideia deste projeto muito antes que as parcerias necessárias para este surgissem;

Ao Dr. Cláudio Mafra, pela amizade, paciência e sabedoria. Por ter aberto as portas de seu laboratório para mim, permitindo-me grande crescimento profissional e pessoal. Pela importante orientação e coordenação técnica deste trabalho, desde a concepção do desenho experimental, desenvolvimento até a análise crítica dos dados obtidos;

Aos amigos e irmãos do Laboratório de Parasitologia de Epidemiologia Molecular (LAPEM/DBB/UFV): Rafael Mazioli, Higo Nasser, Raphael Klein, Carlos Montadon, Edvaldo, Paulo, Gabriel, Fernanda, Juliana, Cíntia e Natasha, pela amizade como também pela troca de experiência. Especialmente Higo, Rafael e Raphael Klein que me prestaram valiosíssima ajuda durante as etapas de realização desse projeto;

Ao Pedro Vidigal pelas suas contribuições no meu aprimoramento no campo da bioinformática, seus conselhos e ajuda foram valiosos;

Ao Eduardo Monteiro, secretário do departamento de bioquímica, pelo cuidado e zelo com todos os assuntos relacionados à minha vida de pós graduando em Viçosa;

Aos amigos do Departamento de Bioquímica e demais amigos de Viçosa, pela valiosa amizade;

Aos meus amigos do Centro Universitário de Caratinga em especial ao Dr. Walter Roberto Linardi e o Dr. João Batista dos Reis pelos aconselhamentos durante o mestrado e encorajamento para a realização desse doutorado;

Ao amigo Leo Mussi, pela ajuda na tradução do artigo e apoio nas tarefas realizadas;

Aos meus demais amigos e familiares que, apesar da distância, sempre me ajudaram com suas palavras de motivação e suas orações;

Enfim, agradeço a todas aquelas pessoas que, de alguma forma, estiveram do meu lado durante este grande desafio, contribuindo para realização desta importante conquista.

BIOGRAFIA

Adriano Carlos Soares, nascido em 13 de Setembro de 1972, na cidade de Governador Valadares, Minas Gerais, Brasil, filho de Pedro Celestino Soares e Luiza Carlos Soares, casado com Waldinéia Dulce Reis Soares, pai de três filhos, Rodrigo José Reis Soares, Felipe Augusto Reis Soares e Ana Laura Reis Soares.

No ano de 1992, ingressou no Bacharelado em Farmácia e Bioquímica pela Universidade Federal de Ouro Preto, obtendo o título de Bacharel em Farmácia e Bioquímica com habilitação em análises clínicas em 11 de janeiro de 1997. Durante a graduação, exerceu diversas atividades como aluno de iniciação científica e monitor nas disciplinas de Química Orgânica I, II e III, Fisiologia I e II e Imunologia Clínica. Em 10 de Outubro de 1998, casou-se e passou a atuar como empresário no ramo de análises clínicas e drogaria.

Após realizar três pós graduações *lato sensu*, ingressou em 2008 na carreira docente ministrando diversas disciplinas nos cursos de graduação em Farmácia, Medicina, Nutrição, Enfermagem, Fisioterapia, Engenharia Ambiental, Licenciatura em Ciências Biológicas, Licenciatura e Bacharelado em Química. Em Julho de 2008 iniciou suas atividades no Programa de Pós-graduação em Ciências Naturais e da Saúde no Centro Universitário de Caratinga (UNEC), com o projeto dissertação de mestrado intitulado: Qualidade de amostras de urina sob conservação para a realização de Urinálise e urocultura, defendido em Julho de 2010.

Em agosto de 2010, iniciou suas atividades no Programa de Pós graduação em Bioquímica Agrícola, atualmente Bioquímica Aplicada, da Universidade Federal de Viçosa, como aluno não vinculado e em Março de 2011 tornou-se doutorando pelo departamento de Bioquímica e Biologia Molecular da Universidade Federal de Viçosa no Laboratório de Parasitologia e Epidemiologia Molecular (LAPEM), sob a orientação do prof. Dr. Cláudio Mafra.

SUMÁRIO

LISTA DE FIGURAS	ix
LISTA DE TABELAS.....	xii
LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS.....	xiii
RESUMO.....	xv
ABSTRACT.....	xvii
REDE DE PESQUISAS EM CARRAPATOS.....	1
1. INTRODUÇÃO	3
2. JUSTIFICATIVA	5
3. HIPÓTESE	7
4. FUNDAMENTAÇÃO TEÓRICA	9
4.1 Aspectos taxonômicos do carrapato	9
4.2 Ciclo de desenvolvimento dos carrapatos vetores	9
4.3 Rickettsia como agente Etiológico da Febre maculosa	10
4.3.1 Aspectos Imunológicos das Doenças Riquetsiais	11
4.3.2 Aspectos Patológicos das Enfermidades Riquetsiais	12
4.4 Lipocalinas, proteínas ubíquas multifuncionais	13
4.4.1 Lipocalinas: Relação entre a sequência e a estrutura	14
4.4.2 Lipocalinas: potencial farmacológico e terapêutico de uma família multifuncional levando à uma pluralidade de alvos moleculares	15
4.4.2.1 Lipocalinas como alvo molecular no desenvolvimento de biofármacos com atividade anti-histamínica e anti-coagulante	16
4.5 Papel da glândula salivar, intestino e ovário de <i>Amblyomma sculptum</i> na interface microrganismo-vetor-hospedeiro	17

4.6 Genes de referência para aplicação em RT-qPCR.....	18
4.7 Bioinformática na predição de modelos proteicos tridimensionais	22
5. OBJETIVOS.....	26
5.1 OBJETIVO GERAL.....	26
5.2 OBJETIVOS ESPECÍFICOS.....	26
6. MATERIAL E MÉTODOS.....	27
6.1 Refino e alinhamento dos transcritos referentes a lipocalinas advindas da base de dados transcriptômicos do carrapato <i>A. sculptum</i>	27
6.2 Predição de domínios nos transcritos de lipocalinas de <i>A. sculptum</i>	27
6.3 Análise filogenética dos transcritos de lipocalinas preditos em transcriptomas de glândulas salivares de <i>Amblyomma sculptum</i>	28
6.4 Extração do RNA total das glândulas salivares, intestinos e ovários do carrapato <i>A. sculptum</i>	28
6.5 Síntese do cDNA para avaliação da expressão gênica por RTq-PCR.....	28
6.6 Desenho de <i>primers</i> , teste de concentração e eficiência para validação da qPCR	29
6.7 Predição estrutural <i>in silico</i> das lipocalinas com domínio SHBP provenientes dos genes selecionados para a RT-qPCR;.....	30
7. RESULTADOS E DISCUSSÃO.....	36
7.1 Análise filogenética das lipocalinas de <i>A. sculptum</i>	36
7.2 Predição <i>in silico</i> dos domínios SHBP.....	41
7.3 Refino da seleção de transcritos de lipocalinas de <i>Amblyomma sculptum</i> com domínio SHBP	44
7.4 Resultado da extração do RNA total das Glândulas salivares, intestinos e ovários do carrapato <i>A. sculptum</i>	46
7.5 Análise e validação dos <i>primers</i> por testes de concentração e eficiência para uso em RTq-PCR.....	47

7.5.1 Estudo dos genes alvo e dos transcritos selecionados para a amplificação e identificação de lipocalinas com domínio SHBP para uso em RTq-PCR.....	51
7.5.2 Predição do modelo estrutural e das estruturas secundárias das lipocalinas 4, 7 e 24 com domínio SHBP por aplicativos de bioinformática	51
7.5.3 Validação do teste de concentração e eficiência dos primers de lipocalinas e dos candidatos a genes de referência para uso em RTq-PCR ...	73
7.6 Análise da expressão gênica de lipocalinas com SHBP <i>domain</i> e genes constitutivos por quantificação relativa utilizando RTq-PCR	75
8. CONCLUSÃO	90
9. REFERÊNCIAS BIBLIOGRÁFICAS	93
ANEXOS	104

LISTA DE FIGURAS

- Figura 1.** Modelo estrutural da lipocalina demonstrando o cálice e as folhas da estrutura secundária em barril β (A–H). 15
- Figura 2.** Estratégia utilizada para a predição e avaliação dos modelos 3D das lipocalinas LIP4, LIP7 e LIP24. 31
- Figura 3.** Estratégia utilizada no aplicativo *Phyre2* para a comparação e avaliação dos modelos 3D das lipocalinas LIP4, LIP7 e LIP24 preditos pelo aplicativo *RaptorX*. 32
- Figura 4.** Estratégia utilizada para a predição da presença de hélices transmembrana nos modelos 3D das lipocalinas LIP4, LIP7 e LIP24 preditos pelo aplicativo *RaptorX* e *Phyre2*. 33
- Figura 5.** Estratégia utilizada para a avaliação da estrutura terciária dos modelos 3D das lipocalinas LIP4, LIP7 e LIP24 preditos pelo aplicativo *RaptorX* e *Phyre2*. 34
- Figura 6.** Diagrama de Ramachandran com a predição das estruturas secundárias mostrando a densidade de valores do par (ψ, ϕ) . 35
- Figura 7.** Árvore bayesiana, das 75 sequências de transcritos de lipocalinas de glândulas salivares de *A. sculptum*, resultante do alinhamento com deleção de gaps e normalização pelo aplicativo Gblocks. 37
- Figura 8.** Árvore bayesiana, das 105 sequências de lipocalinas selecionadas no *Swiss-prot* advindas do transcriptoma das glândulas salivares de *Amblyomma cajennense* 43
- Figura 9.** Alinhamento entre as sete sequências com *SHBP domain* selecionadas do *Swiss-prot* que se agruparam na figura 3 somadas ao transcrito As_6240, não depositado, proveniente do RNA-seq. de glândulas salivares de *A. sculptum sculptum* e confrontadas com as sequências de lipocalinas de *A. cajennense* depositadas no UniProt 45
- Figura 10.** Eletroforese em gel de agarose 1% de amostras de RNAt (dois *pulls* de 75 glândula salivar, intestino e ovário) provenientes de transcriptomas de glândulas salivares, intestinos e ovários do carrapato *Amblyomma sculptum* armazenados a -80°C. 47
- Figura 11.** Distribuição gráfica da participação dos resíduos na formação das estruturas secundárias compondo as α hélices, folhas β e *loop* da Lipocalina 4 preditas pelo aplicativo *RaptorX*. Em (a) modelo 3D predito pelo aplicativo *RaptorX*. 52
- Figura 12.** Distribuição gráfica da participação dos resíduos de aminoácidos da Lipocalina 4 na predição dos parâmetros de desordem e acessibilidade do solvente preditas pelo aplicativo *RaptorX*. 55
- Figura 13.** Distribuição gráfica da participação dos resíduos na formação das estruturas secundárias compondo as α hélices, folhas β e espirais da Lipocalina 7 preditas pelo aplicativo *RaptorX*. 57
- Figura 14.** Distribuição gráfica da participação dos resíduos de aminoácidos da Lipocalina 7 na predição dos parâmetros de desordem e acessibilidade do solvente preditas pelo aplicativo *RaptorX* 59
- Figura 15.** Distribuição gráfica da participação dos resíduos na formação das estruturas secundárias compondo as α hélices, folhas β e espirais da Lipocalina 24 preditas pelo aplicativo *RaptorX* 61

Figura 16. Distribuição gráfica da participação dos resíduos de aminoácidos da Lipocalina 24 na predição dos parâmetros de desordem e acessibilidade do solvente preditas pelo aplicativo <i>RaptorX</i>	64
Figura 17. Predição das estruturas secundárias da LIP4.	66
Figura 18. Predição das estruturas secundárias da LIP7.	67
Figura 19. Predição das estruturas secundárias da LIP24.	68
Figura 20. Análise de domínio transmembrana e presença de peptídeo sinal das lipocalinas 4 (a), 7(b) e 24(c) predita pelo memsat-svm do PSI-PRED.	69
Figura 21. Diagrama de ramachandram construído pelo aplicativo RAMPAGE demonstrando a densidade de valores do par (ϕ, ψ) dos aminoácidos da lipocalina 4.	70
Figura 22. Diagrama de ramachandram construído pelo aplicativo RAMPAGE demonstrando a densidade de valores do par (ϕ, ψ) dos aminoácidos da lipocalina 7.	71
Figura 23. Diagrama de ramachandram construído pelo aplicativo RAMPAGE demonstrando a densidade de valores do par (ϕ, ψ) dos aminoácidos da lipocalina 24.	72
Figura 24. Curva de dissociação (curva de <i>melt</i>) dos <i>primers</i> de lipocalinas desenhados, as concentrações ideais para cada primer encontra-se entre os colchetes [].	73
Figura 25. Curva de dissociação (curva de <i>melt</i>) dos <i>primers</i> dos genes constitutivos desenhados, as concentrações ideais para cada primer encontra-se entre os colchetes [].	74
Figura 26. Expressão relativa do log2 RQ (<i>Relative Quantity</i> , RQ) dos genes <i>lip4</i> , <i>lip7</i> e <i>lip24</i> de lipocalinas com <i>SHBP domain</i> em glândulas salivares, intestinos e ovários de fêmeas do carrapato <i>A. sculptum</i>	76
Figura 27. Expressão relativa do log2 RQ (<i>relative quantity</i> , RQ) do gene <i>lip4</i> de lipocalinas com <i>SHBP domain</i> em glândulas salivares, intestinos e ovários de fêmeas do carrapato <i>A. sculptum</i> .	77
Figura 28. Expressão relativa do log2 RQ (<i>relative quantity</i> , RQ) do gene <i>lip4</i> de lipocalinas com <i>SHBP domain</i> em ovários de fêmeas do carrapato <i>A. sculptum</i> , confrontada com novas amostras de glândula salivar, intestino e ovário não infectado	78
Figura 29. Expressão relativa do log2 RQ (<i>relative quantity</i> , RQ) do gene <i>lip7</i> de lipocalinas com <i>SHBP domain</i> em glândulas salivares, intestinos e ovários de fêmeas do carrapato <i>A. sculptum</i> .	79
Figura 30. Expressão relativa do log2 RQ (<i>Relative Quantity</i> , RQ) do gene <i>lip7</i> de lipocalinas com <i>SHBP domain</i> em ovários de fêmeas do carrapato <i>A. sculptum</i> , confrontada com novas amostras de intestino e ovário não infectado.	80
Figura 31. Expressão relativa do log2 RQ (<i>Relative Quantity</i> , RQ) do gene <i>lip24</i> de lipocalinas com <i>SHBP domain</i> em glândulas salivares, intestinos e ovários de fêmeas do carrapato <i>A. sculptum</i>	81
Figura 32. Expressão relativa do log2 RQ (<i>Relative Quantity</i> , RQ) do gene <i>lip24</i> de lipocalinas com <i>SHBP domain</i> em glândulas salivares intestinos e ovários de	82

fêmeas do carrapato *A. sculptum*, confrontada com novas amostras de intestino e ovário não infectado

- Figura 33.** Resultado da análise de *Two-way* ANOVA (a) e teste *t* (b) dos valores de Cp dos genes candidatos a genes de referência. **86**
- Figura 34.** Resultado da análise de *Two-way* ANOVA (a) e teste *t* (b) dos valores de Cp dos genes candidatos a genes de referência em tecido não infectado. **87**
- Figura S1.** Resultados da análise de real time RT-qPCR do teste de eficiência do primer *lip4* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). **110**
- Figura S2.** Resultados da análise de real time RT-qPCR do teste de eficiência do primer *lip7* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). **110**
- Figura S3.** Resultados da análise de real time RT-qPCR do teste de eficiência do primer *lip24* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *lip24*. **111**
- Figura S4.** Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk1* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk1*. **111**
- Figura S5.** Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk2* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk2*. **112**
- Figura S6.** Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk4* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk4*. **112**
- Figura S7.** Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk5* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk6*. **113**
- Figura S8.** Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk6* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk6*. **114**

LISTA DE TABELAS

Tabela 1. Relação das 121 sequências de lipocalinas advindas do transcriptoma de glândula salivar de <i>A. sculptum</i> .	36
Tabela 2. Análise funcional do clado I dos transcritos de lipocalinas de glândulas salivares de <i>Amblyomma sculptum</i> submetida ao PFAM.	38
Tabela 3. Análise funcional dos clados II, III e IV dos transcritos de lipocalinas de glândulas salivares de <i>Amblyomma sculptum</i> submetida ao PFAM.	40
Tabela 4. Análise funcional do clado V dos transcritos de lipocalinas de glândulas salivares de <i>Amblyomma sculptum</i> submetida ao PFAM.	41
Tabela 5. Característica dos <i>primers</i> para amplificação de fragmentos de lipocalinas com presença de SHBP <i>domain</i> desenhados pelo aplicativo <i>Primer express</i> 3.0 para aplicação em qPCR.	48
Tabela 6. Característica dos <i>primers</i> para amplificação de fragmentos de lipocalinas desenhados pelo aplicativo <i>Primer express</i> 3.0 para os candidatos a genes de referência para aplicação em qPCR.	49
Tabela 7. Resultados das curvas de eficiência dos genes de lipocalinas selecionados na etapa de determinação da concentração e curva de eficiência dos genes candidatos a genes de referência.	50
Tabela 8. Resultados dos parâmetros do aplicativo <i>RaptorX</i> na predição dos modelos estruturais das lipocalinas LIP4, 7 e 24.	65
Tabela 9. análise do diagrama de Ramachandram apresentando o cálculo dos ângulos diedrais do par (ϕ, ψ) para a lipocalina 4.	71
Tabela 10. análise do diagrama de Ramachandram apresentando o cálculo dos ângulos diedrais do par (ϕ, ψ) para a lipocalina 7.	72
Tabela 11. análise do diagrama de Ramachandram apresentando o cálculo dos ângulos diedrais do par (ϕ, ψ) para a lipocalina 24.	73
Tabela 12. Estatística descritiva dos 3 genes candidatos a genes de referência calculados pelo <i>Bestkeeper</i> baseados nos valores de <i>crossing point</i> (Cp).	85
Tabela S1: Sequência de aminoácidos dos transcritos de lipocalinas identificadas no RNA-seq de glândulas salivares de <i>Amblyomma sculptum</i> .	104

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

LAPEM	Laboratório de Parasitologia e Epidemiologia Molecular
FMRP-USP	Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo
UFG	Universidade Federal de Goiás
UFU	Universidade Federal de Uberlândia
UFV	Universidade Federal de Viçosa
LMVR-NIH	<i>Malaria and Vector Research, National Institute of Health</i>
GENOPROT	Genoma e Proteoma
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
DBB/UFV	Departamento de Bioquímica e Biologia Molecular – UFV
RNA-seq	Sequência de ácido ribonucléico
RT-qPCR	<i>Real time Reverse transcriptase quantitative Polimerase Chain Reaction</i>
SHBP	<i>Serotonine and histamine binding protein</i>
<i>g6pdh</i>	Gene glicose-6-fosfato desidrogenase
<i>3ohcoadh</i>	Gene 3-hidroxi-coenzima A desidrogenase
<i>gapdh</i>	Gene gliceraldeído-6-fosfato desidrogenase
BAB	<i>biogenic amin binding</i>
cDNA	Ácido desoxirribonucleico complementar
HBP	<i>Histamine binding proteins</i>
NK	<i>Natural Killer</i>
IL-12	Interleucina 12
INF γ	<i>Interferon gama</i>
Linf. T CD4+	Linfócitos T CD4+
Linf. T CD8+	Linfócitos T CD8+
CD	<i>Cluster of differentiation</i>
Kda	<i>Kilodaltons</i>
SRC	<i>Structure regions conserved</i>
NGAL	lipocalina de neutrófilo associada a gelatinase
5-HT	5-hidróxitriptamina, serotonina
qPCR	<i>Real time Polimerase chain reaction quantitative</i>
Cp	<i>Crossing point</i>
Ct	<i>Cycle threshlold</i>
Δ Cp	Delta crossing point
$\Delta\Delta$ Cp	Delta delta crossing point
PMP	<i>Protein Model Portal</i>
Expasy	<i>Expert Protein Analysis System</i>
Phyre2	<i>Protein Homology/analogy Recognition Engine V.2.0</i>
IntFOLDTS	<i>Integrated web resource for protein fold recognition</i>
PDB	Protein data bank
USeqID	<i>un-identical residues alignment sequence</i>
SeqID	<i>identical residues alignment sequence</i>
GDT	<i>global distance test</i>
uGDT	<i>un-normalized global distance test</i>
ACC	Acessibilidade ao solvente
PSI-BLAST	<i>Position-Specific Iterative Basic Local Alignment Search Tool</i>
PSSM	matriz específica de posição e pontuação
HHsearch1	<i>Homologues Hiden Search 1</i>
HMM	<i>Hiden Markov Models</i>
EMBL-EBI	<i>European Bioinformatcs Institute</i>
SIB	<i>Swiss Institute of Bioinformatic</i>
TrEMBL	<i>Translate Sequence for European Molecular Data Library</i>
PIR	<i>Protein Sequence Database</i>
PSD	Protein Sequence Databse
iProClass	<i>International Protein Class</i>
UniProt	<i>Resource Protein Universal</i>
NIH	<i>National Institute Health</i>
EMBL	<i>European Molecular Biology Laboratory</i>
BHF	<i>Britsh Heart Fundation</i>
PDUK	<i>Parkinson's Disease United Kingdom</i>

NSF	<i>National Science Foundation</i>
UniProtKB	<i>UniProt Knowledge Base</i>
UniRef	<i>UniProt Reference Clusters</i>
UniParc	<i>UniProt Archive</i>
MAFFT	<i>Multiple alignment program for amino acid or nucleotide sequence</i>
NuBiomol/UFV	Nucleo de biomoléculas da Universidade Federal de Viçosa
IDT	<i>Integrated DNA Technologies</i>
NTC	<i>Negative template control</i>
uL	Microlitro
uM	Micromolar
SVM	<i>Support vector Machine</i>
RMSD	Resultado da energia média quadrática
RAMPAGE	<i>Ramachandran Plot Analysis</i>
ϕ, ψ	Par de ângulos diedrais phi, psi
GBLOCKS	Blocos de genes
RT	<i>Reverse transcriptase</i>
RaHBP	<i>Ripicephalus apendiculatus histamin binding protein</i>
PFAM	<i>The Protein Family Database</i>
<i>hk</i>	<i>Housekeeping genes</i>
PSI-PRED	<i>Position-Specific Iterative prediction</i>
DISO-PRED	<i>Disorder prediction</i>
Log2RQ	Log2 da <i>Relative Quantity</i>
INT	intestino
OVA	ovário
GS	Glândula salivar
GST	glutathiona s-transferase

RESUMO

SOARES, Adriano Carlos, D.Sc., Universidade Federal de Viçosa, Junho de 2015. **IDENTIFICAÇÃO DE LIPOCALINAS COM DOMÍNIO DE LIGAÇÃO A HISTAMINA E SEROTONINA E VALIDAÇÃO DE GENES DE REFERÊNCIA EM TRANSCRIPTOMAS DE GLÂNDULA SALIVAR, INTESTINO E OVÁRIO DO CARRAPATO *Amblyomma sculptum***. Orientador: Cláudio Lisias Maфра de Siqueira.

As lipocalinas constituem uma família de pequenas proteínas extracelulares com cerca de 200 resíduos de aminoácidos e massa molecular de 20 Kda possuem forte divergência em suas sequências de aminoácidos mas mantêm a sua estrutura tridimensional estritamente conservada. Sua diversidade funcional e reconhecimento molecular, contribuem para a utilização dessas proteínas como base para o desenvolvimento de alguns biofármacos. Os genes constitutivos participam de diversas funções do metabolismo celular e são utilizados como controle endógeno sendo fundamentais para validação dos resultados no estudo da expressão utilizando *real time* RTq-PCR. Em relação a esses genes, ainda não foram publicados trabalhos que identifiquem e validem genes constitutivos como genes de referência para a espécie *Amblyomma sculptum* (Complexo *Amblyomma cajennense*). Atualmente, transcriptomas de órgãos e tecidos de carrapatos do gênero *Amblyomma* têm sido publicados identificando uma variedade de transcritos de lipocalinas com domínio de ligação a histamina e serotonina (SHBP). Neste trabalho, realizamos por predição *in silico* a identificação de nove lipocalinas com domínio de ligação a serotonina e histamina estabelecendo possíveis relações filogenéticas funcionais de transcritos de lipocalinas identificadas no transcriptoma de glândula salivar do *A. sculptum*. Realizamos o desenho e a validação de 3 *primers* de genes de lipocalinas identificadas, *lip4*, *lip7* e *lip24*. Foram analisados e validados 3 *primers* de genes constitutivos (*g6pdh*, *3ohcoadh* e *gapdh*) candidatos potenciais a genes de referência em glândulas salivares, intestinos e ovários do carrapato *A. sculptum* em situação de não infecção/infecção com *R. amblyommii*. Foram validados 3 modelos estruturais correspondentes às proteínas lipocalina 4, lipocalina 7 e lipocalina 24. E ainda, realizou-se a análise da expressão diferencial de lipocalinas e potenciais genes de referência em glândula salivar, intestino e ovário de *A. sculptum*. Os dados obtidos na análise de expressão gênica por *real time* RT-qPCR suportam a hipótese de que os três genes de lipocalinas interferem na interação microrganismo-vetor sendo necessários trabalhos adicionais para verificação da validade desta hipótese, uma vez

que dentre as 121 lipocalinas preditas no transcriptoma de glândulas salivares, somente nove apresentam domínios SHBP sugerindo um papel especial na função destas nos tecidos estudados.

ABSTRACT

SOARES, Adriano Carlos, D.Sc., Universidade Federal de Viçosa, June, 2015. **IDENTIFICATION OF SEROTONINE HISTAMINE BINDING PROTEIN LIPOCALINS AND VALIDATION OF REFERENCE GENES IN SALIVARY GLAND, GUT AND OVARIAN *Amblyomma sculptum* TICK TRANSCRIPTOMES.** Adviser: Claudio Lisias Mafra de Siqueira.

The lipocalin's family compound a small extracellular protein of about 200 amino acid residues and a molecular mass of 20 kDa present strong divergence in their amino acid sequences but retain its strictly conserved three dimensional structure. Its functional diversity and molecular recognition, contributing to the use of these proteins as a basis for the development of some biopharmaceuticals. The reference genes involved in various cellular metabolism functions and are used, as endogenous control is necessary to validate the results in the study of the expression for real RTQ-PCR analysis. For these genes have not yet been published works to identify and validate constitutive genes as the reference genes for *Amblyomma sculptum* (*Amblyomma cajennense* complex) species. Currently, transcriptomes of organs and tissues of the genus *Amblyomma* ticks have been published identifying a variety of lipocalins transcripts with binding domain histamine and serotonin (SHBP). We carried out by prediction *in silico* identification of nine lipocalins with serotonin and histamine binding domain setting possible functional phylogenetic relationships of lipocalins transcripts identified in salivary gland transcriptome of *A. sculptum*. We carry out the design and validation of three primers identified lipocalin genes, *lip4*, *lip7* and *lip24*. They were analyzed and validated three primers of constitutive genes (*g6pdh*, *3ohcoadh* and *gapdh*) potential candidate reference genes in the salivary glands, ovaries and intestines *A. sculptum* tick for a situation of no infection / infection with *R. amblyommii*. Were validated three corresponding structural models to lipocalins 4, 7 and 24. Yet, there was the analysis of differential expression of lipocalins and potential reference genes in the salivary gland, gut and ovarian *A. sculptum* tick. The results of analysis in experiment of gene expression by real time RT-qPCR support the hypothesis that the three lipocalins genes interfere with the microorganism-vector interactions. Finally it is necessary a further work to verify the validity of this hypothesis, since among the 121 lipocalins predicted in transcriptome salivary glands, only nine have SHBP domains suggesting a special role in the function of these tissues studied.

REDE DE PESQUISAS EM CARRAPATOS

A REDE DE PESQUISAS EM CARRAPATOS consiste em um projeto de pesquisas em genômica, transcriptômica e proteômica das principais espécies de carrapatos de importância médico-veterinária do Brasil. Este projeto atualmente está sendo executado por uma rede composta por pesquisadores de quatro Universidades brasileiras (Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, FMRP-USP; Universidade Federal de Goiás, UFG; Universidade Federal de Uberlândia, UFU e Universidade Federal de Viçosa, UFV) em convênio com o *Laboratory of Malaria and Vector Research, National Institute of Health (LMVR-NIH)* nos EUA, um centro de referência mundial em estudos em genômica funcional de artrópodes de importância médico-veterinária. Contando em 2009 com suporte financeiro do Programa GENOPROT do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (Edital MCT/CNPq/CT-AGRO/CT-BIOTEC No 42/20, A REDE DE PESQUISAS EM CARRAPATOS iniciou atividades de parceria entre grupos correlatos no Brasil. Parte destes recursos foi destinada ao grupo do Departamento de Bioquímica e Biologia Molecular – UFV (DBB/UFV) para sequenciamento em massa de transcriptomas de glândulas salivares de fêmeas do carrapato *Amblyomma sculptum* (Complexo *Amblyomma cajennense*), nas condições de infecção e não infecção com *Rickettsia amblyommi*. Esta tese de doutorado surgiu em decorrência dessa rede, apresentando um estudo do perfil de expressão de genes de lipocalinas com domínio de ligação a histamina e serotonina, bem como, a validação de *primers* de lipocalinas referentes a essa seleção e de potenciais candidatos a genes de referência advindos do RNA-seq de glândula salivar do carrapato *Amblyomma sculptum* (complexo *A. cajennense*), considerado o principal vetor da Febre Maculosa Brasileira, sob condições de infecção e não infecção com *Rickettsia amblyommii*. Adicionalmente, utilizou-se o RNA total de intestino e ovário (Proc. N°: APQ-01552-11), para validação dos *primers* de lipocalinas e genes constitutivos candidatos a genes de referência.

O estudo dos genes de lipocalinas com domínio de ligação a histamina e serotonina na glândula salivar do carrapato *A. sculptum* é uma estratégia que busca compreender melhor a genômica e a transcriptômica do processo de hematofagia deste carrapato, uma vez que sintetiza uma saliva rica em biomoléculas (anti-histamínicos, vasodilatadores, anticoagulantes e imunomoduladores) essenciais para o sucesso do repasto sanguíneo, além de uma série de outros compostos ainda não caracterizados que intermedeiam este processo.

O conhecimento mais específico da expressão dos genes das lipocalinas selecionadas, bem como dos genes constitutivos nas glândulas salivares estudadas, constitui um passo fundamental para a seleção de genes candidatos à prospecção de biomoléculas com atividade anti-histamínica através da obtenção de produtos recombinantes.

Os genes de lipocalinas selecionados tornam-se portanto, moléculas com potencial aplicação farmacológica e biotecnológica, com análise da expressão destes e validação de *primers* para estudos posteriores contribuindo para uma maior compreensão da interação metabólica das glândulas salivares, intestinos e ovários desse parasito, esclarecendo algumas lacunas sobre a manutenção, proliferação e dispersão deste agente para diversos hospedeiros vertebrados, inclusive o homem. Detectando novas moléculas que sirvam de alvo para o desenvolvimento de produtos biotecnológicos que possam ser aplicados em estratégias de controle do *Amblyomma sculptum*.

Dessa forma, este projeto teve como objetivo analisar e validar a expressão de transcritos relacionados às lipocalinas com domínio de ligação à Serotonina e Histamina e de possíveis genes constitutivos candidatos a genes de referência em transcriptomas de glândula salivar, intestino e ovário do carrapato *A. sculptum*.

1. INTRODUÇÃO

O avanço das ferramentas biotecnológicas e de bioinformática e a aplicação das mesmas na busca de biofármacos determinaram a implantação das mais variadas estratégias para identificação de moléculas candidatas à aplicação biotecnológica. Essas estratégias partem de pontos fundamentais, sendo os principais: abordagens computacionais associadas à biologia de sistemas; abordagens de novas tecnologias de medição e análise de bancos de dados eletrônicos e de repositórios na tentativa de mineração de possíveis moléculas candidatas (CSERMELY et al., 2013).

O bloqueio do processo inflamatório depende de diversas moléculas com atividades anti-histamínica, dentre as quais um grupo de proteínas conhecidas como lipocalinas (PAESEN et al., 1999). Estas são proteínas que se encontram expressas em diversas situações fisiológicas e participam das mais variadas funções, estando presente tanto em animais quanto plantas e diversas espécies de artrópodes hematófagos, dentre estas, o carrapato *Amblyomma sculptum* (complexo *A. cajennense*), pertencente à família Ixodidae. As lipocalinas têm sido descritas em estudos relacionados à busca pela descoberta de novos compostos bioativos com potencial biotecnológico que vão desde marcadores de doenças até biofármacos (SKERRA, 2000; FLOWER, 2000; MARCHAL et al. 2012).

Alguns estudos têm relatado a presença de uma gama de lipocalinas presentes em carrapatos argasídeos e ixodídeos. As lipocalinas relatadas nesses carrapatos têm o chamado padrão de ligação de amina biogênica, *biogenic amin binding* (BAB) (BEAUFAYS et al., 2008). As lipocalinas que apresentam este domínio se ligam à histamina e serotonina inativando-as nos locais de alimentação do carrapato, podendo ser consideradas como possíveis alvos moleculares para o desenvolvimento de biofármacos com ação anti-histamínica, uma vez que essa ação inibe o desencadeamento da resposta imune do hospedeiro.

De acordo com Chmelar et al. (2008), estudos da saliva de carrapato já haviam demonstrado a presença de várias moléculas inibidoras de protease, bem como apresentando atividade anticoagulante (NARASIMHAM et al., 2002; FRANCISCHETTI, MATHER & RIBEIRO, 2003), anti-inflamatória e atividade imunossupressora (LEBOULLE et al., 2002; KOTSYFAKIS et al., 2006), constituindo-se, portanto, em moléculas com potencial farmacológico. Como responsáveis por tais atividades foram encontradas prostaglandinas (RIBEIRO et al., 1985; RIBEIRO, MAKOUL &

ROBINSON, 1988), apirases (MANS et al., 1998) e as proteínas de ligação pertencentes às lipocalinas (PAESEN et al., 1999; MANS, 2005).

Neste trabalho, buscou-se realizar por predição *in silico* a identificação de transcritos de lipocalinas em transcriptoma de glândulas salivares do carrapato *A. sculptum* com domínio de ligação a serotonina e histamina, desenhar e validar os *primers* para o estudo da expressão de lipocalinas e genes constitutivos candidatos a genes de referência identificados pelas análises de bioinformática, bem como, analisar a expressão dos transcritos identificados e dos genes candidatos em glândulas salivares, intestinos e ovários de fêmeas do carrapato *Amblyomma sculptum*.

2. JUSTIFICATIVA

Em trabalhos recentes foram obtidas quatro bibliotecas de cDNA da glândula salivar do carrapato *A. sculptum* (MOREIRA, 2012), e oito bibliotecas de cDNA de intestino e ovário de carrapato *A. sculptum* em situação de infecção e não infecção por *R. amblyommii*. Dentre os achados destes trabalhos, foi demonstrado que durante o repasto sanguíneo do *A. sculptum*, as glândulas salivares apresentam superexpressão de moléculas com objetivo de subjugar as barreiras de defesa imunológica do hospedeiro, constituindo assim um arsenal de moléculas bioativas a serem utilizadas pelo vetor durante sua alimentação. No transcriptoma de glândulas salivares, 213 clusters (9,5%) foram anotados como codificadores de lipocalinas.

Conforme discutido por Kovar (2004), os carrapatos precisam lançar mão de seu vasto repertório de biomoléculas, como peptídeos anti-inflamatórios, para driblar os mecanismos de resposta imune do hospedeiro vertebrado. Neste contexto, algumas lipocalinas têm sido caracterizadas como proteínas ligantes de histamina, apresentando ação anti-inflamatória (PAESEN et al., 2000), apresentando como característica funcional a capacidade de atuarem como ligantes de uma grande variedade de biomoléculas como nucleotídeos, aminas bioativas (histamina e serotonina), agentes anti-coagulantes, tromboxanas, leucotrienos, inibidores do sistema do complemento e imunoglobulinas (SCHLEHUBER & SKERRA, 2005; FRANCISCHETTI, MATHER, RIBEIRO, 2003; MANS & RIBEIRO et al, 2008a; 2008b).

Apesar da existência de alguns estudos sobre as lipocalinas onde podemos destacar estudos da diversidade funcional (SCHLEHUBER & SKERRA, 2005; FRANCISCHETTI, MATHER, RIBEIRO, 2003; SANGAMNATDEJ et al., 2002; MANS & RIBEIRO et al., 2008a; 2008b), não há até o momento, pesquisas voltadas à definição das funções individuais exercidas pelas proteínas advindas do carrapato *A. sculptum*, muito menos estudos da atividade anti-histamínica exercida pelas mesmas.

Esse fato justifica a importância desse trabalho na busca de dados que suportem o desenvolvimento futuro de biofármacos anti-histamínicos, valendo ser ressaltado poucos relatos na literatura da otimização e validação de *primers* para amplificação de gene de referência nessa espécie, bem como a identificação de transcritos de lipocalinas em glândulas salivares, intestino e ovário do carrapato *A. sculptum*.

Diante da política atual das agências de fomento valorizando a pesquisa aplicada visando o desenvolvimento de novos produtos biotecnológicos em saúde, torna-se um objetivo a ser alcançado, o estudo *in silico* e a análise da expressão de transcritos relacionados às lipocalinas com domínio de ligação a serotonina e histamina

e de possíveis candidatos a genes de referência do carrapato *A. sculptum*. Nosso trabalho contribui para futuros estudos de aplicação destas moléculas em busca de compostos que atuem como anti-histamínicos.

3. HIPÓTESE

Durante o seu ciclo de vida, os carrapatos precisam lançar mão de seu vasto repertório de biomoléculas para escapar dos mecanismos de resposta imune do hospedeiro vertebrado. Têm-se portanto que devido à necessidade de imunossupressão do hospedeiro durante o repasto sanguíneo, o *A. sculptum*, assim como ocorre em outras espécies de carrapato, libera através da secreção de suas glândulas salivares um arsenal de substâncias com ação farmacológica, dentre as quais encontram-se as proteínas da família das lipocalinas com atividade anti-histamínica (HBPs).

Até o presente trabalho, existem apenas 3 artigos publicados abordando transcriptoma de glândula salivar do carrapato *Amblyomma cajennense* (sinonímia do *A. sculptum*) (BATISTA et al., 2008; ANATRIELO et al., 2009 e GARCIA et al., 2014), havendo predições de sequências proteicas de lipocalinas em todos eles, com depósito de sequências destas disponíveis no *GenBank*, *GenPept*, *Protein Data Bank* e *Swiss-Prot*.

Em carrapatos, as lipocalinas foram caracterizadas como proteínas ligantes de histamina, exercendo ação anti-histamínica durante o repasto sanguíneo contribuindo na imunossupressão do hospedeiro infestado. Haveria diferença no nível de expressão dessas proteínas na glândula salivar, intestino e ovário? Essas proteínas apresentam-se diferentemente expressas nesses órgãos? E quando o carrapato é infectado com *Rickettsia amblyommii*, o nível de expressão será alterado no contexto de órgão/tecido infectado e não infectado?

Quanto ao intestino o mesmo é responsável pela digestão do sangue ingerido, porém devido à complexidade sanguínea, são necessárias diversas moléculas bioativas para o sucesso da digestão, sendo necessária ações multifuncionais das proteínas presentes nesse órgão, a presença de lipocalinas nesse pode indicar papel multifuncional dessas proteínas e seu estudo ampliará nossa compreensão dos processos metabólicos que auxiliem na sobrevivência do artrópode.

No ovário, a presença de proteínas ovarianas em situação de infecção tem sido relacionada a diminuição da reprodução do carrapato, essa pode ser uma abordagem futura para o desenvolvimento de medidas de controle (HEEKIN et al., 2013b).

Dentre os genes validados para o gênero *Amblyomma* o *gapdh* é o mais utilizado, porém seria este o melhor controle endógeno em todos os órgãos/tecidos de carrapato? A interpretação da análise da expressão gênica em carrapato *Amblyomma*

sculptum pode não condizer com a realidade pois até o momento nenhum gene de referência foi validado para essa espécie.

4. FUNDAMENTAÇÃO TEÓRICA

4.1 Aspectos taxonômicos do carrapato

Os carrapatos são artrópodes hematófagos obrigatórios do subfiló Chelicerata, da classe Arachnida, subclasse Ixodida, da ordem Acari. Em nível mundial, temos cerca de 893 espécies de carrapatos, divididas em três famílias: Ixodidae (701 espécies), Argasidae (191 espécies) e Nuttalliellidae (uma espécie) (KEIRANS, 1992; FONSECA & MARTINS, 2007; GUGLIELMONE, 2010; PFAFFLE et al., 2013). De acordo com Aragão & Fonseca (1961) e atualizado por Dantas-Torres (2009), existem no Brasil 61 espécies divididas em nove gêneros (*Argas*, *Antricola*, *Carios*, *Ornithodoros*, *Amblyomma*, *Dermacentor*, *Haemaphysalis*, *Ixodes* e *Rhipicephalus*), sendo cinco pertencentes a família Ixodidae e quatro a família Argasidae (DANTAS-TORRES, 2009).

Os carrapatos pertencentes a família Argasidae realizam parasitismo em aves e morcegos. Quanto aos carrapatos da família Ixodidae, onde se insere o gênero *Amblyomma*, encontram-se espécies parasitas de diversos animais, dentre eles bovinos, equinos, cães, capivaras e antas, além do homem como hospedeiro acidental (LABRUNA et al., 2004; FONSECA & MARTINS, 2007).

O carrapato *A. sculptum* é uma das seis espécies que compõe o complexo *Amblyomma cajennense* (*A. cajennense sensu stricto*, *A. interandium*, *A. mixtum*, *A. patinoi*, *A. tonelliae* e *A. sculptum*) (NAVA et al., 2014). No Brasil esta espécie é o principal vetor para a transmissão *Rickettsia rickettsii* agente causadora da febre maculosa (FMB) (DIAS & MARTINS, 1939) e pode ser encontrado desde as áreas amazônicas, em todos os estados da região sudeste, na região sul, no sul estado do Paraná, na região centro-oeste no Mato grosso, Mato Grosso do Sul e na região central de Goiás, no nordeste é encontrado principalmente no norte do Piauí e Pernambuco (BEATI et al., 2013; NAVA et al., 2014) sendo que os trabalhos publicados pelo complexo *A. cajennense* com amostras brasileiras tornaram-se sinonímia do *Amblyomma sculptum*.

4.2 Ciclo de desenvolvimento dos carrapatos vetores

Os carrapatos da família Ixodidae apresentam quatro estádios evolutivos: ovo, larva, ninfa e adulto, necessitando nos estágios ativos parasitar um hospedeiro vertebrado para prosseguir com o desenvolvimento de seu ciclo de vida. De característica heteroxena, o gênero *Amblyomma* necessita três hospedeiros para a conclusão de seu ciclo, realizando as mudas entre os estádios de larva, ninfa e adulto

fora do hospedeiro. Neste parasitismo, nas fases de larva e ninfa, estes ácaros são menos específicos ou exigentes quanto ao hospedeiro, podendo vir a se fixar e alimentar em diferentes espécies animais, o que comumente não ocorre para o estágio de adulto, cuja especificidade está restrita a algumas espécies de vertebrados, sendo os humanos atingidos por casos de parasitismo acidental (DIAMANT & STRICKLAND, 1965; FONSECA & MARTINS, 2007).

Depois de fecundada e ingurgitada, a fêmea do *Amblyomma* desprende-se do hospedeiro, caindo no solo, onde iniciará a oviposição após 3-5 dias, com uma produção de cerca de 5.000-10.000 ovos, em um período de aproximadamente 11-13 dias (GUGLIELMONE et al., 2006). Após o período de incubação destes ovos, aproximadamente 30 dias, ocorrerá a eclosão dos ovos liberando as larvas, as quais permanecem na vegetação até o encontro do hospedeiro onde realizam o repasto de linfa, sangue ou tecidos digeridos, vindo a desprender-se deste primeiro hospedeiro após três a seis dias de fixação, caindo ao solo para sofrer ecdise. Após 18 a 26 dias, as larvas transformam-se em ninfas, indo em busca do (segundo) hospedeiro, onde permanecerão por cinco a sete dias fixadas e alimentando-se de sangue. Após este período, as ninfas desprendem-se e caem no solo, onde sofrerão nova ecdise por 23 a 25 dias, para então se transformarem em indivíduo adulto, que após sete dias sairão em busca de novos hospedeiros.

Uma vez neste hospedeiro vertebrado, machos e fêmeas acasalam-se e realizam o repasto sanguíneo. A fêmea fertilizada e ingurgitada se desprende da pele do hospedeiro, retornando ao solo para realização da postura, iniciando assim um novo ciclo de vida (GUGLIEMONE et al., 2006; FONSECA & MARTINS, 2007).

4.3 Rickettsia como agente Etiológico da Febre maculosa

As riquetsias são bactérias parasitas intracelulares obrigatórias, Gram (-), da Ordem *Rickettsiales*, Família *Rickettsiaceae*, causando doenças em diversos hospedeiros invertebrados e vertebrados, inclusive no homem (WEISS & STRAUSS, 1991; CARDOSO, 2004). Descrita pela primeira vez como causa da Febre Maculosa das Montanhas Rochosas (*Rocky Mountain Spotted Fever* – RMSF) nos Estados Unidos da América (FMMR) em 1906, pelo Dr. Howard Taylor Ricketts, esta enfermidade teve elucidado seu agente etiológico, a bactéria *Rickettsia rickettsii*, e a transmissão desta por carrapatos (BACELLAR et al., 1996; CARDOSO et al., 2004). As riquetsias são classificadas dentro de quatro grupos baseados em suas características biológicas, genéticas e antigênicas: Grupo da Febre Maculosa, Grupo do Tifo, Grupo de Transição

e Grupo Ancestral (GILLESPIE et al., 2007). A *R. amblyommii* é uma α proteobactéria incompleta pertencente ao grupo da febre maculosa, a mesma tem sido identificada em diversas espécies de carrapato do gênero *Amblyomma* (*A. americanum*, *A. cajennense* e *A. maculatum*). Estudos apontam que a mesma infecta todas as fases de vida do vetor quer alimentado ou não alimentado, quer ingurgitado ou não ingurgitado e constitui um microrganismo para avaliação experimental dessa interface microrganismo hospedeiro (SAYLER et al., 2014).

4.3.1 Aspectos Imunológicos das Doenças Riquetsiais

As doenças riquetsiais humanas compreendem uma variedade de manifestações clínicas causadas por α -proteobactérias do gênero *Rickettsia*, *Orientia*, *Ehrlichia* e *Anaplasma*. Organismos intracelulares obrigatórios, as riquetsias invadem as células endoteliais de revestimento de pequenos e grandes vasos sanguíneos, atingindo monócitos, macrófagos e hepatócitos, provocando a destruição das células invadidas e a disseminação da infecção para todo o endotélio da árvore vascular. Essa capacidade de invadir diferentes células participantes do endotélio vascular, bem como de células envolvidas na resposta imune do hospedeiro vertebrado, leva a um conceito de heterogeneidade de células endoteliais vasculares, demonstrando uma necessidade do estudo das bases biológicas relacionadas a afinidade do microrganismo às mesmas (MANSUETO et. al, 2012).

Seus carrapatos vetores possuem em sua saliva substâncias capazes de evadir o sistema imune do hospedeiro vertebrado através de um mecanismo de modulação, permitindo que os mesmos se alimentem por vários dias ou semanas, inoculando no hospedeiro componentes antihemostáticos, dentre outros compostos, o que vem a impedir ou suprimir a resposta imune local, aumentando desta maneira o potencial de transmissão do agente patogênico para o hospedeiro vertebrado (KOVAR, 2004; KOTSYFAKIS et al., 2006; WALKER & ISMAIL, 2008). Dentre várias atividades, a saliva do carrapato diminui a proliferação de células T, a atividade de macrófagos e células *Natural Killer* (NK) e a produção de citocinas (IL-12, INF γ) (GILLESPIE et al., 2001; WALKER & ISMAIL, 2008), inibindo as funções dos neutrófilos e interferindo no sistema do complemento (KOTSYFAKIS et al., 2006; VALENZUELA, 2004; WALKER & ISMAIL, 2008).

O aspecto imunológico que envolve a doença é denominado de “vasculite riquetsial” e se inicia com o dano ao endotélio, levando a uma disfunção e ativação, seguida por uma resposta de fase aguda com alterações no sistema de coagulação e

na secreção de citocinas, o que provoca uma desregulação do sistema imune transiente, apresentando uma diminuição de linfócitos T CD4+ periféricos e infiltração de linfócitos T CD4+, T CD8+, linfócitos B e macrófagos (FENG et. al., 1997; RYDKINA et. al., 2005). Essa cadeia de eventos se inicia com a indução da ativação de fatores de transcrição da família NF-κB, responsável pela regulação da expressão de citocinas, quimiocinas, moléculas de adesão e genes de resposta iniciais. Estes fatores de transcrição permanecem no citoplasma ligado ao inibidor de IκB, que mediante a fosforilação, volta ao proteossoma para a degradação, resultando na exposição dos sinais de translocação nuclear de NF-κB. A *R. rickettsii* ativa diretamente NF-κB (VALBUENA et al., 2002).

Biópsias de tecido cerebral humano de pacientes acometidos de Febre maculosa das montanhas rochosas (FMMR) revelam um infiltrado rico em linfócitos T CD4+ e TCD8+ circundados de monócitos/macrófagos nas células endoteliais infectadas por riquetsias (VALBUENA et. al., 2002). Por ser uma bactéria intracelular obrigatória, acredita-se que a riquetsia consiga se evadir da resposta imune humoral.

Estudos demonstram que estes mecanismos envolvem a mediação de anticorpos anti-riquetsiais, revelando a ocorrência de um aumento da adesão riquetsial em células-alvo e opsonização destas, bem como, inibição ou morte de células infectadas. Isto aparentemente demonstra um efeito protetor da porção Fc-dependente, o que não é verificado na ligação com fragmentos da porção Fab (FENG et al., 1994; VALBUENA et al., 2002). Aspectos regulatórios da resposta imune riquetsial revelam que a IL-10, citocina anti-inflamatória, possui um importante papel no controle de ativação do sistema imune frente a infecções riquetsiais (dados suportados por experimentos com camundongos infectados por *R. conorii* C3H/HeN, nas fases iniciais, desenvolvimento e convalescência) e que a ativação do mesmo consiste em dano ao hospedeiro. Elevados níveis séricos de IL-10, correlacionados com baixos níveis de IL-12, sugerem a ocorrência de um período de imunossupressão transiente entre o 10º e 15º dias pós-infecção (VITALE et. al., 2001; VALBUENA et al., 2002).

4.3.2 Aspectos Patológicos das Enfermidades Riquetsiais

Na análise patológica das doenças riquetsiais, verifica-se a ocorrência de aumento da permeabilidade vascular em alguns órgãos devido a presença de numerosas células endoteliais da microcirculação infectadas pelo patógeno, com o extravasamento do líquido a partir da corrente sanguínea e acumulação nos tecidos circundantes, provocando edema e conseqüentemente hipovolemia do paciente

(HARREL & AIKAWA, 1949; DAVISON et al., 1990; WALKER et al., 2003). Esta hipovolemia quando localizada no pulmão ou cérebro, leva o paciente a risco de morte, visto esses órgãos dependerem exclusivamente de vasos linfáticos para a remoção do fluido intersticial. No cérebro, por exemplo, ocorre aumento da pressão intracraniana e diminuição do suprimento sanguíneo, comprometendo as funções cerebrais. No pulmão, ocorre redução de sua capacidade devido ao comprometimento dos espaços alveolares e da troca gás-líquido, o que leva a hipoxemia.

E ainda, a hipovolemia desencadeia uma diminuição da perfusão dos rins (WALKER & MATTERN, 1980; CONLON et al., 1996; WALKER et al., 2003). Lesões multifocais apresentadas no sistema nervoso central causam a perda da função neurológica de acordo com a área envolvida com a lesão (perda de controle motor devido a efeito extra-piramidal ocorrida na região do estriado, levando a acatsia, discenesia etc, por exemplo). Têm-se que as lesões focais não levam a incapacidade dos órgãos, mas à uma redução de sua função. Por exemplo, no fígado ocorre morte focal de hepatócitos, não se instalando porém a insuficiência hepática (RANDALL & WALKER, 1984; WALKER et al., 2003).

4.4 Lipocalinas, proteínas ubíquas multifuncionais

As lipocalinas constituem uma família de pequenas proteínas extracelulares apresentando cerca de 200 resíduos de aminoácidos e massa molecular de 20 Kda (FLOWER, 1993; GANFORNINA, 2000), ubíquas na natureza, que apresentam forte divergência em suas sequências de aminoácidos mantendo a sua estrutura tridimensional estritamente conservada (SKERRA, 2000). Essas proteínas se ligam a moléculas lipofílicas presentes em plantas, animais e bactérias. As lipocalinas são divididas em duas subfamílias de acordo com a presença ou não e o número de *motifs*, denominados de *Structurally Regions Conserved* (SRC's).

Assim, de acordo com a presença destes SRCs, as lipocalinas são subdivididas em lipocalinas do núcleo (*Kernel lipocalins*), contendo três SRC's conservados (SCR1, SRC2 e SRC3) e lipocalinas divergentes (*outliers lipocalins*), as quais podem conter duas das três SRC's (SRC1 associado a SCR2 ou SCR3) (FLOWER et al, 2000). As lipocalinas presentes em artrópodes hematófagos pertencem às *Outliers lipocalins* (MANS & RIBEIRO, 2008a).

Quanto à estrutura dessas proteínas, as mesmas compõem-se de oito folhas antiparalelas formando um barril- β fechado em uma das extremidades pela porção N-terminal e na outra pela porção C-terminal (uma α -hélice), possui ainda próximo à

porção N-terminal uma pequena hélice denominada de hélice 3¹⁰ (SKERRA, 2000). O nome lipocalina faz referência a capacidade dessa família de proteínas de abrigar em sua estrutura terciária (em forma de cálice) pequenas moléculas, sobretudo de natureza lipofílica, daí os termos em grego “kalyx” e “lipos”. Essas proteínas possuem pesos moleculares variados, com alguns autores apontando para uma superfamília de proteínas, as calicinas. Essa inclui além das lipocalinas, proteínas de ligação a ácidos graxos, triabinas, avidinas/estreptovidinas e uma classe de inibidores de metaloproteases. Os representantes dessa família possuem cadeias β antiparalelas formando uma estrutura de β -barril em uma conformação semelhante a um cálice. A triabina é a única exceção em que a disposição antiparalela promove uma troca na cadeia β , em relação ao número de cadeias para a formação do β -barril as proteínas de ligação a ácidos graxos possuem 2 cadeias β adicionais (têm-se portanto 10 cadeias em vez de 8) para a formação da estrutura em β -barril quando comparadas as outras calicinas (FLOWER, 1996; BESTE et al., 1999; GRZYB, et al., 2006; WU et al., 2012).

4.4.1 Lipocalinas: Relação entre a sequência e a estrutura

Estruturalmente, o cálice das lipocalinas é formado cadeias β constituindo oito folhas beta antiparalelas, cuja topologia é classificada de A-H, com variações em relação a um grupo específico de lipocalinas denominadas de lipocalinas divergentes (*outliers* lipocalinas) (FLOWER et al., 2000; SKERRA et al., 2000).

O cálice possui uma extremidade aberta e outra fechada, o que é determinante para a presença de diferentes conformações estruturais e funções (FLOWER et al., 2000). As cadeias β formam as folhas β que possuem conexão transversal por ligações de hidrogênio, sendo a topologia H responsável pela estabilização do domínio.

Estruturalmente as lipocalinas possuem um *loop* especial entre as cadeias β que formam as folhas β A e B, denominado como *loop* Ômega, sendo o mesmo mais largo e flexível definindo a abertura e fechamento do cálice. As demais cadeias β , que determinam a formação das folhas β antiparalelas de C a H, são estabilizadas por *loops* entre as folhas formadas ligadas à α hélice próximo à extremidade C-terminal. Todos esses *loops* se localizam na extremidade fechada do cálice. Na extremidade aberta encontra-se próximo a porção N-terminal o Ω -loop entre as cadeias β que formam as folhas β antiparalelas A e B, C e D, E e F, e G e H, com as folhas se encontrando com as folhas presentes na porção que leva a extremidade aberta do cálice. Somado a essa estrutura, as lipocalinas possuem duas α -hélices ancoradas à

estrutura de barril β formado pelas folhas β , sendo uma α -hélice mais larga e a outra menor, possuindo três voltas denominadas de hélice 3¹⁰.

Sua estrutura terciária é estabilizada através da presença de uma, duas ou três pontes dissulfeto. Essa conformação forma uma cavidade dentro do barril β que é revestido por um grupo de resíduos de aminoácidos hidrofóbicos. O interior dessa cavidade apresenta alteração dimensional, divergindo ao longo da evolução das lipocalinas (GRZYB et al., 2006). A Figura 1 apresenta o modelo estrutural definido da β -lactoglobulina.

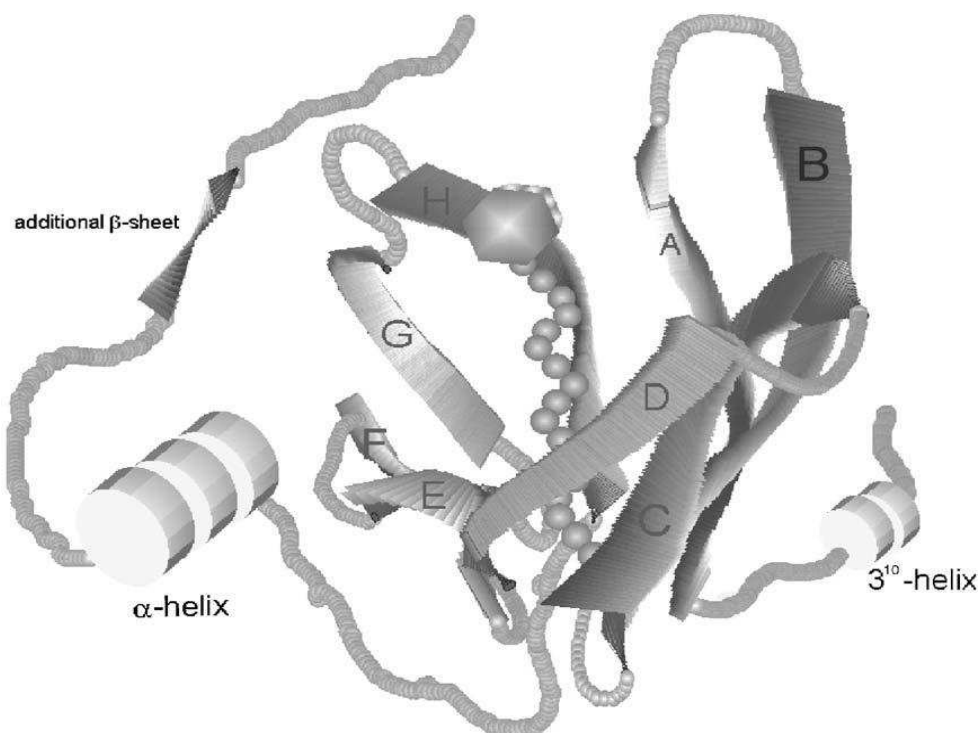


Figura 1: Modelo estrutural da lipocalina demonstrando o cálice e as folhas da estrutura secundária em barril β (A–H). Essa figura foi baseada na estrutura cristalina da β -lactoglobulina (Adaptado de KONTOPIDIS et al., 2002).

4.4.2 Lipocalinas: potencial farmacológico e terapêutico de uma família multifuncional levando à uma pluralidade de alvos moleculares

Especificamente, existe uma variabilidade de funções atribuídas as lipocalinas, dentre as quais se incluem o transporte de retinol, a participação no processo de o transporte de feromônios (p. ex. *Odorant-binding protein*) (TEGONI et al., 1996), a síntese de prostaglandinas, a participação da regulação da resposta imune e a ação como mediador celular homeostático (FLOWER, 1996; FLOWER et al., 2000). Além dessas funções, algumas lipocalinas específicas têm sido investigadas como prováveis marcadores de patologias, por exemplo, a lipocalina de neutrófilo associada a

gelatinase (NGAL), utilizada como biomarcador quantitativo alternativo presente na urina quando ocorre nefrotoxicidade pela cisplatina (MISHRA et al., 2004).

A lipocalina é uma das famílias de proteínas mais importantes em carrapatos da família Argasidae. Quer pelo número de membros ou pelos seus níveis de expressão, visto estarem envolvidas em sistemas de modulação da resposta inflamatória do hospedeiro por captura de mediadores inflamatórios (MANS & RIBEIRO, 2008a; MANS & RIBEIRO, 2008b; MANS et al., 2008; DÍAZ-MARTÍN et al., 2011), bem como com ação no sistema do complemento e agregação de plaquetas (WAXMAN & CONNOLLY, 1993; NUNN et al., 2005; DÍAZ-MARTÍN et al., 2011).

De acordo com Chmelar (2008), estudos da saliva de carrapatos já haviam demonstrado a presença de várias moléculas inibidores de proteases, estas apresentam atividade anticoagulante (NARASIMHAM et al., 2002; FRANCISCHETTI, MATHER & RIBEIRO, 2003), anti-inflamatória e atividade imunossupressora (LEBOULLE et al., 2002; KOTSYFAKIS et al., 2006), constituindo-se, portanto, em moléculas com potencial farmacológico. Como responsáveis por tais atividades foram encontrados prostaglandinas (RIBEIRO et al., 1985; RIBEIRO, MAKOUL & ROBINSON, 1988), apirases (MANS et al., 1998) e as proteínas de ligação pertencentes às lipocalinas (PAESEN et al., 1999; MANS, 2005).

4.4.2.1 Lipocalinas como alvo molecular no desenvolvimento de biofármacos com atividade anti-histamínica e anti-coagulante

A histamina é uma amina biogênica resultante da descarboxilação da histidina possuindo um papel importante nos processos inflamatórios e alérgicos. Sua produção ocorre nos mastócitos (basófilos que migram da corrente sanguínea para o tecido em caso de lesão tecidual), sendo secretadas mediadas por IgE ou por ativação não específica da ação do sistema imune inato (presença de proteínas do Sistema do Complemento ou em decorrência de dano tecidual). Após secreção, essas aminas ligam-se a receptores acoplados a proteína G presentes na membrana das células alvo (receptores H1 e H2) desencadeando uma ação efetora pró-inflamatória (FALUS, 1994; RAIBLE et al., 1994).

Durante o repasto sanguíneo do carrapato apesar de existirem receptores cutâneos de histamina no hospedeiro, esse tem sua reação cutânea inflamatória diminuída o que prolonga a alimentação do artrópode. Caso contrário, a presença de histamina e ligação desta ao receptor específico provoca de maneira direta ou indireta uma diminuição da alimentação, conseqüentemente reduzindo o êxito na reprodução

desse parasito, visto a fêmea depender do repasto sanguíneo para a mesma (KEMP & BOURNE, 1980; WIKEL, 1982). Sendo assim, esse parasito teve que desenvolver estratégias que pudessem suprimir a resposta do hospedeiro.

De fato, estudos realizados a partir do cDNA de glândula salivar de *Rhipicephalus appendiculatus* revelaram a presença de três *Histamin Binding Protein* (HBPs), proteínas da família das lipocalinas capazes de capturar a histamina secretada no local da picada suprimindo a resposta do hospedeiro (PAESEN et al., 1999). Neste carrapato foram isolados três HBPs, denominadas de *Rhipicephalus appendiculatus Histamin Binding Protein*, RaHBP1, RaHBP2 e RaHBP3, apresentando alternadamente ao longo do seu ciclo biológico, sendo que carrapatos fêmeas adultas produzem RaHBP1 e RaHBP2, enquanto que RaHBP3 é exclusivo de larvas, ninfas e carrapatos machos (KELLER et al., 1993; ZHANG et al., 1998).

Além das RaHBPs, estudos realizados na espécie *Dermacentor reticulatus* revelaram uma Proteína de Ligação à Histamina e Serotonina (5-hidroxitriptamina ou 5-HT) SHBP, uma lipocalina com dois sítios de ligação, um para histamina e outro para a 5-HT (PAESEN et al., 1999; SANGAMNATDEJ et al., 2002; BEAUFAYS et al., 2008). Em estudo semelhante, foram identificados os mesmos sítios de ligação (para histamina e/ou 5-HT) em *Argas reflexus*, *Argas monolakensis* e *Ornithodoros savignyi* (MANS, RIBEIRO & ANDERSEN, 2008).

4.5 Papel da glândula salivar, intestino e ovário de *Amblyomma sculptum* na interface microrganismo-vetor-hospedeiro

A estudo da interface microrganismo-vetor-hospedeiro ainda é pouco conhecida, especialmente durante a fase parasitária do carrapato, quando ocorre passagem do patógeno para o hospedeiro vertebrado susceptível, inclusive o homem. Durante o repasto sanguíneo os carrapatos secretam diversas substâncias bioativas através de sua saliva, distribuídas nas mais diversas classes, sendo as principais: anticoagulantes, imunomoduladores e compostos anti-inflamatórios. Todas essas classes têm como finalidade diminuir os mecanismos de defesa no local da picada no hospedeiro, facilitando desta maneira o ingurgitamento do artrópode e desenvolvimento do mesmo em seu ciclo evolutivo (WALKER & ISMAIL, 2008).

Ao se alimentar, os artrópodes precisam transpor as respostas fisiológicas de defesa do hospedeiro. Para tanto faz-se necessário por parte do artrópode, a secreção de substâncias que vão diminuir a resposta inflamatória, manter atividade anti-coagulante bem como destoxificar as defesas do hospedeiro. Do contrário, seria

impossível proceder a um repasto sanguíneo sem ativar mecanismos de defesa do hospedeiro, afetando negativamente o processo de alimentação do parasito (ANDERSEN et al., 2005).

O *A. cajennense* possui em sua saliva substâncias que permitem sua evasão do sistema imune do hospedeiro vertebrado através de um mecanismo de modulação. Isto permite que o mesmo se alimente por vários dias ou semanas, inoculando no hospedeiro componentes anti-hemostáticos, dentre outros compostos, o que vem a impedir ou suprimir a resposta imune local, aumentando desta maneira o potencial de transmissão do agente patogênico para o hospedeiro vertebrado (KOVAR, 2004; KOTSYFAKIS et al., 2006; WALKER & ISMAIL, 2008).

Dentre as várias atividades imunomoduladoras exercidas através da saliva do carrapato inclui-se a diminuição da proliferação de células T, a atividade de macrófagos e células *Natural Killer* (NK) e a produção de citocinas (IL-12, INF γ) (GILLESPIE et al., 2007; WALKER & ISMAIL, 2008), inibindo as funções dos neutrófilos e interferindo no sistema do complemento (KOTSYFAKIS et al., 2006; VALENZUELA, 2004; WALKER & ISMAIL, 2008).

O trato digestivo do *Amblyomma cajennense* é dividido em três porções, sendo, intestino anterior, médio e posterior, esse órgão é responsável pela digestão do sangue ingerido durante o repasto do artrópode no hospedeiro. A ultraestrutura do intestino do carrapato apresenta organelas e inclusões citoplasmáticas sugerindo uma organização complexa de funções que vão desde a alimentação até estratégias de combate a infecção ocorrido no trato digestivo do carrapato, visto que, a digestão ocorre de maneira intraluminal e no intestino dos carrapatos não existem enzimas digestivas, o alimento, no caso o sangue, deve ser absorvido pelas células do lúmen por internalização de substâncias (BALASHOV, 1967; CAPERUCCI et al., 2010).

O trato digestivo fica à mercê das infecções por microrganismos, esse é o habitat das *Rickettsias* ssp no carrapato contribuindo para a manutenção da infecção e também para a transmissão ovariana (BALASHOV, 1967; CAPERUCCI et al., 2010).

4.6 Genes de referência para aplicação em RT-qPCR

Os genes constitutivos são responsáveis por diversas funções do metabolismo celular, encontram-se envolvidos nos mais variados processos sendo necessários a sua expressão em taxas constantes em diversos tecidos durante o desenvolvimento do organismo. Por apresentarem uma taxa de expressão constante na maioria dos tecidos/órgãos estudados, estes são utilizados como controle endógeno (genes de

referência) e são fundamentais para validação dos resultados no estudo da expressão utilizando *reverse transcriptase real time PCR quantitative* (RT-qPCR) (BUSTIN et al., 2009).

Os genes de referência devem por definição apresentar um nível de expressão de mRNA estável diante de uma variedade de condições experimentais (tratado/não tratado, infectado/não infectado) e em diferentes tecidos/órgãos. Atualmente entende-se que este gene ideal não existe, sendo necessária a validação de um gene de referência em cada condição e tecido estudado (VANDESOMPELE et al., 2002; ANDERSEN et al., 2004; BROWNING et al., 2012). A condição mais importante a ser obedecida por um gene de referência é de que o nível da expressão não deve ser afetado pela condição experimental, que exista uma variabilidade mínima na sua expressão dentre os tecidos/órgãos e estados fisiológicos estudados. E ainda, esse gene deve refletir mudanças mínimas de expressão quando técnicas ou tecnologias são mal empregadas durante a qPCR. Geralmente, os genes que satisfazem essa condição são aqueles envolvidos em processos essenciais de sobrevivência celular, conhecidos como *housekeeping genes* (genes constitutivos), porém muitos desses não participam somente de processos básicos mas estão implicados em processos secundários sofrendo alterações das taxas de expressão por esses (TYBURSKI et al., 2008).

Diversos trabalhos têm sido publicados envolvendo a análise da expressão gênica com as mais variadas finalidades, porém vários equívocos de interpretação por falha na execução da metodologia levaram à publicação de vários manuscritos contendo recomendações mínimas para a execução da *quantitative* PCR (qPCR), destes há de destacar os trabalhos de Bustin et al., 2009, que estabelece os critérios mínimos de informação para publicações de experimentos por qPCR e Kozera & Rapacz, 2013 que aborda uma profunda revisão dos genes de referência usados em qPCR.

De maneira a preencher os requisitos de uma qPCR com resultados robustos e confiáveis deve-se utilizar ao menos um controle endógeno visando a normalização dos erros da reação de amplificação, esses são denominados genes de referência e são utilizados como uma amplificação interna de controle, possuindo uma sequência diferente do gene alvo e sofrendo a menor variabilidade possível em sua expressão quando considerados os diferentes tecidos e/ou condições de infecção/não infecção ou tratado/não tratado (CHERVONEVA et al., 2010). Os genes de referência são controles internos que sofrem variações de acordo com os tratamentos estabelecidos durante o

experimento da mesma forma que os genes de interesse (BUSTIN et al., 2009), sendo assim, o melhor gene de referência será aquele que apresentar uma menor variação da expressão durante o experimento. Tal escolha é realizada através de *softwares* estatísticos da análise da expressão gênica que avaliam os valores de Cp (*crossing point*) (RASMUSSEN, 2001) ou *Cycle threshold* (Ct) (LIVAK, 2001), entre vários aplicativos destaca-se o *GeNorm* (VANDESOMPELE et al., 2002), *NormFinder* (ANDERSEN; JENSEN; ONRTOFT, 2004) e o *Bestkeeper* (PFAFFL et. al., 2004). Esses calculam a estabilidade dos genes candidatos a gene de referência durante as condições experimentais levando a uma escolha robusta do melhor gene candidato analisado.

Uma vez escolhido o melhor gene candidato a gene de referência, à amplificação sofrida pelo mesmo vai normalizar os erros experimentais da reação de amplificação. Essa é realizada utilizando a média geométrica da expressão gênica convertida em logaritmo fazendo com que o erro experimental sofrido pelo gene de referência seja considerado para todos os outros desde que nas mesmas condições e com a mesma amostra. Assim obtêm-se uma comparação mais robusta dos resultados obtidos pois o erro amostral foi normalizado significando uma minimização da variância inata ao experimento. De maneira geral, quanto maior o número de genes de referência utilizados numa reação, menor seria a variância inata. Como essa prática é dispendiosa e impraticável em termos experimentais recomenda-se utilizar um gene de referência que tenha sido validado para a situação e tecido em estudo (DHEDA et al., 2004).

Dentre os genes de referência mais utilizados destaca-se o *gliceraldeído-6-fosfato desidrogenase*, *gapdh*. Devido ao grande número de manuscritos publicados com o mesmo, este, de certa maneira tornou-se “clássico” como escolha de um controle endógeno (KOZERA & RAPACZ, 2013). Estudo recente envolvendo a escolha de genes de referência para o carrapato *A. maculatum*, apresentou resultados divergentes na expressão do *gapdh* em diferentes tecidos considerados o que corrobora para uma necessidade de validação do mesmo quando se opta pela sua utilização como gene de referência (BROWNING et al., 2012).

A quantificação dos resultados da qPCR pode ser realizada de forma absoluta ou relativa. Na quantificação absoluta utiliza-se um padrão de mRNA resultante da expressão absoluta do RNA transcrito *in vitro*, seguida de uma etapa de transcrição reversa para a construção de uma fita de cDNA plasmidial dos fragmentos do gene alvo dentro de um vetor plasmidial através da PCR convencional. Esse método é utilizado

quando a quantificação absoluta é necessária para a interpretação dos resultados (por exemplo, na quantificação de DNA viral) (MARTELL et al., 1999; MOODY et al., 2000). Uma das vantagens desse método é que os clones dos plasmídeos contendo o vetor de expressão para o obter-se os fragmentos do gene alvo, construídos pela PCR convencional, podem ser facilmente preparados em maior quantidade, permitindo a realização de um maior número de experimentos utilizando amostras diluídas, aliquotadas e estocadas a -20°C , essas não sofrem degradação significativa quando estocadas por longos períodos. As principais desvantagens da quantificação absoluta residem na necessidade da realização de etapas trabalhosas na construção do cDNA plasmidial dificultada pela falta de controle da variação da eficiência no processo de transcrição reversa do mRNA e a necessidade da obtenção de um cDNA com estabilidade semelhante ao DNA do gene alvo (GIULIETTI et al., 2001).

Outro método de quantificação da qPCR consiste na quantificação relativa, esse é realizado pelo método da curva padrão ou pelo método de comparação do limite do ciclo de detecção (*Crossing point*, C_p , termo conhecido como *Cycle threshold*, C_t , no equipamento *StepOne* da *applied biosystems*). No método da curva padrão utiliza-se uma amostra de concentração conhecida, essa pode ser de DNA dupla fita (dsDNA) purificado em plasmídeo, DNA fita simples (ssDNA) obtido por síntese *in vitro* ou utilizando uma amostra de cDNA resultante da expressão do gene alvo. A concentração das amostras de DNA ou RNA é medida por espectrofotometria em comprimento de onda (λ) a 230, 240 e 260 nm e convertido para o número de cópias de ácido nucleico usando o peso molecular do DNA ou RNA. No método de comparação do limite do ciclo de detecção (C_p) são utilizadas fórmulas aritméticas para o cálculo do nível da expressão relativa comparado a um calibrador, que pode ser uma amostra controle não tratada. Em seguida, a quantidade do gene alvo é normalizada por um gene endógeno em relação ao calibrador, tendo então a expressão $2^{-\Delta\Delta C_p}$ onde o $\Delta\Delta C_p$ consiste na diferença do ΔC_p (amostra) pelo ΔC_p (calibrador), sendo o ΔC_p a diferença entre o C_p do gene alvo pelo C_p do controle do gene endógeno. Essa equação representa a expressão normalizada do gene alvo na amostra desconhecida em relação a expressão do gene alvo na amostra normalizada pelo calibrador. Esse método pode ser usado quando a eficiência dos iniciadores (*primers*) do gene alvo e do controle endógeno são iguais (ou variam em 10%). Portanto, deve-se avaliar o gene constitutivo através da validação do mesmo, o que envolve determinar a melhor concentração e a eficiência do *primer* a ser utilizado (TYBURSKI et al., 2008). Quando obedecidas estas condições e sendo as eficiências comparáveis, pode-se utilizar o método do $2^{-\Delta\Delta C_p}$ para

a análise da expressão (LIVAK & SCHMITTGEN, 2001), caso contrário utiliza-se o método da curva padrão (GIULIETTI et al., 2001).

4.7 Bioinformática na predição de modelos proteicos tridimensionais

As análises de bioinformática constituem em ferramentas indispensáveis para a análise estrutural dos transcritos das proteínas referentes aos genes alvo em estudo. Para a predição de estruturas secundárias e possível conformação da estrutura terciária existem diversos pacotes de aplicativos, muito desses presentes no *Protein Model Portal* (PMP) pertencente à plataforma de aplicativos de bioinformática *Expert Protein Analysis System (Expasy)* (http://www.expasy.org/structural_bioinformatics) dentre estes destacam-se o *RaptorX* (<http://raptorx.uchicago.edu/StructurePrediction/>) (KALLBERG et al., 2012), *Protein Homology/analogy Recognition Engine V.2.0 (Phyre2)* (<http://www.sbg.bio.ic.ac.uk/phyre2>) (KELLEY et al., 2015), *SwissModel* (<http://swissmodel.expasy.org/>), *Integrated web resource for protein fold recognition (IntFOLDTS)* (<http://www.reading.ac.uk/bioinf/servlets/nFOLD/>) (HASS et al., 2013).

O *RaptorX* é um aplicativo de predição da estrutura tridimensional (3D) de proteínas. Esse aplicativo prediz as estruturas 3D de sequências proteicas em comparação com proteínas homólogas no PDB. Os parâmetros utilizados no *RaptorX* para a análise são: análise de domínio, predição de estruturas secundárias, acessibilidade a solvente, *score* de alinhamento, relação entre resíduos de aminoácidos não idênticos pelos resíduos de aminoácido idênticos no alinhamento (*un-identical residues alignment sequence*, USeqID, *identical residues alignment sequence*, SeqID), *global distance test* (GDT), *un-normalized global distance test* (uGDT) e *p-value* (KALLBERG et al., 2012).

Para a análise de domínio, o mesmo deve ser interpretado no *RaptorX* levando em consideração que nesse aplicativo um domínio constitui uma unidade de modelagem, assim a sequência submetida é dividida em unidades baseadas no *template* depositados no *Protein Data Bank* PDB. Desse modo deve-se levar em conta que, dois domínios podem ser modelados como únicos caso compartilhem um molde comum (o que cria um viés de identidade), assim o modelo 3D resultante da análise torna-se confiável quando partilham do mesmo molde, sendo necessário a adição de novos parâmetros para suporte e corrobore com os resultados (MA et al., 2013).

A previsão da estrutura secundária é realizada para a verificação de 3 estados: presença de hélices, folhas β e *loop*, no *RaptorX* esses parâmetros são representados respectivamente pelas letras H, E e C. O parâmetro de acessibilidade do solvente

(ACC) é separado no *RaptorX* em 3 estados divididos pelas faixas de < 10% (letra B no *output* do *RaptorX*), correspondendo a uma estrutura não acessível a solventes, entre 10 e 42% correspondendo à uma faixa intermediária de acessibilidade (letra M no *output* do aplicativo) e > 42% correspondendo a uma estrutura que permite um grande acesso a solventes (letra E no *output* do aplicativo) (PENG & XU, 2011a).

De maneira a diminuir o viés de identidade e concomitantemente avaliar a qualidade do modelo 3D previsto o *RaptorX* utiliza-se do *score* de alinhamento, uSeqID e SEQID, uGDT e GDT e o *p-value*. O *score* de alinhamento avalia a relação entre o tamanho da sequência analisada com a parte da mesma alinhada ao *template* depositado. Sendo que um valor de 0 ou próximo do mesmo demonstra falha na determinação da estrutura 3D. No intuito de diminuir o viés de identidade, utiliza-se a relação entre uSeqID e SEQID. O uSeqID representa o número de resíduos idênticos no alinhamento (o que pode gerar uma modelagem com unidades sobrepostas) assim realiza-se a normalização em percentagem deste valor com a SEQID. Desse modo, quanto maior a razão uSeqID/SEQID melhor a qualidade da estrutura prevista. E ainda se SEQID > 30% e a proteína (ou domínio) for > 200 resíduos, os resultados indicam que a predição dos dobramentos da proteína em estudo apresentada no modelo está correta (PENG & XU, 2011b).

Os parâmetros, uGDT e GDT avaliam a qualidade do modelo absoluto através da estimação do erro de modelagem. Assim, para uma proteína > 100 resíduos uGDT > 50 constitui em um bom indicador. E ainda, se GDT < 50 mesmo que uGDT > 50 os resultados indicam que somente uma pequena porção do modelo é verdadeira. O aplicativo *RaptorX* avalia a probabilidade da predição ao acaso do modelo 3D resultante através do *p-value*, portanto, quanto menor o *p-value* melhor a qualidade do modelo obtido. Para proteínas que apresentem predomínio de estrutura secundária em α hélice, o *p-value* < 10E-3 é o ideal sendo que para as proteínas que apresentem predomínio de estrutura em folhas β , espera-se um valor de *p-value* < 10E-4 (KALLBERG et al., 2012).

O aplicativo *PHYRE2* (<http://www.sbg.bio.ic.ac.uk/phyre2>) trabalha na predição de modelos estruturais de proteínas a partir da sequência linear inserida como *input*. Essa sequência é comparada às sequências presentes no *Position-Specific Iterative Basic Local Alignment Search Tool* (PSI-BLAST) o mesmo trabalha com uma matriz específica de posição e pontuação (PSSM) aliada a um perfil de alinhamento múltiplo de comparação das sequências proteicas e validado por um *score* de pontuação mínima. Essa matriz é utilizada para atualizar a base de dados quando ocorre a

inserção de novas sequências, permitindo um meio de avaliação da correlação de distância entre proteínas suportando através dessas relações inferências quanto a função biológica dessas proteínas nos organismos em que ocorrem (ALTSCHUL et al., 1997). E ainda, o *PHYRE2* possui um algoritmo de alinhamento perfil a perfil de proteína, através da relação proteína *input* buscando um modelo a ser predito/proteína, *template* depositado no PSI-BLAST, usado devido à similaridade de sequência, que avalia a paridade com os modelos do algoritmo presente no *Homologues Hidden Search 1* (HHsearch1) que tem como base o algoritmo *Hidden Markov Models* (HMM) (KROGH et al., 1994; EDDY, 1998; KARPUS et al., 2001; SODING, 2005) melhorando a precisão, confiabilidade e a taxa de execução do alinhamento obtido.

Para a modelagem de estruturas cujo *template* não é conhecido o *PHYRE2* incorpora uma nova ferramenta de avaliação dos resultados do dobramento predito, o aplicativo *POING2* (<http://www.sbg.bio.ic.ac.uk/phyre2/poing2>), capaz de modelar as regiões de proteínas sem homologia detectável com sequências conhecidas, podendo combinar vários modelos. Os resultados gerados pelo *PHYRE2* nos fornece uma visão geral da estrutura 3D da proteína em estudo, uma visão da estrutura secundária e da desordem relacionado ao modelo predito, uma análise de domínio, detalhamento do modelo, visualização do alinhamento utilizado para busca de um provável *template* e possível localização de porção com hélice transmembrana, quando presente, na proteína em estudo (KELLEY et al., 2015).

Em 1996, o *European Bioinformatics Institute* (EMBL-EBI) e o *Swiss Institute of Bioinformatics* (SIB) constituíram o *Swiss-Prot* e o *Translate Sequence for European Molecular Data Library TrEMBL*, além desses, outro banco de dados de proteínas o *Protein Sequence Database* (PIR) compilava a sequência do Protein Sequence Database (PSD) compondo o banco PIR-PSD. Com o projeto genoma, o volume de dados a ser processados ultrapassou a capacidade de análise do *Swiss-prot* sendo assim foi criado o *TrEMBL*. A seção *TrEMBL* foi introduzida na composição do banco de dados do *Swiss-prot* e contém anotações de alta qualidade dos registros analisados. O *input* pelo *Swiss-prot*, é realizado de forma manual, o que dificulta o processamento de grande volume de dados sendo assim, o banco não cobria a demanda disponível. No caso do *TrEMBL* a alimentação e registro de dados é realizado de forma automática tanto a anotação quanto a classificação da proteína em estudo, permitindo uma avaliação de grande volume de dados. Assim, foram mantidos dois bancos de dados de modo que os resultados não são sobrepostos, o que ampliou o volume de sequências proteicas depositados nos dois bancos de proteínas disponibilizando os registros para

o público rapidamente. Em paralelo manteve-se o banco PIR-PSD e sua base de dados relacionados, incluindo o *International Protein Class* (iProClass), um banco de dados de sequências e classificação de famílias proteicas (POUX et al., 2014).

Em 2002, o *Swiss-prot* e o *TrEMBL* se fundiram passando a constuir o *Resource Protein Universal* (UniProt) resultado da formação de um consórcio entre o *European Bioinformatics Institute* (EMBL-EBI), o *Swiss Institute of Bioinformatics* (SIB) e o *Protein Information Resource* (PIR). O financiamento desse consórcio advém de diversos institutos de pesquisas, o *National Institute Health* (NIH), o *European Molecular Biology Laboratory* (EMBL), a *British Heart Foundation* (BHF), a *Parkinson's Disease United Kingdom* (PDUK) além desses institutos o governo federal da Suíça e o *National Science Foundation* (NSF) completam o recurso desse consórcio. O UniProt compila os resultados experimentais, *inputs* e conclusões científicas compondo 2 seções de bancos distintos, o *UniProtKB/SwissProt* (alimentado manualmente) e *UniProtKB/TrEMBL* (alimentado automaticamente). As bases de dados do *uniprot* são a *UniProt Knowledge Base* (UniProtKB), o *UniProt Reference Clusters* (UniRef), e *UniProt Archive* (UniParc) (LEINONEN et al., 2004; SUZEK et al., 2007; POUX et al., 2014).

5. OBJETIVOS

5.1 OBJETIVO GERAL

Identificar transcritos de lipocalinas com domínio de ligação a histamina e serotonina e potenciais candidatos a genes de referência provenientes dos transcriptomas de glândulas salivares, intestinos e ovários do carrapato *Amblyomma sculptum* (complexo *Amblyomma cajennense*, sinonímia *A. cajennense*) em situação de não infecção e infecção com *Rickettsia amblyommii*.

5.2 OBJETIVOS ESPECÍFICOS

5.2.1 Realizar o alinhamento dos transcritos de lipocalinas do transcriptoma de glândula salivar do carrapato *A. sculptum*;

5.2.3 Realizar a análise funcional dos transcritos de lipocalinas do transcriptoma de glândula salivar do carrapato *A. sculptum*;

5.2.4 Confrontar os transcritos do RNA-seq de glândulas salivares de *A. sculptum* com os depositados no *Swiss-prot*;

5.2.5 Identificar no RNA-seq de glândulas salivares de *A. sculptum* sequências proteicas de lipocalinas com domínios de ligação a serotonina e histamina;

5.2.5 Desenhar os *primers* necessários para análise da expressão de lipocalinas provenientes do RNA-seq de glândulas salivares de *A. sculptum*;

5.2.6 Desenhar os *primers* necessários para análise da expressão de genes constitutivos potenciais candidatos a gene de referência;

5.2.7 Analisar os genes alvo selecionados na validação da RT-qPCR para estudo da expressão relativa de lipocalinas com domínio SHBP

5.2.8 Realizar a predição estrutural *in silico* das lipocalinas com domínio SHBP dos genes selecionados para a RT-qPCR;

5.2.9. Optimizar os *primers* necessários para análise da expressão de lipocalinas;

5.2.10. Validar os *primers* necessários para análise da expressão de genes constitutivos potenciais candidatos a gene de referência;

5.2.11. Analisar a expressão diferencial de lipocalinas e potenciais genes de referência em glândula salivar intestino e ovário de *A. sculptum*.

6. MATERIAL E MÉTODOS

6.1 Refino e alinhamento dos transcritos referentes a lipocalinas advindas da base de dados transcriptômicos do carrapato *A. sculptum*

A partir dos resultados do transcriptoma de *A. sculptum* verificou-se a anotação de 189 transcritos de lipocalinas (MOREIRA, 2012) sendo que após verificação do códon de iniciação e eliminação das sequências truncadas selecionou-se 121 (Tabela S1). Em seguida, realizou-se uma análise comparativa entre as sequências de proteínas anotadas no transcriptoma de *A. sculptum* com aquelas depositadas no banco de dados de proteínas *Swiss-Prot*, resultando em 128 sequências de lipocalinas de *A. cajennense* oriundas do trabalho de Garcia, 2014. Após retiradas as sequências truncadas e incompletas essa etapa de refino e alinhamento resultou na seleção de um *dataset* contendo 105 sequências de lipocalinas.

Todos os alinhamentos deste manuscrito foram realizados nos aplicativos MAFFT versão 7.0, (mafft.cbrc.jp/alignment/server/) e CLUSTAL W (THOMPSON et al., 1994), para eliminar posições e regiões divergentes utilizamos o aplicativo Gblocks (molevol.cmima.csic.castresana.org/Gblocks_server) (CASTRESANA, 2000). As análises de *Maximum-likelihood* (ML) foram realizadas no aplicativo MEGA versão 5.0 (KUMAR et al., 2008) e as análises bayesianas no aplicativo Mr Bayes versão 3.1 (MB) (HUELSENBECK & RONQUIST, 2001; RONQUIST & HUELSENBECK, 2003). Para predição do melhor modelo de substituição de aminoácidos foi utilizado o aplicativo ProtTest 2.4.server (darwin.uvigo.es/software/prottest2_server.html), (ABASCAL, 2005). Todas as análises de bioinformática foram processadas no Núcleo de Biologia de Moléculas (NuBiomol/UFV), utilizando os processadores da Divisão de Suporte ao Desenvolvimento Científico e Tecnológico da Universidade Federal de Viçosa, Brasil.

6.2 Predição de domínios nos transcritos de lipocalinas de *A. sculptum*

A fim de verificarmos a presença de possíveis domínios funcionais, em especial *Histamin Binding Protein* (HBP) e *Serotonin Histamin Binding Protein* (SHBP) nos transcritos das 121 lipocalinas preditas de glândulas salivares de *A. sculptum*, procedemos à verificação individual das sequências submetendo-as à análise de predição de domínios por *Hidden Markov Models* (HMM) usando o banco de dados *Protein Family database*, PFAM 27.0 (FINN et al., 2014).

6.3 Análise filogenética dos transcritos de lipocalinas preditos em transcriptomas de glândulas salivares de *Amblyomma sculptum*

Para a análise filogenética entre as 105 sequências de lipocalinas do transcriptoma de *A. cajennense* depositados no *Swiss-prot* foi realizado um novo alinhamento utilizando como probabilidade *um bootstrap* até 1.000 vezes replicado (HASEGAWA & KISHINO, 1994). Em seguida, construímos uma árvore bayesiana, utilizando o método de Monte Carlo – Cadeia de Markov, com 1.000.000 de gerações, contendo quatro cadeias e duas corridas (nchains=4; nruns=2) (HUELSENBECK & RONQUIST, 2001; RONQUIST & HUELSENBECK, 2003).

6.4 Extração do RNA total das glândulas salivares, intestinos e ovários do carrapato *A. sculptum*

Após a extração, os órgãos internos foram divididos em *pools*, de acordo com os tratamentos (infectado e não infectado), resultando em 6 *pools* de amostra de RNA total, distribuídos em 75 glândulas salivares 75 intestinos e 75 ovários buscando a obtenção de 3 a 5 ug de RNA total de cada pool de órgãos representado, e imediatamente imersos em solução de estoque *RNA later* (*Ambion, Austin, TX, EUA*) e mantidos a 4°C por 1 hora sendo depois armazenados em freezer -80°C para as etapas posteriores de extração de RNA. Toda a bancada e material de trabalho (lupas, pinças, placas, etc) foram tratados com solução *RNAZap* (*Invitrogen, San Diego, CA, USA*) para remoção de RNAses, evitando-se a degradação do RNA total das amostras.

A extração do RNA total da glândula salivar, intestino e ovário foi realizada através do kit *SV Total RNA Isolation kit* (*Promega*®) de acordo com as recomendações descritas no manual do fabricante. A qualidade do RNA total extraído foi avaliada pela integridade das bandas do RNA ribossômico em gel de agarose 1,0%, corado com brometo de etídeo (SAMBROOK et al., 2001). A pureza e concentração das amostras foram determinadas pela leitura de 1 ul de amostra em espectrofotômetro *NanoDrop 2000* (*ThermoScientific*®), que fornece automaticamente as concentrações de RNA em ng/uL, bem como as razões A260/A280 e A260/A230 indicando assim a qualidade do RNA extraído, nessa etapa buscou-se um resultado na faixa de 1,8 a 2,2 indicando uma ótima qualidade de amostra para utilização em etapas posteriores. Após extração, quantificação e verificação de sua integridade, as amostras de mRNA foram armazenadas a -80°C até a síntese de cDNA.

6.5 Síntese do cDNA para avaliação da expressão gênica por RTq-PCR

A síntese do cDNA foi obtida através da transcrição reversa do mRNA, utilizando o MMLV *reverse transcriptase* (Invitrogen), de acordo com as recomendações do fabricante. Detalhadamente, a partir de aproximadamente 3 ng de RNA total extraído das glândulas salivares, intestino e ovário provenientes dos carrapatos alimentados e não infectados e alimentados infectados, todos livres de patógenos cedidos pela UFU, realizou-se a síntese de cDNA tendo por molde o mRNA. Para minimizar o viés inerente da técnica de qPCR, algumas etapas foram verificadas. Inicialmente, após a extração do RNA total da amostra a mesma foi purificada com a utilização do Kit *RNAqueous*, *Invitrogen*[®], de acordo com as recomendações descritas no manual do fabricante, foram avaliadas a concentração e a integridade do produto extraído. Para manter a qualidade e a integridade da amostra foi adicionado ao RNA extraído a DNase I (proveniente do *Kit RNAOUT*, *Invitrogen*[®]). A concentração do cDNA sintetizado foi avaliado pela leitura de 1 ul de amostra em espectrofotômetro *NanoDrop 2000* (*ThermoScientific*[®]), tendo como ideal um resultado na faixa de 1,8 a 2,2 indicando uma ótima concentração de amostra para utilização em etapas posteriores.

6.6 Desenho de *primers*, teste de concentração e eficiência para validação da qPCR

Após a síntese do cDNA foram desenhados 15 *primers* utilizados na amplificação do cDNA dos genes alvo. Para tanto utilizamos o aplicativo *Primer express 3.0* (<https://products.appliedbiosystems.com>). Todos os *primers* foram sintetizados pela empresa *Integrated DNA Technologies*, IDT (Coralville, Iowa, USA). Foram desenhados nove *primers* de lipocalinas que apresentaram, *in silico*, *Serotonin* e *Histamin Bindin Protein Domain* (*lip1*, *lip4*, *lip7*, *lip10*, *lip14*, *lip17*, *lip21*, *lip24* e *lip27*), para nove genes de lipocalinas (A0A023FS71, A0A023FUB8, A0A023FFA2, A0A023FD30, A0A023FD26, As_6240, A0A023FUX8, A0A023FS44 e A0A023FU99) e seis pares de *primers* (*hk1*, *hk2*, *hk3*, *hk4*, *hk5* e *hk6*) para três genes candidatos a genes de referência (A0A023FLV9, *glicose-6-fosfato-desidrogenase*, *g6pd*; A0A023FJ24, *3-hidroxi-coenzima-a-desidrogenase*, *3ohcoadh*; A0A023FLU5, *gliceraldeído-3-fosfato-desidrogenase*, *gapdh*).

As reações de qPCR foram realizadas no aparelho *StepOne* (*Applied Biosystems*) relacionando os valores de *Cp*, parâmetro que indica o ponto em que a fluorescência aumenta acima do ruído, ou seja onde a fluorescência atinge níveis aumentados indicando a presença de uma fluorescência maior que a do corante ROS utilizado na normalização de qPCR de DNA dupla fita (*Crossing point*, ou *Cycle threshold*). A quantificação relativa da expressão foi determinada pelo método da curva

de calibração normalizada pelo aplicativo *Relative expression software tool* (REST) visto que as eficiências dos *primers* não foram comparáveis (LIVAK & SCHMITTGEN, 2001). Para detecção da reação utilizou-se o sistema *SYBR Green Master Mix* (Thermocientific®). Todas as amostras dos experimentos de qPCR foram realizadas em replicatas, em cada placa de reação tendo um controle negativo (*Negative Template Control*, NTC) para cada alvo estudado além do controle endógeno para cada tipo de tecido.

Após a síntese dos *primers* realizou-se o teste de concentração e eficiência dos mesmos através de qPCR. Para o teste de concentração foram preparadas soluções de *primers forward* e *reverse* na concentração de 4 uM e 10 uM (soluções estoque), tendo como mix para uma reação 6 uL de 2X *SYBR Green Master Mix* (Thermocientific®), 3 uL de *primer F/R* nas concentrações de 0,4; 0,6; 0,8 e 1,0 uM, 2 uL de água *nuclease free* (Ambion®) e 1 uL de cDNA. Para o teste de eficiência utilizou-se as diluições seriadas de 200, 100, 50, 25, 12,5, 6,25, e 3.125 uM de cDNA. O coeficiente angular da reta (*slope*) foi calculado pela equação: Eficiência PCR = $(10^{(1/-slope)} - 1) \times 100$ através do aplicativo *StepOne* versão 2.3 (*Life Technologies*®). A análise estatística dos resultados foi realizada no aplicativo *GraphPad Prism* (GraphPad software Inc., San Diego, CA), foram utilizados os testes de ANOVA-way com um nível de confiança de 95% (5% de erro amostral) na comparação dos grupos de glândulas salivares, intestinos e ovários em situação de infecção e não infecção com *R. amblyommii*.

O programa para amplificação no teste de concentração e de eficiência foi de 2 passos sendo *Holding* 95 °C, 10 min.; 40 ciclos de desnaturação a 95 °C por 15 seg.; anelamento e extensão a 60 °C por 1 min. Após amplificação de 40 ciclos realizou-se a desnaturação gradual com vistas a realização da curva de dissociação (curva de *melt*) utilizando o aumento de 0,3 °C por 1 min partindo de 60 °C até o limite de 95 °C. As curvas de concentração das análises das amostras de glândulas salivares, intestinos e ovários foram realizadas com 7 pontos de concentração sendo 200, 100, 50, 25, 12,5, 6,25 e 3,125 uM. Analisadas em replicata contendo como controle endógeno o *gapdh* na concentração otimizada de 0,8 uM.

6.7 Predição estrutural *in silico* das lipocalinas com domínio SHBP provenientes dos genes selecionados para a RT-qPCR;

A sequência das proteínas referentes aos transcritos dos genes alvo dos *primers* selecionados no item 6.6 (lip 4, 7 e 24) foram analisados quanto à presença de

estruturas secundárias e possível conformação da estrutura terciária através do pacote de aplicativos presentes no *Protein Model Portal* (PMP) (HASS et al., 2013) pertencente a plataforma de aplicativos de bioinformática *Expert Protein Analysis System* (Expasy) (http://www.expasy.org/structural_bioinformatics). No PMP as sequências proteicas advindas dos genes *lip 4*, *7* e *24* foram analisadas pelos aplicativos RaptorX (<http://raptorx.uchicago.edu/StructurePrediction/>) (KALLBERG et al., 2012), *Protein Homology/analogy Recognition Engine* V.2.0 (*Phyre2*) (<http://www.sbg.bio.ic.ac.uk/phyre2>) (KELLEY et al., 2015), e *SwissModel* (<http://swissmodel.expasy.org/>).

A figura 2 indica a estratégia abordada na utilização do aplicativo *RaptorX*.

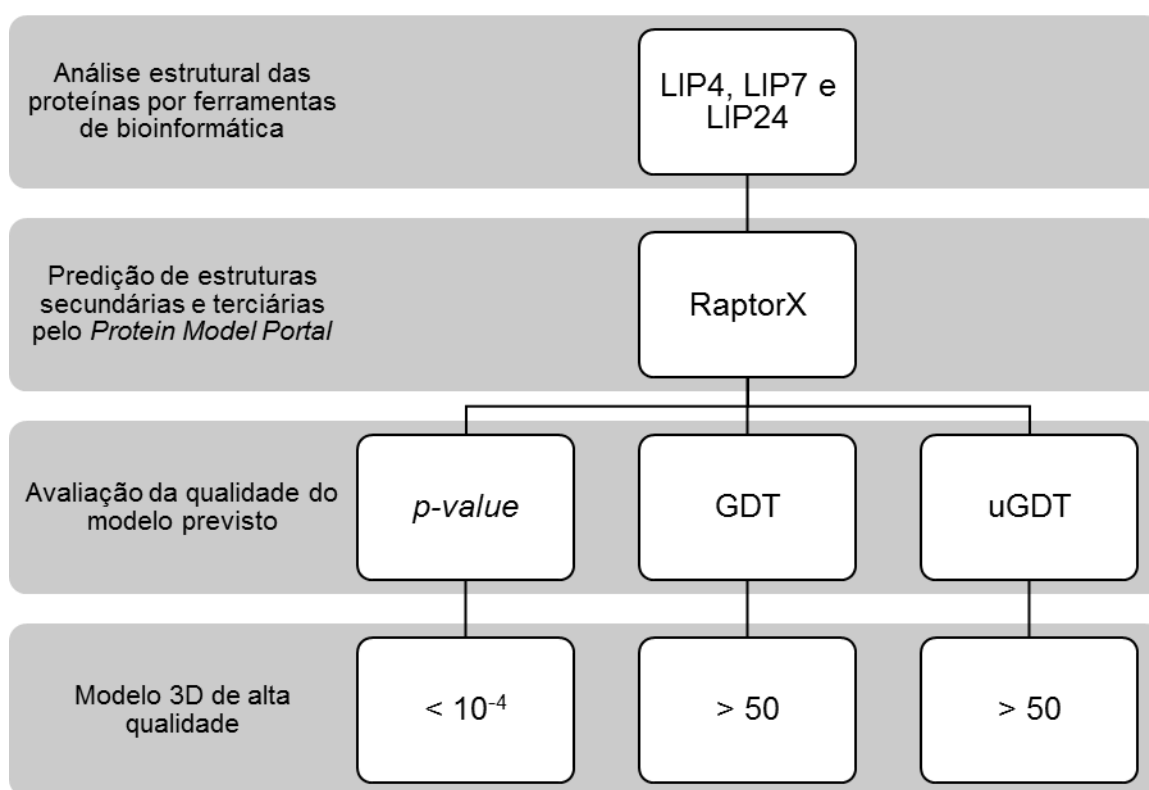


Figura 2: Estratégia utilizada para a predição e avaliação dos modelos 3D das lipocalinas LIP4, LIP7 e LIP24. *p-value*: valor de p; GDT: *Global Distance Test*; uGDT: *un-normalized Global Distance Test* (uGDT).

A figura 2 apresenta o fluxo de análise para uma predição e avaliação do modelo estrutural das proteínas LIP4, LIP7 e LIP24. O aplicativo *RaptorX*, foi utilizado para prever e avaliar a qualidade do modelo previsto. Esse aplicativo, utiliza principalmente 3 parâmetros fundamentais o *p-value*, o uGDT e o GDT, estes em conjunto possuem maior peso na avaliação. Um modelo que apresente um *p-value* pequeno (<10E-4), e uGDT e GDT ambos > 50 constitui-se num modelo 3D de alta qualidade. Através do

input individual das sequências LIP 4, 7 e 24, foram preditos a estrutura secundária e terciária, prováveis sítios de ligação, acessibilidade a solventes e prováveis regiões desordenadas resultantes da composição dos modelos para os transcritos citados. Nessa análise, o *RaptorX* atribui valores de credibilidade aos modelos previstos confrontando o *template* já depositado no PDB com o modelo sugerido para a sequência analisada. Para tanto, realiza-se uma análise global da qualidade do modelo através do estabelecimento do *p-value*, *global distance test* (GDT), *un-normalized global distance test* (uGDT), esse último para avaliar o erro de modelagem de cada resíduo de aminoácido presente (KALLBERG et al., 2012).

A figura 3 indica a estratégia abordada na utilização do aplicativo *Phyre2*.

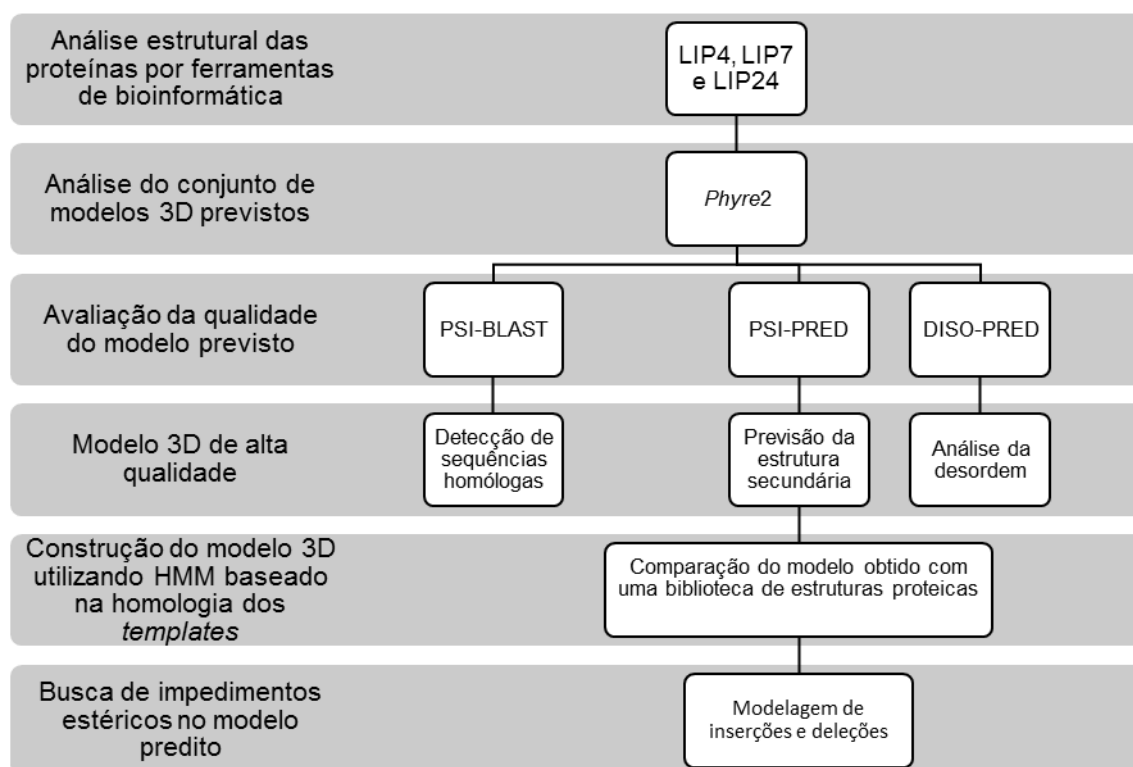


Figura 3: Estratégia utilizada no aplicativo *Phyre2* para a comparação e avaliação dos modelos 3D das lipocalinas LIP4, LIP7 e LIP24 preditos pelo aplicativo *RaptorX*. PSI-BLAST: *Position Specific Iterative Basic Local Alignment Search Tool*; PSI-pred: *Position-Specific Iterative prediction*; Diso-pred: *disorder prediction*.

A figura 3 apresenta o fluxo de análise para a predição e avaliação do modelo estrutural das proteínas LIP4, LIP7 e LIP24. O aplicativo *Phyre2*, foi utilizado para avaliar a qualidade dos modelos preditos. A análise das sequências proteicas referentes aos transcritos *lip 4, 7 e 24* no *Phyre2* produziu um conjunto de modelos 3D do *input*, baseadas em *templates* depositados, feito pela detecção de sequências homólogas com o PSI-BLAST, em seguida fez-se uma previsão da estrutura secundária e análise da desordem com o *Position-Specific Iterative prediction* (PSI-pred) e *disorder prediction* (Diso-pred). Após essa etapa de predição e análise da desordem foi

construído um modelo utilizando *Hidden Markov Models* (HMM) baseado na homologia dos *templates* depositados seguido da comparação do modelo obtido com uma biblioteca de estruturas proteicas preditas por HMM e experimentalmente caracterizadas. Finalmente foi realizado um ajuste através da modelagem de inserções e deleções utilizando uma biblioteca de *loop* associado a comparação de buscas de prováveis impedimentos estéricos, resultando no modelo final (KELLEY et al., 2015).

A figura 4 indica a estratégia utilizada para predição da topologia de hélices transmembrana nas proteínas estudadas.

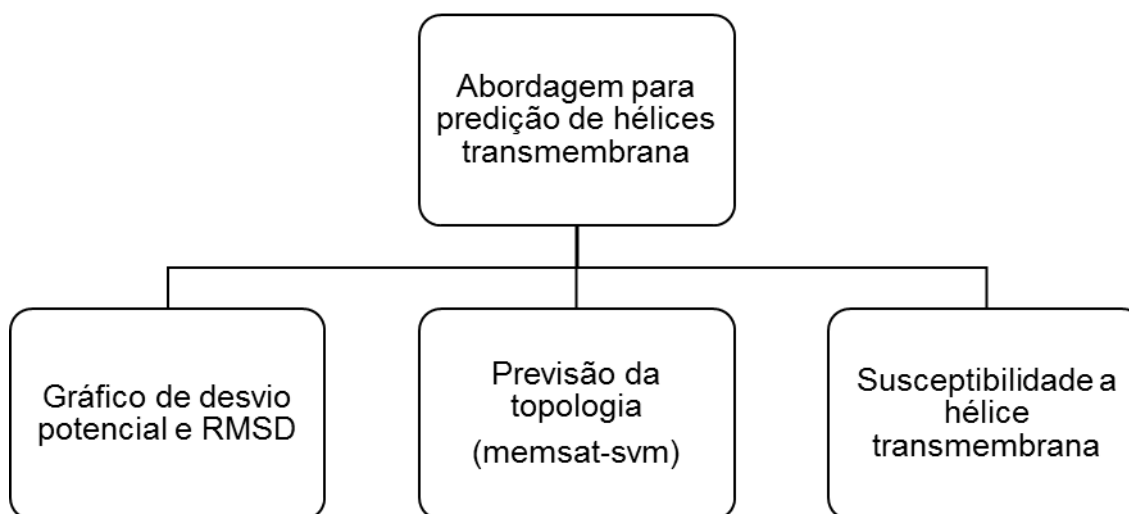


Figura 4: Estratégia utilizada para a predição da presença de hélices transmembrana nos modelos 3D das lipocalinas LIP4, LIP7 e LIP24 preditos pelo aplicativo *RaptorX* e *Phyre2*. Memsat-svm: *Support Vector Machine* (memsat-svm); RMSD: Desvio potencial da energia média quadrática.

A figura 4 apresenta o fluxo de análise para a predição da topologia de hélices transmembrana caso sejam detectadas no modelo estrutural das proteínas LIP4, LIP7 e LIP24. Para a previsão dessas hélices utilizou-se pontuações das medidas gráficas de desvio potencial e da energia média quadrática (RMSD) para efeito de comparação com os modelos depositados, seguido da previsão da topologia pelo aplicativo *Support Vector Machine* (memsat-svm) que determina se as sequências preditas são susceptíveis a conterem hélices transmembrana e a topologia destas. A abordagem utilizada por essa ferramenta possui uma precisão média de 89% (NUGENT & JONES, 2009; KELLEY et al., 2015).

A figura 5 indica a estratégia a avaliação da estrutura terciária predita pelos aplicativos *RaptorX* e *Phyre2* pelo diagrama de Ramachandram.

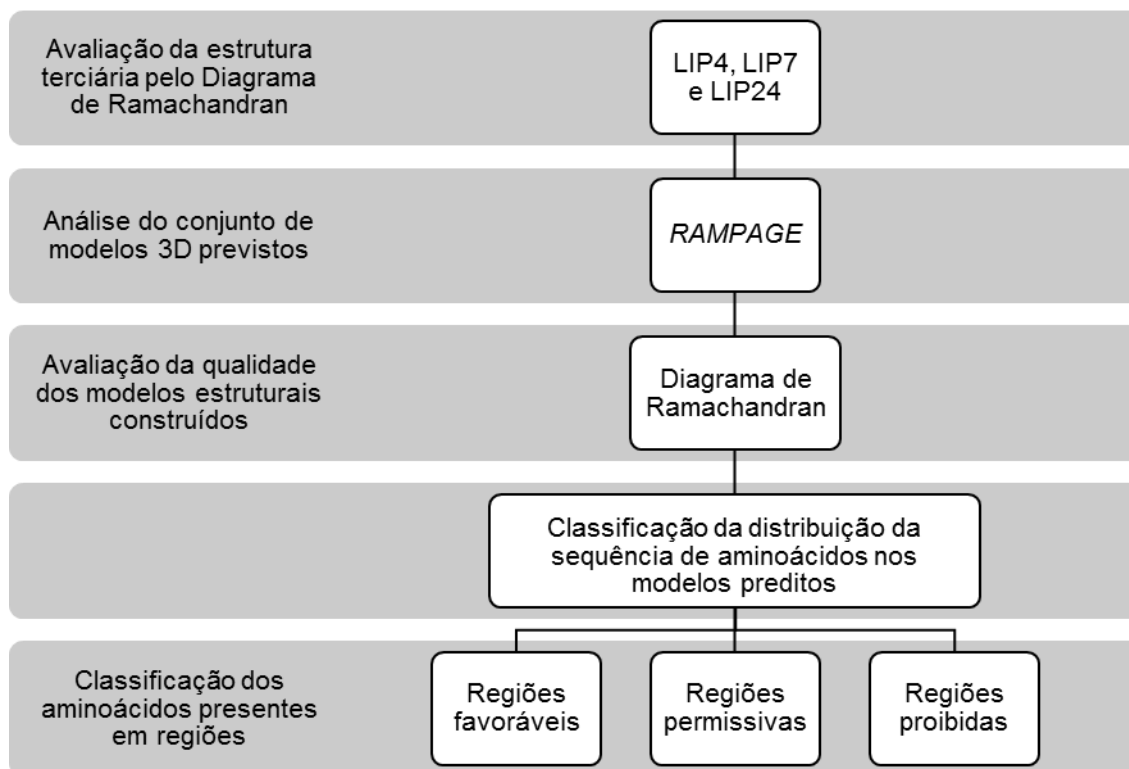


Figura 5: Estratégia utilizada para a avaliação da estrutura terciária dos modelos 3D das lipocalinas LIP4, LIP7 e LIP24 preditos pelo aplicativo *RaptorX* e *Phyre2*. RAMPAGE: *Ramachandran Plot Analysis*.

A figura 5 apresenta o fluxo de análise para a avaliação da qualidade das estruturas terciárias preditas no modelo 3D das proteínas LIP4, LIP7 e LIP24. Adicionalmente para a avaliação da estrutura terciária predita pelos aplicativos utilizados, os modelos gerados foram avaliados pelo aplicativo *Ramachandran Plot Analysis* (RAMPAGE) (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>), que fornece informações sobre a qualidade dos modelos estruturais de proteínas com base no diagrama de Ramachandran (figura 6). Este classifica a distribuição da sequência de aminoácidos presentes na proteína em regiões favoráveis, permitidas e proibidas atribuindo *scores* percentuais às mesmas corroborando na avaliação do modelo predito em outros aplicativos (LOVELL et al., 2003).

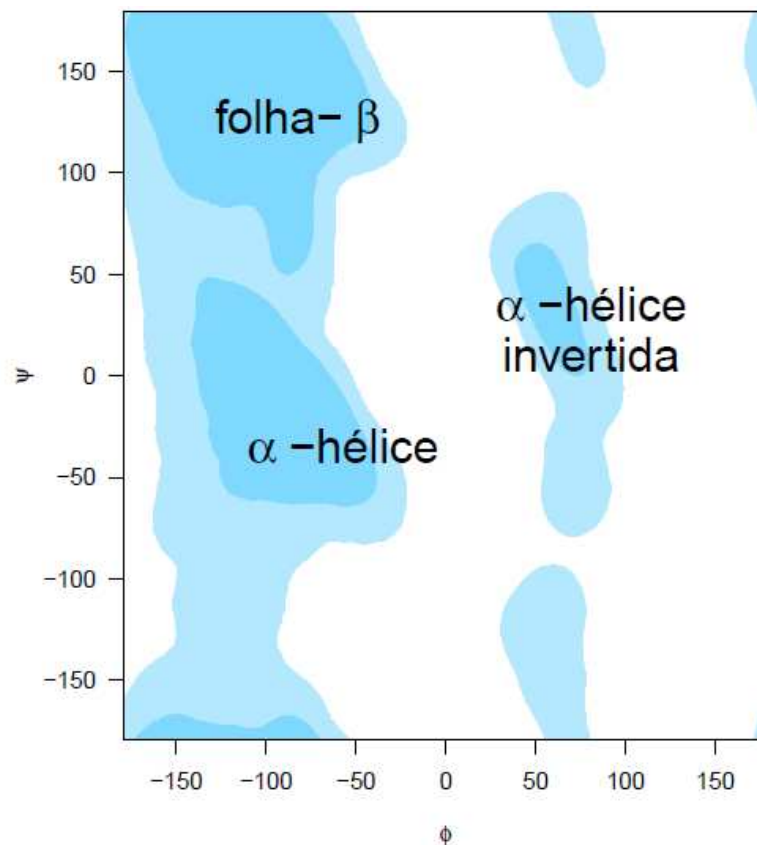


Figura 6: Diagrama de Ramachandran com a predição das estruturas secundárias mostrando a densidade de valores do par (ψ, ϕ) (BERG et al., 2002).

A partir da análise dos ângulos diedrais representados pelo par (ψ, ϕ) é contruído um diagrama distribuindo os aminoácidos em posições favoráveis permitidas e proibidas gendo um score e assim validando a estrutura terciária predita em outros aplicativos (BERG et al., 2002).

7. RESULTADOS E DISCUSSÃO

7.1 Análise filogenética das lipocalinas de *A. sculptum*

Através da busca de sequências de lipocalinas no *uniprot* verificamos a existência de 105 sequências de *A. cajennense* depositadas (GARCIA et. al., 2014). Nas bases de dados obtidas em nosso laboratório para *A. sculptum* foram identificados 121 transcritos de lipocalinas em glândula salivar (tabela 1). A tabela S1 no material suplementar apresenta todas as sequências dos transcritos do RNA-seq utilizadas nesse trabalho.

Tabela 1: Relação das 121 sequências de lipocalinas advindas do transcriptoma de glândula salivar de *A. sculptum*.

No.	Ref. seq	Proteína	No.	Ref. seq	Proteína	No.	Ref. seq	Proteína
1	As_10488	As_LIP1	51	As_7195	As_LIP51	101	As_31300	As_LIP101
2	As_30706	As_LIP2	52	As_1555	As_LIP52	102	As_32490	As_LIP102
3	As_30883	As_LIP3	53	As_2043	As_LIP53	103	As_10679	As_LIP103
4	As_10489	As_LIP4	54	As_15657	As_LIP54	104	As_12499	As_LIP104
5	As_32787	As_LIP5	55	As_55154	As_LIP55	105	As_102988	As_LIP105
6	As_40720	As_LIP6	56	As_44527	As_LIP56	106	As_31301	As_LIP106
7	As_61324	As_LIP7	57	As_97593	As_LIP57	107	As_51348	As_LIP107
8	As_6243	As_LIP8	58	As_10703	As_LIP58	108	As_81142	As_LIP108
9	As_23634	As_LIP9	59	As_6776	As_LIP59	109	As_44528	As_LIP109
10	As_6242	As_LIP10	60	As_17867	As_LIP60	110	As_110797	As_LIP110
11	As_7608	As_LIP11	61	As_50138	As_LIP61	111	As_56745	As_LIP111
12	As_53994	As_LIP12	62	As_84148	As_LIP62	112	As_10927	As_LIP112
13	As_55838	As_LIP13	63	As_45076	As_LIP63	113	As_98550	As_LIP113
14	As_6241	As_LIP14	64	As_17859	As_LIP64	114	As_40031	As_LIP114
15	As_7785	As_LIP15	65	As_56562	As_LIP65	115	As_80509	As_LIP115
16	As_5909	As_LIP16	66	As_17864	As_LIP66	116	As_31881	As_LIP116
17	As_6240	As_LIP17	67	As_114525	As_LIP67	117	As_5660	As_LIP117
18	As_86866	As_LIP18	68	As_17873	As_LIP68	118	As_106712	As_LIP118
19	As_31297	As_LIP19	69	As_115809	As_LIP69	119	As_21521	As_LIP119
20	As_76441	As_LIP20	70	As_113877	As_LIP70	120	As_21522	As_LIP120
21	As_6244	As_LIP21	71	As_8440	As_LIP71	121	As_95768	As_LIP121
22	As_113729	As_LIP22	72	As_108939	As_LIP72			
23	As_55591	As_LIP23	73	As_54515	As_LIP73			
24	As_61325	As_LIP24	74	As_53938	As_LIP74			
25	As_30884	As_LIP25	75	As_17862	As_LIP75			
26	As_30885	As_LIP26	76	As_17863	As_LIP76			
27	As_6243	As_LIP27	77	As_2042	As_LIP77			
28	As_89672	As_LIP28	78	As_108578	As_LIP78			
29	As_6240	As_LIP29	79	As_98549	As_LIP79			
30	As_30880	As_LIP30	80	As_8430	As_LIP80			
31	As_6244	As_LIP31	81	As_54075	As_LIP81			
32	As_89125	As_LIP32	82	As_56070	As_LIP82			
33	As_30881	As_LIP33	83	As_55222	As_LIP83			
34	As_65847	As_LIP34	84	As_56563	As_LIP84			
35	As_61325	As_LIP35	85	As_80499	As_LIP85			
36	As_88490	As_LIP36	86	As_17865	As_LIP86			
37	As_83441	As_LIP37	87	As_58208	As_LIP87			
38	As_80624	As_LIP38	88	As_120249	As_LIP88			
39	As_117390	As_LIP39	89	As_56677	As_LIP89			
40	As_17860	As_LIP40	90	As_120181	As_LIP90			
41	As_115808	As_LIP41	91	As_56843	As_LIP91			
42	As_96613	As_LIP42	92	As_7196	As_LIP92			
43	As_54516	As_LIP43	93	As_113700	As_LIP93			
44	As_40237	As_LIP44	94	As_99270	As_LIP94			
45	As_33701	As_LIP45	95	As_54682	As_LIP95			
46	As_78707	As_LIP46	96	As_7788	As_LIP96			
47	As_28374	As_LIP47	97	As_80236	As_LIP97			
48	As_17866	As_LIP48	98	As_6953	As_LIP98			
49	As_10888	As_LIP49	99	As_10928	As_LIP99			
50	As_56564	As_LIP50	100	As_50127	As_LIP10			

As 105 proteínas preditas como lipocalinas do RNA-seq de glândulas salivares de *A. sculptum* foram alinhadas, sendo que durante o alinhamento, utilizou-se o

aplicativo GBLOCKS (CASTRESANA, 2000) para eliminar as posições e regiões divergentes o que resultou numa deleção de 29 sequências. Procedeu-se a construção de uma árvore MB com as 75 sequências restantes o que resultou em uma árvore contendo cinco clados (I a V) (Figura 7).

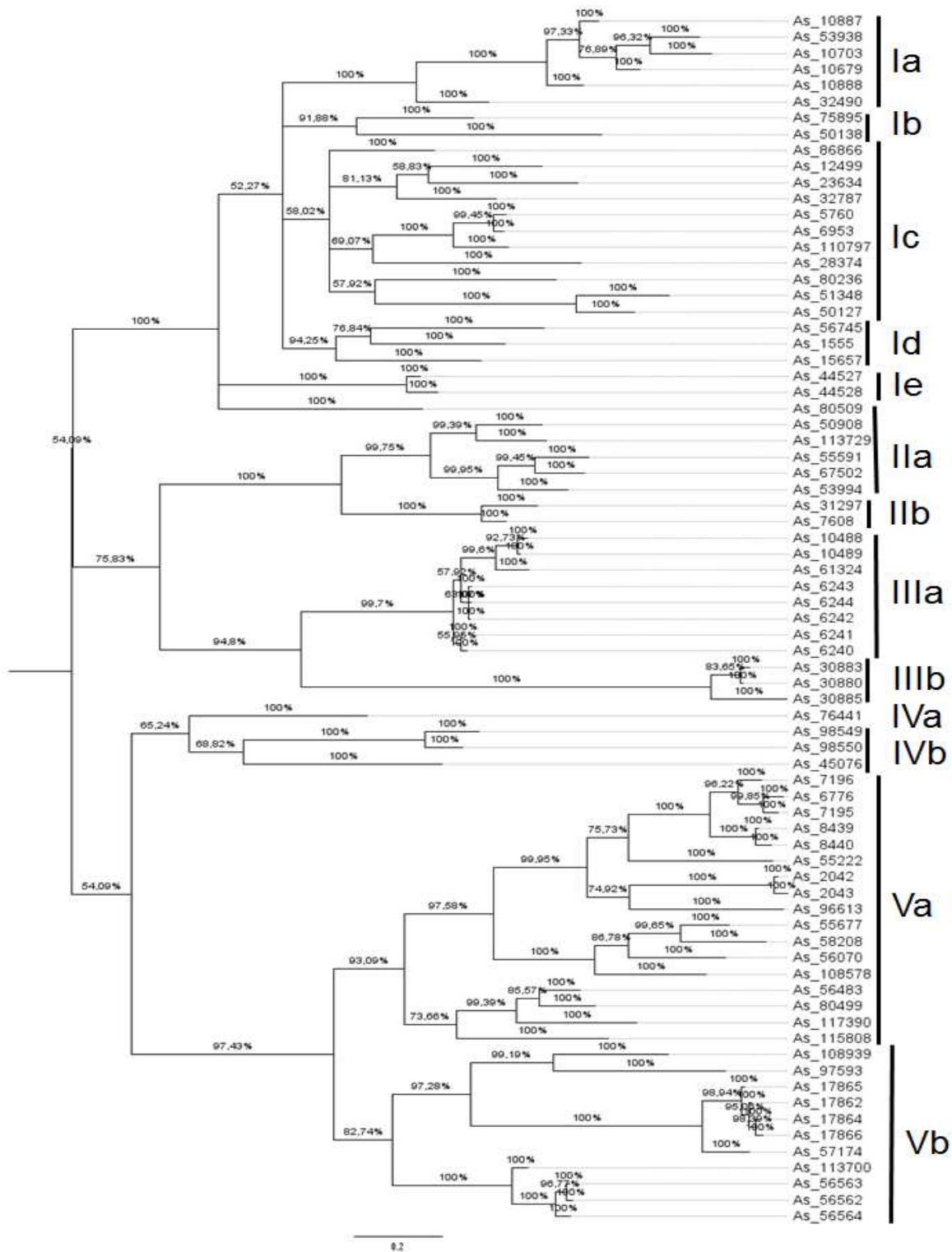


Figura 7: Árvore bayesiana, das 75 sequências de transcritos de lipocalinas de glândulas salivares de *A. sculptum*, resultante do alinhamento com deleção de gaps e normalização pelo aplicativo Gblocks (com deleção de 29 sequências divergentes). As letras As correspondem ao gênero e família (*Amblyomma sculptum*), os números correspondem aos códigos dos *contigs* do transcriptoma referido. Os algarismos romanos (I a V) indicam os clados filogenéticos obtidos e as letras de a até e indicam os subclados

obtidos. Os números nos nodos indicam a probabilidade em percentagem calculada pela análise bayesiana. A barra indica 20% de substituição por sítio.

A árvore bayesiana dos 105 transcritos de lipocalinas de glândulas salivares de *A. sculptum* contendo cinco cladros (I a V) foram comparadas com a análise para predição de domínios funcionais. Para discussão desse resultado, consideramos como *cut of* para domínios significativos um bit score > 30, independentemente do tamanho das mesmas. Assim, bit score elevado (>30) demonstra que os transcritos possuem forte relação com sua descrição correspondente, pertencendo à superfamília das calicinas, onde estão contidas as lipocalinas e possuindo domínio de ligação à histamina (SAWIER & RICHARDSON, 1991; FLOWER, NORTH, ATTWOOD, 1993; FLOWER, 1993; FLOWER, NORTH, SANSOM, 2000).

O clado I contém cinco subclados (Ia até Ie) todas as sequências pertencem à família *His Binding Protein* e à superfamília das calicinas (Tabela 2).

Tabela 2: Análise funcional do clado I dos transcritos de lipocalinas de glândulas salivares de *Amblyomma sculptum* submetida ao PFAM.

Sequencia	Clade	Family	Description	Entry type	Clan	HMM length	Bit score	E-value
As_10887	Ia	His-binding	Tick histamin biding protein	Domain	CL0116	152	29,6	4,9e-7
As_53938	Ia	His-binding	Tick histamin biding protein	Domain	CL0116	152	23,2	4,9e-5
As_10703	Ia	His-binding	Tick histamin biding protein	Domain	CL0116	152	21,6	1,5e-4
As_10679	Ia	His-binding	Tick histamin biding protein	Domain	CL0116	152	24,3	2,2e-5
As_10888	Ia	His-binding	Tick histamin biding protein	Domain	CL0116	152	29,2	4,2e-7
As_32490	Ia	His-binding	Tick histamin biding protein	Domain	CL0116	152	31,1	1,7e-7
As_75895	Ib	n/a ¹	n/a ¹	n/a ¹	Unk ²	n/a ¹	n/a ¹	n/a ¹
As_50138	Ib	His-binding	Tick histamin biding protein	Domain	CL0159	152	29,1	7,3e-7
As_86866	Ic	His-binding	Tick histamin biding protein	Domain	CL0116	152	25,4	1e-5
As_12499	Ic	n/a ¹	n/a ¹	n/a ¹	Unk ²	n/a ¹	n/a ¹	n/a ¹
As_23634	Ic	His-binding	Tick histamin biding protein	Domain	CL0116	152	17,7	2,4e-3
As_32787	Ic	His-binding	Tick histamin biding protein	Domain	CL0116	152	20,6	3e-4
As_5760	Ic	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹
As_6953	Ic	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹
As_110797	Ic	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹
As_28374	Ic	His-binding	Tick histamin biding protein	Domain	CL0116	152	33,4	3,5e-8
As_80236	Ic	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹
As_51348	Ic	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹
As_50127	Ic	His-binding	Tick histamin biding protein	Domain	CL0116	152	24,4	2e-5
As_56745	Id	His-binding	Tick histamin biding protein	Domain	CL0116	152	36,0	5,3e-9
As_1555	Id	His-binding	Tick histamin biding protein	Domain	CL0116	152	25,9	7,1e-6
As_15657	Id	His-binding	Tick histamin biding protein	Domain	CL0116	152	30,1	3,5e-7
As_44527	Ie	His-binding	Tick histamin biding protein	Domain	CL0116	152	35,6	7e-9
As_44528	Ie	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹	n/a ¹

Obs: ¹n/a = no available; ²Unk=Unknown.

Na superfamília das calicinas verificou-se a presença de domínio de ligação à histamina. Todos os transcritos do subclado Ia (As_10887, As_53938, As_10703, As_10679, As_10888, As_32490) possuem o domínio das calicinas (CL0116), com cálice hidrofóbico que caracteriza a superfamília com domínio de ligação a histamina (HBP), conforme descrito por Sawier & Richardson (1991). Durante o repasto sanguíneo do carrapato, apesar de existirem receptores cutâneos de histamina no hospedeiro, verifica-se uma diminuição da reação cutânea inflamatória prolongando a

alimentação do artrópode, o que pode ser interpretado como parte das estratégias desenvolvidas pelo artrópode no intuito de suprimir a resposta do hospedeiro.

Caso contrário, a presença de histamina e ligação desta ao receptor específico provocaria, de maneira direta ou indireta, uma diminuição da alimentação por ativação da resposta inflamatória, com conseqüente redução do êxito reprodutivo desse parasito, visto a fêmea depender do repasto sanguíneo para seu desenvolvimento (KEMP & BOURNE, 1980; WIKEL, 1982).

Neste contexto, na saliva do carrapato *R. appendiculatus* foi identificada uma lipocalina que possui a capacidade de ligação a histamina, a *Ripicephalus appendiculatus histamin binding protein* (RaHBP). Esta compõe-se de 171 aminoácidos e possui dois sítios distintos de ligação de histamina. O primeiro, situado em posição análoga à de outras lipocalinas, ou seja, na extremidade aberta do barril beta, possuindo uma afinidade de ligação de histamina de aproximadamente 2 nM (PAESEN, 1999); com o segundo possuindo baixa afinidade e localizando-se na extremidade oposta do barril β , fato que a diferencia das demais considerando-se que geralmente essa porção encontra-se fechada pela hélice 3¹⁰ na porção N-terminal. Assim, essa RaHBP foi classificada como RaHBP2, devido à ausência da hélice citada.

Também como exemplo em sequências de lipocalinas referentes ao clado I têm-se a rEV131 (Evolutec, Reino Unido), uma proteína recombinante anti-inflamatória, cujo efeito é tão eficaz quanto às drogas de escolha atualmente utilizadas no tratamento da asma humana. Tendo como provável mecanismo de ação a diminuição da ativação e recrutamento de neutrófilos reduzindo desta maneira o bronco espasmo. Estudos clínicos com esta molécula ultrapassaram as fases I e II do desenvolvimento de fármacos, demonstrando segurança e atividade farmacológica promissoras no tratamento de alergia humana, vermelhidão conjutival, espirros e corrimento nasal (COUILLIN et al., 2004; SCHLEHUBER & SKERRA, 2005).

No subclado Ib identificamos a presença do domínio HBP na sequência As_50138. As análises do subclado Ic retornaram com 11 sequências, cinco (As_50138, As_86866, As_23634, As_32787, As_28374 e As_50127) com presença do domínio HBP e as demais (As_12499, As_5760, As_6953, As_110797, As_80236 e As_52348) possuindo domínio desconhecido. O *bit score* das sequências que apresentaram domínio HBP foi próximo a 20, indicando também pouca significância. No subclado Id identificamos três sequências (As_56745, As_1555, As_15657) com a presença do domínio HBP. As análises do subclado le retornaram com duas

sequências (As_44527 e As_44528) sendo que a sequência As_44527, possui o domínio HBP e a As_44528 apresentou domínio desconhecido.

A tabela 3 apresenta os resultados da análise do PFAM das lipocalinas presentes nos clados II, III e IV e seus subclados correspondentes.

Tabela 3: Análise funcional dos clados II, III e IV dos transcritos de lipocalinas de glândulas salivares de *Amblyomma sculptum* submetida ao PFAM.

Sequence	Class	Family	Description	Entry type	Clan	HMM length	Bit score	E-value
As_50908	Ila	His-binding	Tick histamin biding protein	Domain	CL0116	152	50,1	2,5e-13
As_113729	Ila	His-binding	Tick histamin biding protein	Domain	CL0116	152	42,1	7,2e-11
As_55591	Ila	His-binding	Tick histamin biding protein	Domain	CL0116	152	30,2	3,2e-7
As_67502	Ila	His-binding	Tick histamin biding protein	Domain	CL0116	152	18	1,9e-3
As_53994	Ila	His-binding	Tick histamin biding protein	Domain	CL0116	152	38,2	1,1e-9
As_31297	Ilb	His-binding	Tick histamin biding protein	Domain	CL0116	152	45,2	7,7e-12
As_7608	Ilb	His-binding	Tick histamin biding protein	Domain	CL0159	152	45,9	4,8e-12
As_10488	IIIa	His-binding	Tick histamin biding protein	Domain	CL0116	152	122,2	1,6e-35
As_10489	IIIa	His-binding	Tick histamin biding protein	Domain	CL0116	152	121,2	3,2e-35
As_61324	IIIa	His-binding	Tick histamin biding protein	Domain	CL0116	152	119,5	1e-34
As_6243	IIIa	His-binding	Tick histamin biding protein	Domain	CL0116	152	117,7	3,7e-34
As_6244	IIIa	His-binding	Tick histamin biding protein	Domain	CL0116	152	118,3	2,4e-34
As_6242	IIIa	His-binding	Tick histamin biding protein	Domain	CL0116	152	117,5	4,4e-34
As_6241	IIIa	His-binding	Tick histamin biding protein	Domain	CL0116	152	118,9	1,6e-34
As_6240	IIIa	His-binding	Tick histamin biding protein	Domain	CL0116	152	119,4	1,1e-34
As_30883	IIIb	His-binding	Tick histamin biding protein	Domain	CL0116	152	37,5	1,8e-9
As_30880	IIIb	His-binding	Tick histamin biding protein	Domain	CL0116	152	37,4	2,1e-9
As_30885	IIIb	His-binding	Tick histamin biding protein	Domain	CL0116	152	30,8	2,1e-7
As_76441	IVa	His-binding	Tick histamin biding protein	Domain	CL0116	152	19,0	9,3e-4
As_98549	IVb	His-binding	Tick histamin biding protein	Domain	CL0116	152	26,8	3,8e-6
As_98550	IVb	His-binding	Tick histamin biding protein	Domain	CL0116	152	21,6	1,4e-4
As_45076	IVb	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Obs: ¹n/a = no available; ²Unk=Unknown.

No clado II identificamos dois subclados (Ila e Ilb) totalizando sete sequências, sendo as sequências As_50908, As_113729, As_55591, As_67502 e As_53394 pertencentes ao subclado Ila, e as sequências As_31297 e As_7608 pertencentes ao subclado Ilb. Com o auxílio do PFAM verificamos que todas as sequências possuem domínio HBP e pertencem a superfamília das calicinas (CL0116).

No clado III identificamos dois subclados (IIIa e IIIb) totalizando 11 sequências, sendo As_10488, As_10489, As_61324, As_6243, As_6244, As_6242, As_6241 e As_6240 pertencentes ao subclado IIIa, e as sequências As_30883, As_30880 e As_3885 pertencentes ao subclado IIIb. Todas as sequências retornaram como pertencentes à família HBP, tendo como domínio significativo o das calicinas, pertencendo à superfamília das calicinas (CL0116) (Tabela 3). Todas as sequências do subgrupo IIIa apresentaram um *bit score* na faixa de 117 a 122, discrepando das demais pertencentes ao mesmo clado cujo *bit score* máximo foi de 50,1. Todas as sequências do subclado IIIa foram identificadas, *in silico*, neste trabalho como possuidoras do domínio de ligação a serotonina e histamina, SHBP *domain*.

O clado IV também apresentou dois subclados (IVa e IVb), totalizando quatro sequências, As_76441, As_98549, As_98550 e As_45076. A sequência As_45076 retornou com classificação desconhecida (Tabela 3). A sequência As_61324, As_6244 e As_30880, apresentaram domínio HBP. E ainda, a sequência As_45076 não foi classificada pelos parâmetros avaliados pelo PFAM.

A tabela 4 apresenta os resultados da análise do PFAM das lipocalinas presentes nos clados V e seus subclados correspondentes.

Tabela 4: Análise funcional do clado V dos transcritos de lipocalinas de glândulas salivares de *Amblyomma sculptum* submetida ao PFAM.

Sequence	Class	Family	Description	Entry type	Clan	HMM length	Bit score	E-value
As_7196	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_6776	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_7195	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_8439	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_8440	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_55222	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_2042	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_2043	Va	His-binding	Tick histamin biding protein	Domain	CL0116	152	17,2	3,3e-3
As_96613	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_55677	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_58208	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_56070	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_108578	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_56483	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_80499	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_117390	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_115808	Va	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_108939	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_97593	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_17865	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_17862	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_17864	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_17866	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_57174	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_113700-	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_56563	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_56562	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a
As_56564	Vb	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Obs: ¹n/a = no available; ²Unk=Unknown.

No clado V identificamos 28 sequências, com a análise do PFAM revelando que a sequência As_2043 apesar de apresentar domínio HBP não é significativo. Todas as demais 27 sequências apesar de significativas retornaram como desconhecidas (Tabela 4).

7.2 Predição *in silico* dos domínios SHBP

Para a predição, *in silico*, dos domínios SHBP das sequências proteicas de lipocalinas depositadas no *Swiss-prot* para posterior comparação com aquelas anotadas de glândula salivar de *A. sculptum*, foi realizado um alinhamento no aplicativo MAFFT versão 7.0, (mafft.cbrc.jp/alignment/server/) e CLUSTAL W (THOMPSON et al., 1994), em seguida o alinhamento foi transportado para o aplicativo MEGA versão 5.0

(KUMAR et al., 2008). Optou-se pelo critério de não deleção de *gaps* na tentativa de obter maior número de informações. Após o alinhamento foi construída uma árvore resultante da análise bayesiana, a mesma foi realizada no aplicativo Mr Bayes versão 3.1 (MB) (HUELSENBECK & RONQUIST, 2001; RONQUIST & HUELSENBECK, 2003). Para predição do melhor modelo de substituição de aminoácidos utilizou-se o aplicativo ProtTest 2.4.server (darwin.uvigo.es /software/ prottest2_server.html), (ABASCAL, 2005) (figura 8).

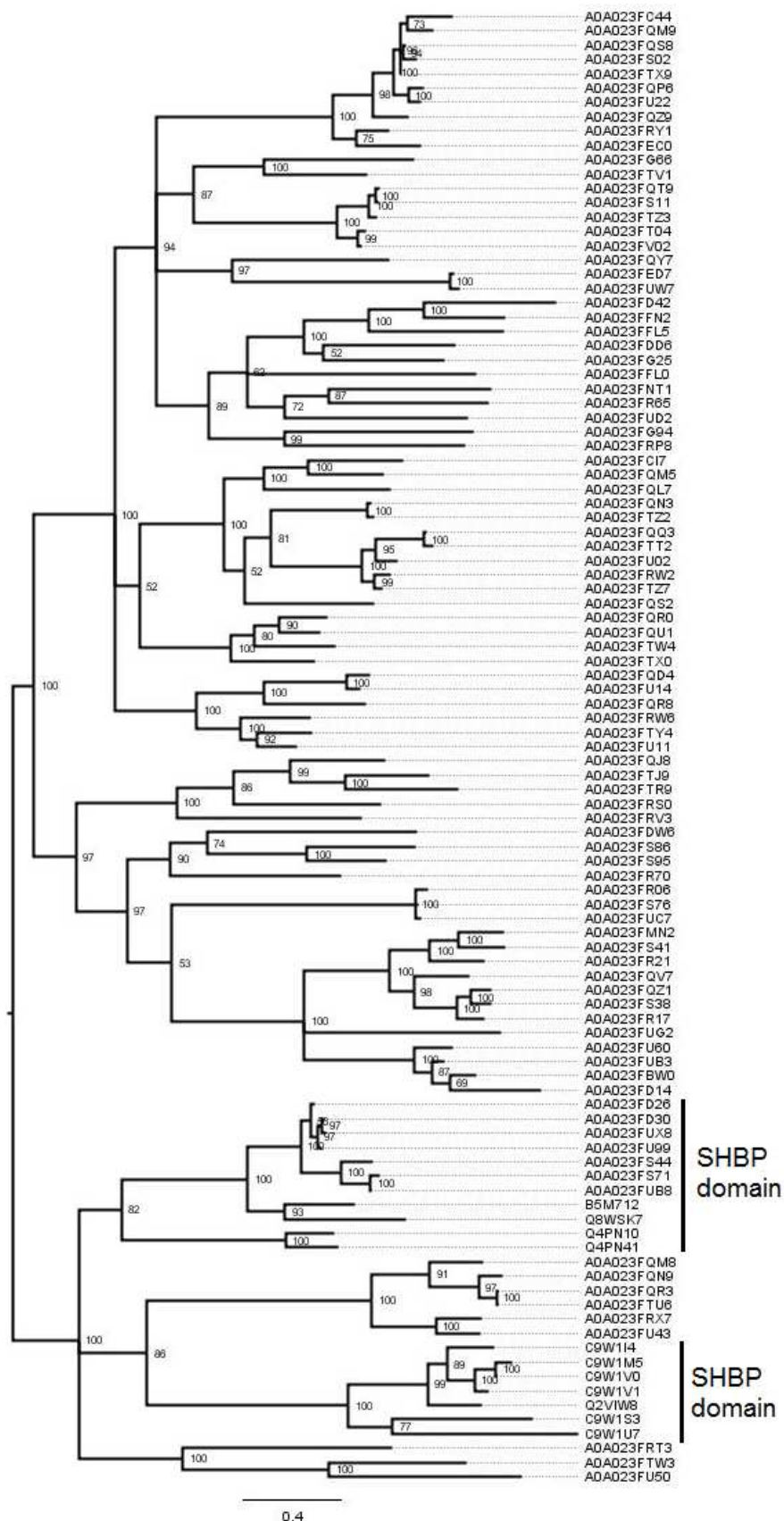


Figura 8: Árvore bayesiana, das 105 sequências de lipocalinas selecionadas no *Swiss-prot* advindas do transcriptoma das glândulas salivares de *Amblyomma cajennense* (GARCIA et al., 2014) confrontadas com sequências de lipocalinas de carrapato com SHBP *domain* depositadas no *Swiss-prot* resultante do alinhamento sem deleção de *gaps*. A barra lateral identifica as sequências que possuem SHBP *domain*.

Por análise bayesiana foram identificadas sete sequências de lipocalinas do *dataset* advindo do *Swiss-prot* (A0A023FS71, A0A023FUB8, A0A023FD30, A0A023FD26, A0A023FUX8, A0A023FS44 e A0A023FU99) agrupadas com as sequências Q8WSK7, SHBP de *Dermacentor appendiculatus* (SANGAMNATDEJ et al., 2002), B5M712, SHBP *A. americanum* (ALJAMALI, et al., 2009), Q4PN10, SHBP de *Ixodes scapularis* (MANS et al., 2008) e Q4PN4, SHBP de *Ixodes scapularis* (MANS et al., 2008) em detrimento das demais, sugerindo uma seleção *in silico* de sete sequências com SHBP *domain* e portanto com potencial atividade anti-histamínica. A Q8WSK7 é uma lipocalina do *Dermacentor reticulatus* uma proteína de ligação à histamina e serotonina (SHBP), que apresenta dois sítios de ligação. Em estudo semelhante, foram identificados os mesmos sítios de ligação para histamina e/ou serotonina em *Argas reflexus*, *A. monolakensis* e *Ornithodoros savignyi* (MANS, RIBEIRO & ANDERSEN, 2008).

7.3 Refino da seleção de transcritos de lipocalinas de *Amblyomma sculptum* com domínio SHBP

O alinhamento das sete lipocalinas selecionadas apresentou 7 sítios conservados, reforçando a ideia de que as sete lipocalinas selecionadas possuem o domínio *SHBP* (Figura 9).

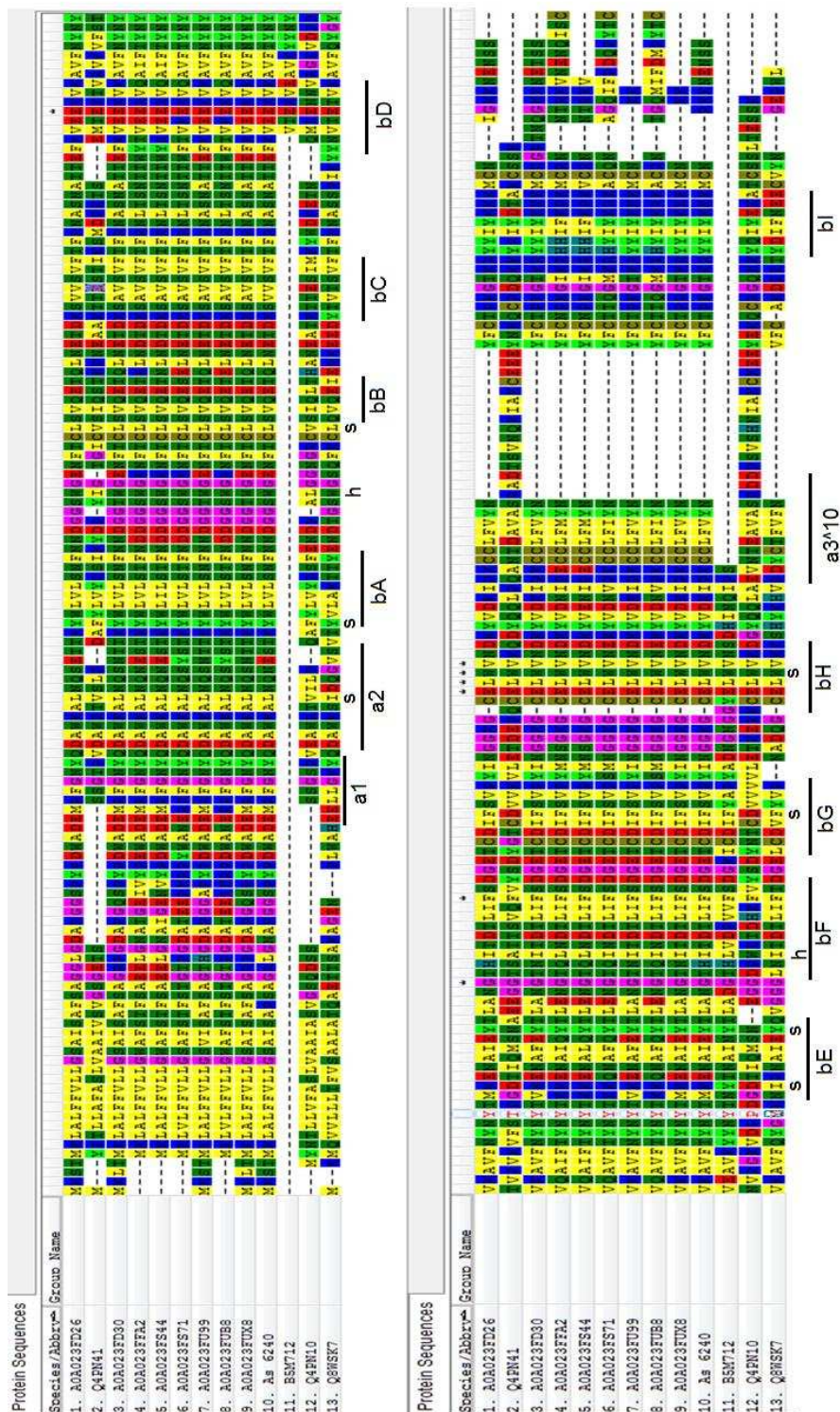


Figura 9: Alinhamento entre as sete seqüências com *SHBP domain* selecionadas do *Swiss-prot* que se agruparam na figura 3 somadas ao transcrito *As_6240*, não depositado, proveniente do RNA-seq. de glândulas salivares de *A. sculptum sculptum* e confrontadas com as seqüências de lipocalinas de *A. cajennense* depositadas no UniProt (GARCIA, 2014). Os asteriscos (*) correspondem aos sítios conservados. As letras h e s indicam a posição dos sítios de ligação para histamina e serotonina, respectivamente. As estruturas secundárias estão representadas sendo **a** (1 -3), representado as α -hélices, a α -hélice 3^{10} está representada como $a3^{10}$ e **b** (A-I) representando as folhas β presentes. As estruturas secundárias foram previstas de acordo com Q8WSK7, SHBP de *Dermacentor appendiculatus* (SANGAMNATDEJ et al., 2002).

Nossos resultados reforçam a hipótese de uma possível conservação evolutiva da função anti-histamínica por meio da participação de lipocalinas na modulação da resposta do hospedeiro. Estas, durante o repasto sanguíneo do carrapato, apesar da existência de receptores cutâneos de histamina no hospedeiro vertebrado, atuam reduzindo a reação cutânea inflamatória, prolongando o período de alimentação do artrópode.

Na presença de histamina e ligação desta ao receptor específico teríamos de maneira direta ou indireta, uma diminuição do período de alimentação, com consequente redução do êxito reprodutivo desse parasito, visto a fêmea depender do repasto sanguíneo para seu desenvolvimento (KEMP & BOURNE, 1980; WIKEL, 1982). Isto vem a demonstrar o desenvolvimento de estratégias destinadas a suprimir a resposta do hospedeiro o que, de acordo com os dados demonstrados, não é diferente em carrapatos *A. sculptum*.

Estudos experimentais para determinar a atividade anti-histamínica apresentada pela saliva do carrapato *A. sculptum*, bem como, a investigação de uma possível interação desta atividade com receptores celulares associados com o padrão de conservação de sua estrutura terciária, constituem uma sólida base para o desenvolvimento de biofármacos. Neste contexto, potencialmente a ação anti-histamínica apresentada pelas lipocalinas do *A. sculptum*, associadas à sua capacidade de abrigar substâncias hidrofóbicas no interior de sua estrutura terciária, tornam estas moléculas candidatas a produto de uso farmacêutico.

7.4 Resultado da extração do RNA total das Glândulas salivares, intestinos e ovários do carrapato *A. sculptum*

Por análise espectrofotométrica avaliamos a concentração do RNA total de glândulas salivares, intestino e ovário armazenados a -80°C . Para tanto, utilizamos a absorbância em 260 e 280 nm, para avaliar a relação 260/280, bem como a concentração das amostras. A integridade das mesmas foi avaliada por eletroforese em gel de agarose a 1% corada com brometo de etídeo (figura 10).

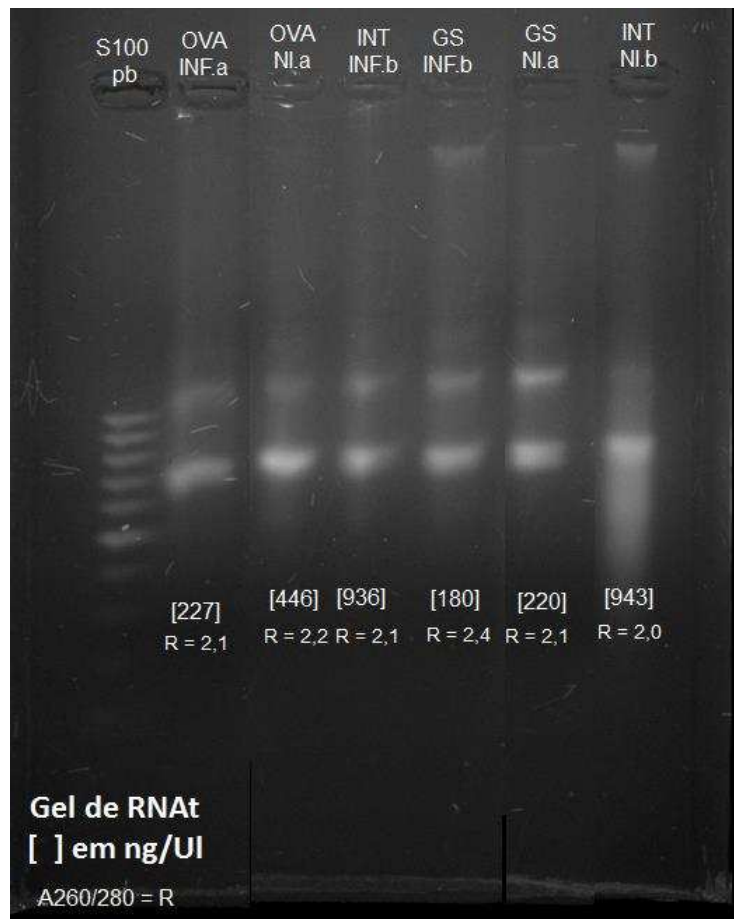


Figura 10: Eletroforese em gel de agarose 1% de amostras de RNAt (dois *pulls* de 75 glândula salivar, intestino e ovário) provenientes de transcriptomas de glândulas salivares, intestinos e ovários do carrapato *Amblyomma sculptum* armazenados a -80°C. As letras correspondem a: GS NI – glândulas salivares não infectadas; GS INF – glândulas salivares infectadas; INT NI – intestino não infectado; INT INF – intestino infectado; OVA NI – ovário não infectado; OVA INF – ovário infectado. Entre colchetes está a concentração de RNA em ng/uL, R é a relação A260/A280, advindos da análise espectrofotométrica no aparelho espectrofotômetro NanoDrop 2000 (*ThermoScientific*®), tendo como ideal um resultado na faixa de 1,8 a 2,2.

7.5 Análise e validação dos *primers* por testes de concentração e eficiência para uso em RTq-PCR

Através do aplicativo *Primer Express 3.0* (*Applied Biosystems*) foram desenhados nove *primers* de lipocalinas que apresentaram, *in silico*, domínio de ligação para serotonina e histamina (*SHBP domain*) (*lip1*(As-10488/A0A023FS71), *lip4* (As-10489/A0A023FUB8), *lip7* (As-61324/A0A023FFA2), *lip10* (As-6242/A0A023FD30), *lip14* (As-6241/A0A023FD26), *lip17* (As-6240/Não depositado), *lip21* (As-6244/A0A023FUX8), *lip24* (As-61325/A0A023FS44) e *lip27* (As-6243/A0A023FU99), para nove genes de lipocalinas (A0A023FS71, A0A023FUB8, A0A023FFA2, A0A023FD30, A0A023FD26, As_6240, A0A023FUX8, A0A023FS44 e A0A023FU99) (Tabela 5).

Tabela 5: Característica dos *primers* para amplificação de fragmentos de lipocalinas com presença de SHBP *domain* desenhados pelo aplicativo *Primer express 3.0* para aplicação em qPCR.

Gene	Code Transcriptome/ Primer name/Swiss-prot	Primer Sequence (5'-3')	Size ^a (Bp)	Tm ^b (°C)	CG (%)	MW	ΔG (25°C)
lip 1	As-10488/ LIP1 A0A023FS71	Foward AGTGATGGAGAGACGTGTGATAT TTT	72	56.0	38.5	8,120.3	1.77
		Reverse CACCCAAAGCTCACACCCTTT	72	58.0	52.4	6,255.5	0.29
lip 4	As-10489/ LIP4 A0A023FUB8	Foward TCTTCTAGGTTGGGCGTTCTCT	81	57.6	50.0	6,714.4	1.02
		Reverse GCCCAGTCAGGACGTTTACG	81	58.2	60.0	6,118.0	-1.15
lip 7	As-61324/ LIP7 A0A023FFA2	Foward TGGCGAGACGTGCGATATT	68	56.7	52.6	5,888.9	-0.46
		Reverse TCACCCAGAGTTCACACCCTTT	66	58.4	50.0	6,590.3	0.75
lip 10	As-6242/ LIP10 A0A023FD30	Foward CTACTGCTACTGTTGTGATGTTT GGTT	74	57.2	40.7	8,293.4	-1.12
		Reverse GTACCGCCGTCATTGTTGAA	74	55.6	50.0	6,108.0	0.14
lip 14	As-6241/ LIP14 A0A023FD26	Foward TGGCGAGACATGCGACAT	67	57.1	55.6	5,548.7	-0.50
		Reverse CACCCAGAGTTCACATCCTTTG	67	55.8	50.0	6,630.4	0.30
lip 17	As-6240/ LIP17 Não depositado	Foward TGGCGAGACATGCGACAT	66	57.1	55.6	5,548.7	-0.50
		Reverse CACCCAGAGTTCACATCCTTTG	66	55.8	50.0	6,630.4	0.30
lip 21	As-6244/ LIP21 A0A023FUX8	Foward TGGCGAGACATGCGACAT	69	57.1	55.6	5,548.7	-0.50
		Reverse TTTACCCAGAGTTCACATCCTTT G	69	55.0	41.7	7,253.8	0.27
lip 24	As-61325/ LIP24 A0A023FS44	Foward TGGCGAGACGTGCGATATT	68	56.7	52.6	5,883.9	-0.46
		Reverse TCACCCAGAGTTCACACCCTTTG	68	59.1	52.2	6,919.5	0.44
lip 27	As-6243/ LIP27 A0A023FU99	Foward TGGCGAGACATGCGACAT	69	57.1	55.6	5,548.7	-0.50
		Reverse TTTACCCAGAGTTCACATCCTTT G	69	55.0	41.7	7,253.8	0.27

^atamanho do *amplicon*; ^btemperatura de *melting* para 50 nM NaCl; MW, peso molecular; os parâmetros dos *primers* desenhados foram fornecidos pelo programa *primer express 3.0*. As letras As seguidas de 4 a 6 números provêm das sequências do transcriptoma de *A. sculptum*.

Adicionalmente, através do aplicativo *Primer Express 3.0* (*Applied Biosystems*) foram desenhados seis pares de *primers* (*hk1/hk2*, *hk3/hk4* e *hk5/hk6*) para três genes constitutivos potenciais candidatos a genes de referência (A0A023FLV9, *glicose-6-fosfato-desidrogenase*, *g6pdh*; A0A023FJ24, *3-hidroxi-Coenzima-A-desidrogenase*, *3ohcoadh*; A0A023FLU5, *gliceraldeído-3-fosfato-desidrogenase*, *gapdh*) (Tabela 6).

Tabela 6: Característica dos *primers* para amplificação de fragmentos de lipocalinas desenhados pelo aplicativo *Primer express 3.0* para os candidatos a genes de referência para aplicação em qPCR.

Gene	Code Transcriptome/ Primer name/Swiss- <i>prot</i>	Primer Sequence (5'-3')	Size ^a (Bp)	Tm ^b (°C)	CG (%)	MW	ΔG (25°C)	
<i>glicose-6-fosfato-desidrogenase (g6pdh)</i>	As-17526/ HK1 A0A023FLV9	Foward TGGTGAGGCCATCTATG TCAA	68	56.1	45.5	6,774.5.2	-1.08	
		Reverse CCGTCTCCTCAATGTGCG AAAG	68	55.4	52.4	6,366.2	-0.25	
	As-17526/ HK2 A0A023FLV9	Foward CCACGGTGCCTCCAAAC T	70	58.1	61.1	5,404.6	- 0.85	
		Reverse CACCGCTCGTTGTTGAT GTAGA	70	57.2	50.0	6,741.4	+ 1.01	
	3-hidroxi-Coenzima-A-desidrogenase (<i>3ohcoadh</i>)	As-50837/ HK3 A0A023FJ24	Foward TTTTGCGGCACTTGACAA AG	66	54.6	45.0	6,132.0	-0.43
			Reverse GGCAGAGAGGAGGTGTT GCT	66	59.5	60.0	6,278.1	-0.93
As-50837/ HK4 A0A023FJ24		Foward TGCAAGCAACACCTCCT CTCT	71	58.7	52.4	6,286.1	+0.67	
		Reverse CACCAAAACGATCCTTC CTTGA	71	55.3	45.5	6,623.4	+0.36	
gliceraldeído-3-fosfato-desidrogenase (<i>gapdh</i>)	As-111072/ HK5 A0A023FLU5	Foward GCCTTTGCCGACCACTT TTA	61	56.0	50	6,018.9	+1.2	
		Reverse TGCCGCAGCTCTCAAAG AA	61	57.3	52.6	5,781.8	+1.6	
	As-111072/ HK6 A0A023FLU5	Foward TCGTCCTGCCACACAAG TTC	66	57.7	55.0	6,012.9	+1.3	
		Reverse CACCGCTTGATCTTTC CAA	66	55.4	47.6	6,341.0	+1.1	

^a tamanho do *amplicon*; ^b temperatura de *melting* para 50 nM NaCl; As letras *As* seguido de 4 a 6 números provêm das sequências do transcriptoma de *A. sculptum* (Moreira, 2012). *Gli6pdh*, Glicose-6-Phosphate-Desidrogenase; *3OHCADH*, 3-Hidroxi-Coenzima-A-Desidrogenase *G3pdh* Glicerol-3-Phosphate-Desidrogenase; MW, molecular weight. Os parâmetros dos primers foram fornecidos pelo programa *primer express 3.0*.

Os testes de concentração e eficiência dos *primers* demonstraram a viabilidade do *Lip 4*, *Lip 7* e *Lip 24* e os *HK 2 (G6PDH)*, *HK 4 (3OHCADH)* e *HK 6 (GAPDH)* (Tabela 7).

Tabela 7: Resultados das curvas de eficiência dos genes de lipocalinas selecionados na etapa de determinação da concentração e curva de eficiência dos genes candidatos a genes de referência.

<i>Target</i>	<i>slope</i>	Intercepto (y)	R ²	Eff	Conc. (uM)	Melt (°C)
Lip4	-3,859	37,568	0,987	0,81	0,6	79,63
Lip7	-4,414	40,22	0,982	0,68	1,0	78,43
Lip24	-3,27	38,014	0,949	1,02	0,8	78,88
Hk1	-1,318	30,014	0,866	4,73	0,8	78,28
Hk2	-2,342	31,951	0,963	1,67	0,8	81,86
Hk4	-1,47	29,186	0,937	3,78	0,8	78,58
Hk5	-2,754	32,262	0,978	1,30	0,8	80,51
Hk6	-4,29	35,76	0,998	0,71	0,8	77,24

Dos nove *primers* de lipocalinas desenhados apenas três *lip 4*, *lip 7* e *lip 24* obtiveram sucesso na etapa de concentração e eficiência obtenção dos *primers* testados. Esse fato demonstra a necessidade da execução de etapas preliminares antes da análise da expressão do experimento em condições biológicas diferentes (p. ex: tratado contra não tratado, infectado contra não infectado) evitando assim conclusões equivocadas quanto expressão de proteínas em determinados órgãos ou tecidos. Os testes de concentração otimizaram a concentração dos *primers lip 4*, *7* e *24*, para utilização em nosso experimento resultando em 0,6 uM; 1,0 uM e 0,8 uM respectivamente para *lip 4*, *7* e *24*. Em relação a etapa de determinação da concentração e eficiência para os genes de referência foram selecionados os *primers hk 2*, *4* e *6* com a concentração ótima de 0,8 uM para todos os genes testados. Os *primers hk2*, *4* e *6* correspondem às sequências depositadas (A0A023FLV9, *glucose-6-fosfato-desidrogenase*, *g6pdh*; A0A023FJ24, *3-hidroxi-coenzima-a-desidrogenase*, *3ohcoadh*; A0A023FLU5, *gliceraldeído-3-fosfato-desidrogenase*, *gapdh*), respectivamente.

Assim, de acordo com os resultados do teste de concentração e eficiência, optou-se pela continuidade dos experimentos a utilização dos *primers* de lipocalinas *lip4*, *lip7* e *lip24* e a viabilidade de uso dos *primers* de gene de referência, *hk2*, *hk4* e *hk6*. Dessa forma, foram escolhidos um *primer* de cada gene de referência a ser estudado *hk2* representando *g6pdh*, *hk4* representando *3-ohcoadh* e o *hk6* representando o gene *gapdh*. Sendo esse último selecionado como controle endógeno a ser utilizado em nosso experimento.

O estudo por *real time* RTq-PCR com *primers* não validados ou validados, somente por análise *in silico*, através dos aplicativos para desenho (p. ex: *Primer Express*®), não refletem a realidade da eficiência destes quando confrontados com seus alvos, sendo necessários requisitos mínimos de checagem antes da amplificação

do experimento em si, corroborando com a ideia de que o teste da concentração dos *primers* desenhados e análise da eficiência destes deve ser rotina para o uso dessa ferramenta. Como demonstrado em nosso trabalho apesar de apresentar ótimos parâmetros para amplificação *in silico* (tabela 7.3.1), dos nove *primers* desenhados para lipocalinas somente três funcionaram satisfatoriamente (BUSTIN et al., 2009). Como as amostras de extração do RNA total e síntese de cDNA foram checadas quanto a integridade e concentração a falha recai sobre a qualidade dos *primers* sintetizado para amplificação das lipocalinas. Os resultados fornecidos pelo aplicativo *StepOne* versão 2.3 (*Liife Technologies*) relacionado a essa etapa encontra-se como material suplementar na seção de Anexos desse trabalho (Figuras S1 a S8).

7.5.1 Estudo dos genes alvo e dos transcritos selecionados para a amplificação e identificação de lipocalinas com domínio SHBP para uso em RTq-PCR

No tópico 7.6 foram selecionados 3 genes (*lip* 4, 7 e 24) para a amplificação dos fragmentos de lipocalinas, esse tópico apresentará a descrição desses genes-alvo, análise dos *primers* construídos para a amplificação dos mesmos, bem como análise estrutural das proteínas referente aos transcritos proveniente dos genes *lip* 4, 7 e 24.

O gene *lip4* está depositado no *European nucleotide archive* (ENA) (<http://www.ebi.ac.uk/ena/data/view/GBBK01000222>), sob o código GBBK01000222, possui 663 bp e seu transcrito possui 220 aminoácidos. O transcrito está depositado no *uniprot* sob o código A0A023FUB8. O gene *lip7* está depositado no *European nucleotide archive* (ENA) (<http://www.ebi.ac.uk/ena/data/view/GBBK010047221>), sob o código GBBK010047221, possui 663 bp e seu transcrito possui 220 aminoácidos. O transcrito está depositado no *uniprot* sob o código A0A023FFA2. O gene *lip24* está depositado no *European nucleotide archive* (ENA) (<http://www.ebi.ac.uk/ena/data/view/GBBK01000248>), sob o código GBBK01000248, possui 639 bp e seu transcrito possui 214 aminoácidos. O transcrito está depositado no *uniprot* sob o código A0A023FS44 (GARCIA et al., 2014).

7.5.2 Predição do modelo estrutural e das estruturas secundárias das lipocalinas 4, 7 e 24 com domínio SHBP por aplicativos de bioinformática

A Figura 11 apresenta os resultados da predição e avaliação de modelos estruturais da proteína LIP4 preditos pelo aplicativo *RaptorX*.

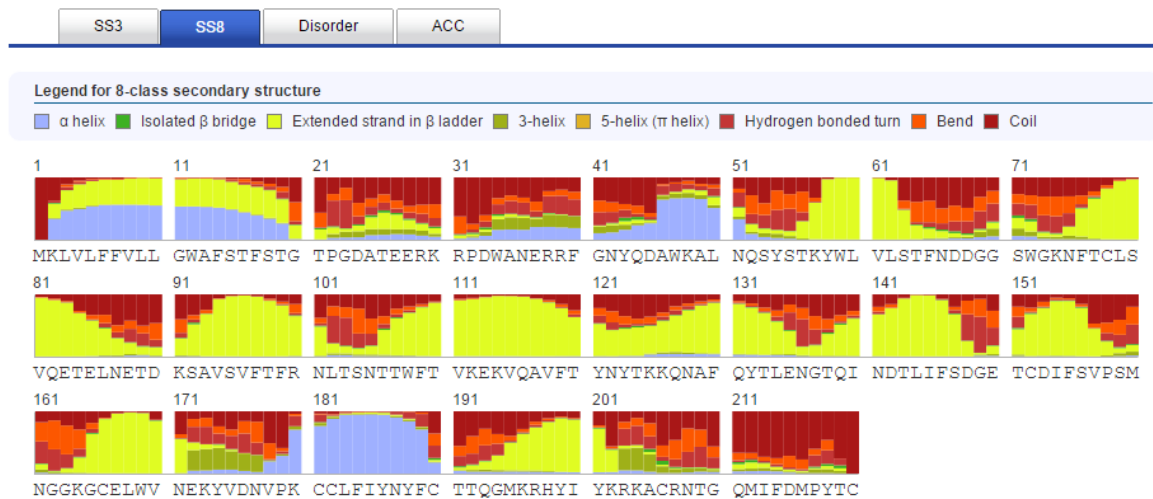


•Estruturas secundárias: 13%H, 35%E, 51%C

a)



b)



c)

Figura 11: Distribuição gráfica da participação dos resíduos na formação das estruturas secundárias compoendo as α hélices, folhas β e *loop* da Lipocalina 4 previstas pelo aplicativo *RaptorX*. Em (a) modelo 3D previsto pelo aplicativo *RaptorX*. Em (b) ocorre a distribuição principal em 3 classes de resíduos (hélices, folhas e *loop*). Em (c) detalhamento da predição estrutural das estruturas secundárias em 8 classes (α hélices, pontes β isoladas, prolongamento de cadeias β , hélice-3, hélice-5, interações de hidrogênio, dobramentos e *loop*). Os números em (b) e (c) representam o número de resíduos de aminoácidos (aa) e as letras representam os resíduos de aa da sequência de lipocalina 4. H: hélice; E: folhas β ; C: estruturas em espiral (*loop*).

A figura 11 (a) apresenta o modelo 3D predito pelo aplicativo RaptorX contendo as porcentagens correspondentes à estrutura secundária da proteína LIP4. A mesma revela que 13% dos resíduos compõe as α hélices (letra H figura 11 (a)), 35% compõe as folhas β (letra E figura 11 (a)) e 51% dos resíduos compõem os *loops* (letra C figura 10 (a)). As folhas β apresentadas no modelo 3D constituem a estrutura em β barril.

A figura 11 (b) resultou no reconhecimento da assinatura de lipocalinas da sequência submetida uma vez que a predição da estrutura secundária realizada pelo aplicativo *RaptorX* revelou a presença de 3 α hélices (duas comuns e uma menor a hélice 3¹⁰) e 8 folhas β revelando a assinatura estrutural de lipocalinas que necessita de pelo menos 8 folhas β (com topologia classificada de A-H) para formação do barril β que corresponderá ao “cálice” para abrigar lipídeos e ainda a presença de 3 α hélices sendo uma menor denominada de hélice-3 e a presença de *loops* característicos para a estabilização das folhas β (FLOWER et al., 2000; SKERRA et al., 2000).

Detalhadamente a predição da estrutura secundária feita pelo aplicativo *RaptorX* representada pela figura 11 (b) revelou a presença de 3 α hélices: a primeira a partir do resíduo de lisina (K) na posição 2 até o resíduo de treonina (T) na posição 18; a segunda iniciando no resíduo de triptofano (W) na posição 34 até o resíduo de asparagina (N) na posição 51; e a terceira, iniciando no resíduo de glutamato (E) na posição 171 até o resíduo de fenilalanina (F) na posição 189. Quanto a folhas β responsáveis pela formação da estrutura de barril β que corresponderá ao cálice foram preditas 8 folhas: a primeira iniciando no resíduo de treonina (T) na posição 56 até o resíduo de aspartato (D) na posição 68; a segunda iniciando no resíduo de asparagina (N) na posição 75 até o resíduo de aspartato (D) na posição 90; a terceira iniciando no resíduo de lisina (K) na posição 91 até o resíduo de leucina (L) na posição 102; a quarta iniciando no resíduo de treonina (T) na posição 106 até o resíduo de glutamato (E) na posição 135; a quinta iniciando no resíduo de glicina (G) na posição 138 até o resíduo de aspartato (D) na posição 148; a sexta iniciando no resíduo de glutamato (E) na posição 149 até o resíduo de glicina (G) na posição 162; a sétima iniciando no resíduo de lisina (K) na posição 164 até o resíduo de asparagina (N) na posição 171 e finalmente a oitava iniciando no resíduo de glicina (G) na posição 191 até o resíduo de lisina (K) na posição 202.

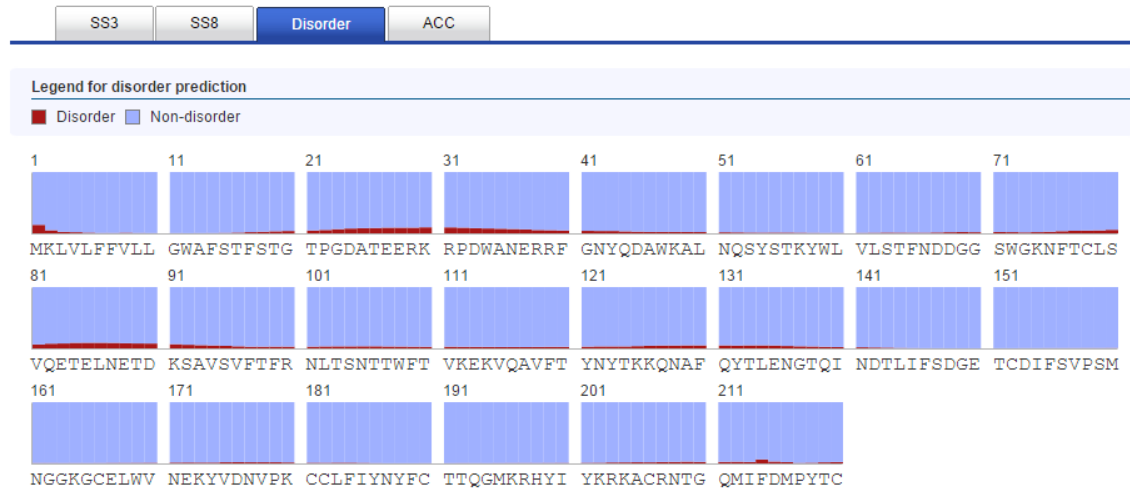
A figura 11 (c) demonstra a presença da hélice 3¹⁰ iniciando no resíduo de glutamato (E) na posição 171 até o resíduo de asparagina (N) na posição 177. E ainda o detalhamento da estrutura secundária evidencia a presença das 3 α hélices descritas acima.

Para a análise da acessibilidade do solvente (parâmetro ACC no aplicativo utilizado) e desordem do modelo 3D predito, o *RaptorX* utiliza uma escala de 3 estados: valores menores que 10% significa que não há acesso do solvente, maiores que 42% significa que a estrutura permite acessibilidade ao solvente e valores médios na faixa 10% a 42%, correspondendo a uma relativa acessibilidade. A Figura 12 apresenta os resultados da predição da desordem e acessibilidade ao solvente dos modelos estruturais da proteína LIP4 predito pelo aplicativo *RaptorX*.

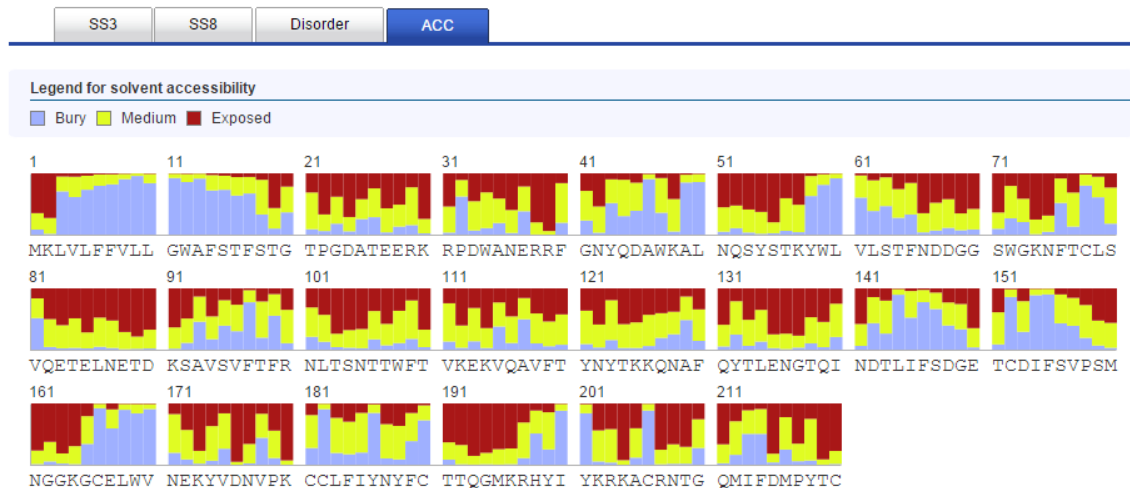


•Acesso ao solvente: 33%E, 40%M, 25%B

a)



b)



c)

Figura 12: Distribuição gráfica da participação dos resíduos de aminoácidos da Lipocalina 4 na predição dos parâmetros de desordem e acessibilidade do solvente previstas pelo aplicativo *RaptorX*. Em (a) modelo 3D predito pelo aplicativo *RaptorX* com as porcentagens nas condições de exposto ao solvente, exposição média ao solvente e pouco acesso ao solvente representadas pelas letras E, M e B, respectivamente. Em (b) ocorre a distribuição gráfica da predição de desordem dos resíduos de aminoácidos da proteína LIP4. Em (c) ocorre a distribuição gráfica da acessibilidade ao solvente dos resíduos de aminoácidos da proteína LIP4. Os números em (b) e (c) representam o número de resíduos de aminoácidos (aa) e as letras representam os resíduos de aa da sequência de lipocalina 4. E: *exposed*; M: *medium*; B: *buried*.

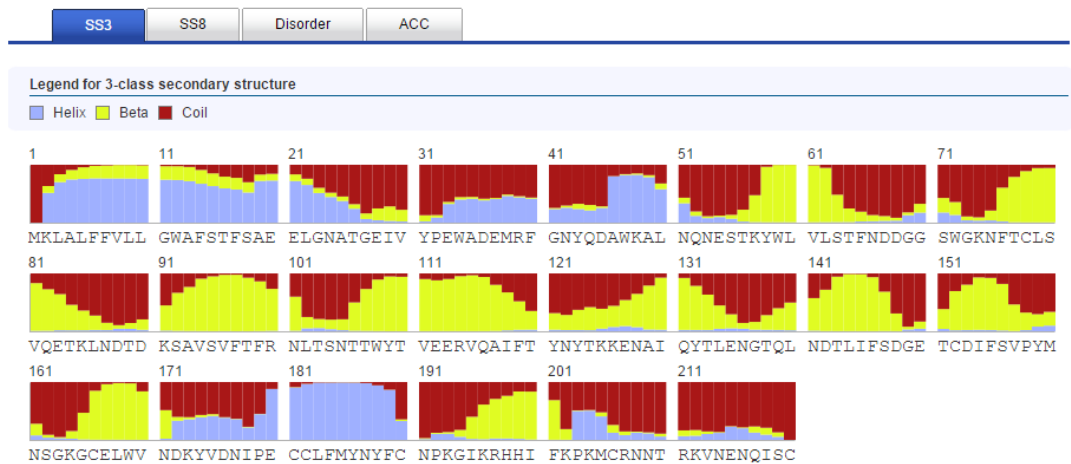
A figura 12 (a) apresenta o modelo 3D predito pelo aplicativo RaptorX contendo as percentagens correspondentes a acesso ao solvente. A LIP4 apresenta 33% de sua estrutura exposta ao solvente, 40% parcialmente exposta e 25% não exposta ao solvente. Portanto, a LIP4 possui uma acessibilidade intermediária (entre 10 a 42 %). A comparação da distribuição gráfica dos resíduos da LIP4 das figuras 12 (b) e (c) resulta em um modelo apresentando ausência de desordem (correspondendo à cor azul da figura 12 (b)) e ainda a figura 12 (c) revela que a maioria dos resíduos do modelo 3D situam-se em regiões de acessibilidade média (cor verde da figura 12 (c)) o que corrobora com as percentagens entre a faixa de 10 a 42% apresentadas no modelo 3D da figura 12 (a).

A Figura 13 apresenta os resultados da predição e avaliação de modelos estruturais da proteína LIP7 preditos pelo aplicativo *RaptorX*.

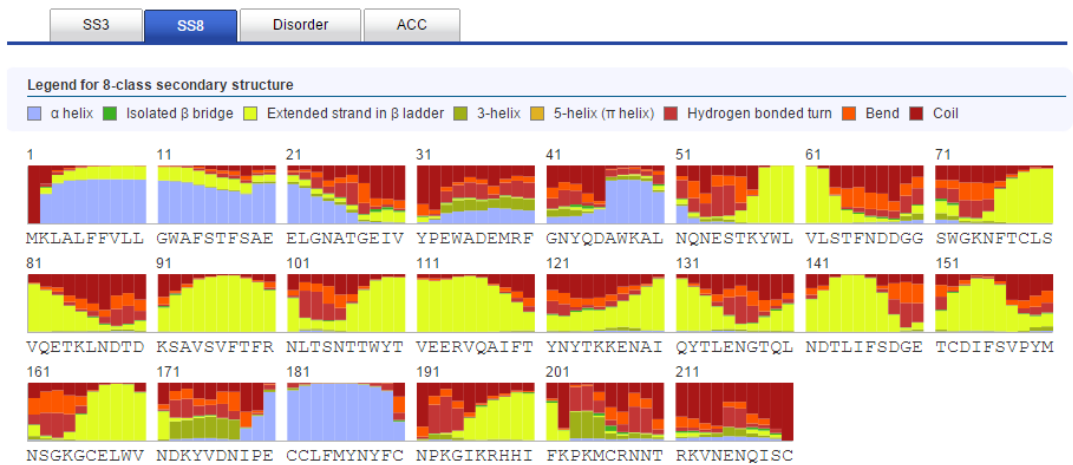


•Estruturas secundárias: 16%H, 33%E, 50%C

a)



b)



c)

Figura 13: Distribuição gráfica da participação dos resíduos na formação das estruturas secundárias compoendo as α hélices, folhas β e espirais da Lipocalina 7 previstas pelo aplicativo *RaptorX*. Em (a) modelo 3D previsto pelo aplicativo *RaptorX*. Em (b) ocorre a distribuição principal em 3 classes de resíduos (hélices, folhas e espirais). Em (c) detalhamento da predição estrutural em 8 classes de estruturas secundárias (α hélices, pontes β isoladas, prolongamento de cadeias β , hélice-3, hélice-5, interações de hidrogênio, dobramentos e *loop*). Os números em (b) e (c) representam o número de resíduos de aminoácidos (aa) e as letras representam os resíduos de aa da sequência de lipocalina 4. H: hélice; E: folhas β ; C: estruturas em espiral (*loop*).

A figura 13 (a) apresenta o modelo 3D predito pelo aplicativo RaptorX contendo as percentagens correspondentes à estrutura secundária da proteína LIP7. A mesma revela que 16% dos resíduos compõe as α hélices (letra H figura 13 (a)), 33% compõe as folhas β (letra E figura 13 (a)) e 50% dos resíduos compõem os *loops* (letra C figura 13 (a)). As folhas β apresentadas no modelo 3D constituem a estrutura em β barril.

A figura 13 (b) resultou no reconhecimento da assinatura de lipocalinas da sequência submetida uma vez que a predição da estrutura secundária realizada pelo aplicativo *RaptorX* revelou a presença de 4 α hélices (duas comuns e duas menores possivelmente hélice 3¹⁰) e 8 folhas β revelando a assinatura estrutural de lipocalinas que necessita de pelo menos 8 folhas β (com topologia classificada de A-H) para formação do barril β que corresponderá ao “cálice” para abrigar lipídeos e ainda a presença de 3 α hélices sendo uma menor denominada de hélice-3 e a presença de *loops* característicos para a estabilização das folhas β (FLOWER et al., 2000; SKERRA et al., 2000).

Detalhadamente a predição da estrutura secundária feita pelo aplicativo *RaptorX* representada pela figura 13 (b) revelou a presença de 4 α hélices: a primeira a partir do resíduo de lisina (K) na posição 2 até o resíduo de triptofano (G) na posição 26; a segunda iniciando no resíduo de prolina (P) na posição 32 até o resíduo de glutamina (G) na posição 52; a terceira, iniciando no resíduo de treonina (T) na posição 56 até o resíduo de aspartato (D) na posição 67 e possivelmente uma quarta hélice iniciando no resíduo de aspartato (D) na posição 172 até o resíduo de lisina (K) na posição 193. Quanto a folhas β responsáveis pela formação da estrutura de barril β que corresponderá ao cálice foram preditas 8 folhas: a primeira iniciando no resíduo de triptofano (T) na posição 56 até o resíduo de aspartato (D) na posição 67; a segunda iniciando no resíduo de asparagina (N) na posição 75 até o resíduo de aspartato (D) na posição 88; a terceira iniciando no resíduo de aspartato (D) na posição 90 até o resíduo de asparagina (N) na posição 105; a quarta iniciando no resíduo de treonina (T) na posição 106 até o resíduo de asparagina (N) na posição 136; a quinta iniciando no resíduo de glicina (G) na posição 137 até o resíduo de aspartato (D) na posição 148; a sexta iniciando no resíduo de glutamato (E) na posição 150 até o resíduo de serina (S) na posição 162; a sétima iniciando no resíduo de lisina (K) na posição 164 até o resíduo de asparagina (N) na posição 171 e finalmente a oitava iniciando no resíduo de glicina (G) na posição 194 até o resíduo de lisina (K) na posição 202.

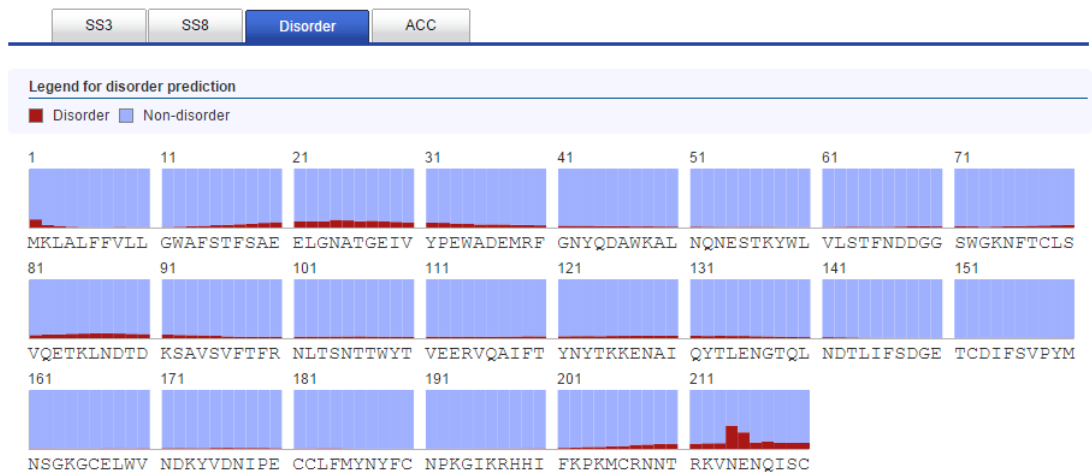
A figura 13 (c) demonstra a presença de possivelmente duas α hélices 3¹⁰, a primeira iniciando no resíduo de aspartato (D) na posição 172 até o resíduo de

asparagina (N) na posição 177 sendo formada por 6 resíduos de aa DKYVDN e a segunda α hélice iniciando no resíduo de prolina (P) na posição 203 até metionina (M) na posição 205 compondo uma hélice de 3 resíduos de aa PKM.

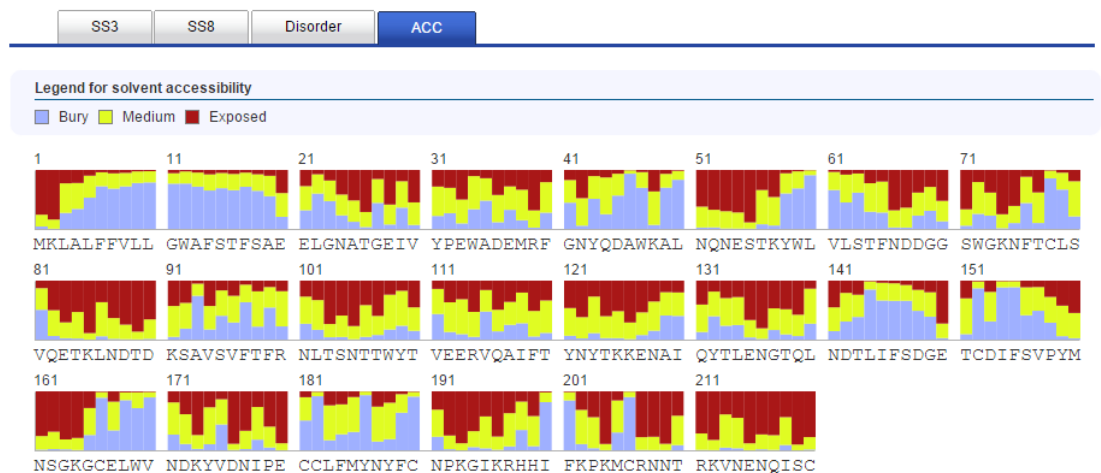
A Figura 14 apresenta os resultados da predição da desordem e acessibilidade ao solvente dos modelos estruturais da proteína LIP7 predito pelo aplicativo *RaptorX*.



a) •Acesso ao solvente: 37%E, 35%M, 27%B



b)



c)

Figura 14: Distribuição gráfica da participação dos resíduos de aminoácidos da Lipocalina 7 na predição dos parâmetros de desordem e acessibilidade do solvente preditas pelo aplicativo *RaptorX*. Em (a) modelo 3D predito pelo aplicativo *RaptorX* com as porcentagens nas condições de exposto ao solvente, exposição média ao solvente e pouco acesso ao solvente representadas pelas letras E, M e B, respectivamente. Em (b) ocorre a distribuição gráfica da predição de desordem dos resíduos de aminoácidos da proteína LIP7. Em (c) ocorre a distribuição gráfica da acessibilidade ao solvente dos resíduos de aminoácidos da proteína LIP7. Os números em (b) e (c) representam o número de resíduos de aminoácidos (aa) e as letras representam os resíduos de aa da sequência de lipocalina 7. E: *exposed*; M: *medium*; B: *buried*.

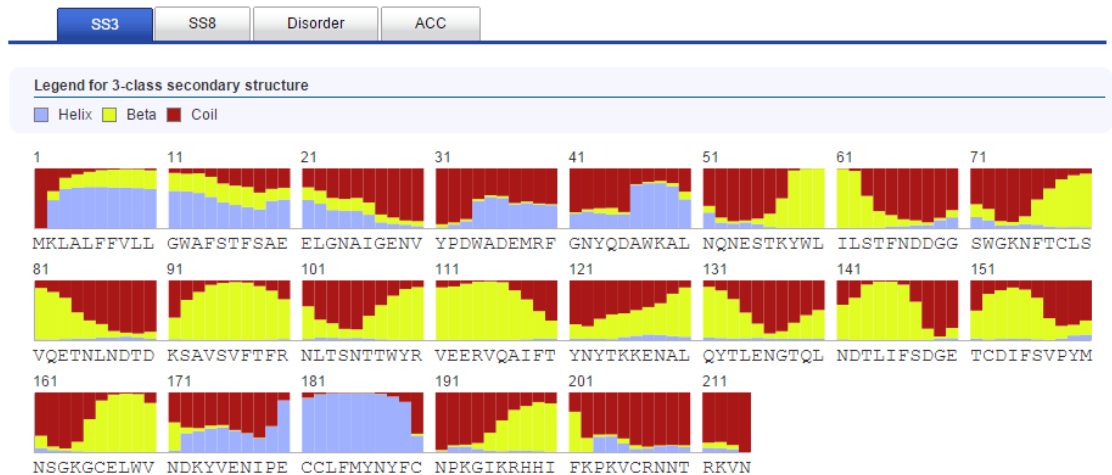
A figura 14 (a) apresenta o modelo 3D predito pelo aplicativo RaptorX contendo as porcentagens correspondentes a acesso ao solvente. A LIP7 apresenta 37% de sua estrutura exposta ao solvente, 35% parcialmente exposta e 27% não exposta ao solvente. Portanto, a LIP7 possui uma acessibilidade intermediária (entre 10 a 42 %). A comparação da distribuição gráfica dos resíduos da LIP7 das figuras 14 (b) e (c) resulta em um modelo apresentando desordem nos resíduos de metionina (M) na posição 1, asparagina (N) na posição 215 e glutamato (E) na posição 216 (correspondendo à cor marrom da figura 13 (b)) e ainda a figura 14 (c) revela que a maioria dos resíduos do modelo 3D situam-se em regiões de acessibilidade média (cor verde da figura 14 (c)) o que corrobora com as porcentagens entre a faixa de 10 a 42% apresentadas no modelo 3D da figura 14 (a).

A Figura 15 apresenta os resultados da predição e avaliação de modelos estruturais da proteína LIP24 preditos pelo aplicativo *RaptorX*.

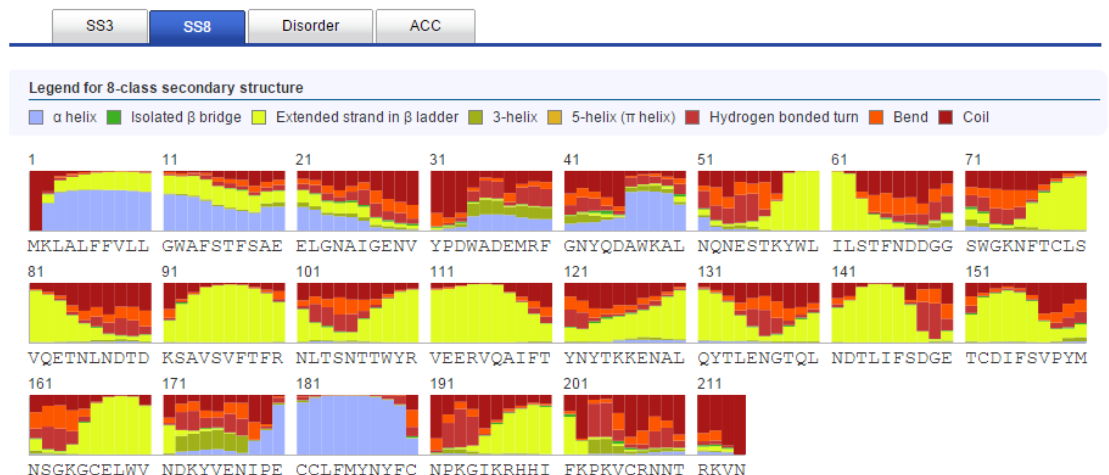


•Estruturas secundárias: 17%H, 30%E, 51%C

a)



b)



c)

Figura 15: Distribuição gráfica da participação dos resíduos na formação das estruturas secundárias compoendo as α hélices, folhas β e espirais da Lipocalina 24 preditas pelo aplicativo *RaptorX*. Em (a) modelo 3D predito pelo aplicativo *RaptorX*. Em (b) ocorre a distribuição principal em 3 classes de resíduos (hélices, folhas e espirais). Em (c) detalhamento da predição estrutural em 8 classes de estruturas secundárias (α hélices, pontes β isoladas, prolongamento de cadeias β , hélice-3, hélice-5, interações de hidrogênio, dobramentos e *loop*). Os números em (b) e (c) representam o número de resíduos de aminoácidos (aa) e as letras representam os resíduos de aa da sequência de lipocalina 24. H: hélice; E: folhas β ; C: estruturas em espiral (*loop*).

A figura 15 (a) apresenta o modelo 3D predito pelo aplicativo RaptorX contendo as percentagens correspondentes à estrutura secundária da proteína LIP24. A mesma revela que 17% dos resíduos compõe as α hélices (letra H figura 15 (a)), 30% compõe as folhas β (letra E figura 15 (a)) e 51% dos resíduos compõem os *loops* (letra C figura 15 (a)). As folhas β apresentadas no modelo 3D constituem a estrutura em β barril.

A figura 15 (b) resultou no reconhecimento da assinatura de lipocalinas da sequência submetida uma vez que a predição da estrutura secundária realizada pelo aplicativo *RaptorX* revelou a presença de 4 α hélices (2 comuns e 2 menores possivelmente hélice 3¹⁰) e 9 folhas β revelando a assinatura estrutural de lipocalinas que necessita de pelo menos 8 folhas β (com topologia classificada de A-H) para formação do barril β que corresponderá ao “cálice” para abrigar lipídeos e ainda a presença de 4 α hélices sendo uma menor denominada de hélice-3 e a presença de *loops* característicos para a estabilização das folhas β (FLOWER et al., 2000; SKERRA et al., 2000).

Detalhadamente a predição da estrutura secundária feita pelo aplicativo *RaptorX* representada pela figura 15 (b) revelou a presença de 4 α hélices: a primeira a partir do resíduo de lisina (K) na posição 2 até o resíduo de asparagina (N) na posição 29; a segunda iniciando no resíduo de prolina (P) na posição 32 até o resíduo de glutamina (Q) na posição 52; a terceira, iniciando no resíduo de aspartato (D) na posição 172 até o resíduo de cisteína (C) na posição 190 e possivelmente uma quarta hélice iniciando no resíduo de arginina (R) na posição 203 até o resíduo de lisina (K) na posição 205. Quanto a folhas β responsáveis pela formação da estrutura de barril β que corresponderá ao cálice foram preditas 9 folhas: a primeira iniciando no resíduo de treonina (T) na posição 56 até o resíduo de aspartato (D) na posição 67; a segunda iniciando no resíduo de glicina (G) na posição 73 até o resíduo de treonina (T) na posição 89; a terceira iniciando no resíduo de aspartato (D) na posição 90 até o resíduo de serina (N) na posição 104; a quarta iniciando no resíduo de treonina (T) na posição 106 até o resíduo de asparagina (N) na posição 122; a quinta iniciando no resíduo de tirosina (Y) na posição 123 até o resíduo de asparagina (N) na posição 136; a sexta iniciando no resíduo de glicina (G) na posição 137 até o resíduo de glicina (G) na posição 149; a sétima iniciando no resíduo de glutamato (E) na posição 150 até o resíduo de glicina (G) na posição 163 a oitava iniciando no resíduo de lisina (K) na posição 164 até o resíduo de asparagina (N) na posição 171 e finalmente a nona iniciando no resíduo de isoleucina (Ile) na posição 195 até o resíduo de lisina (k) na posição 202.

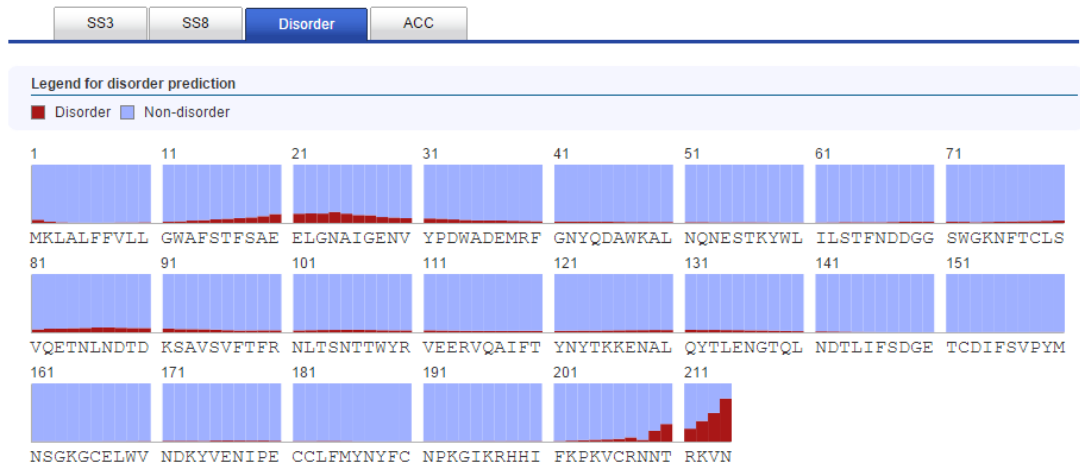
A figura 15 (c) demonstra a presença de possivelmente duas α hélices 3^{10} , a primeira iniciando no resíduo de aspartato (D) na posição 172 até o resíduo de asparagina (N) na posição 177 sendo formada por 6 resíduos de aa NDKYVEN e a segunda α hélice iniciando no resíduo de prolina (P) na posição 203 até valina (V) na posição 205 compondo uma α hélice de 3 resíduos de aa PKM.

A Figura 16 apresenta os resultados da predição da desordem e acessibilidade ao solvente dos modelos estruturais da proteína LIP24 predito pelo aplicativo *RaptorX*.

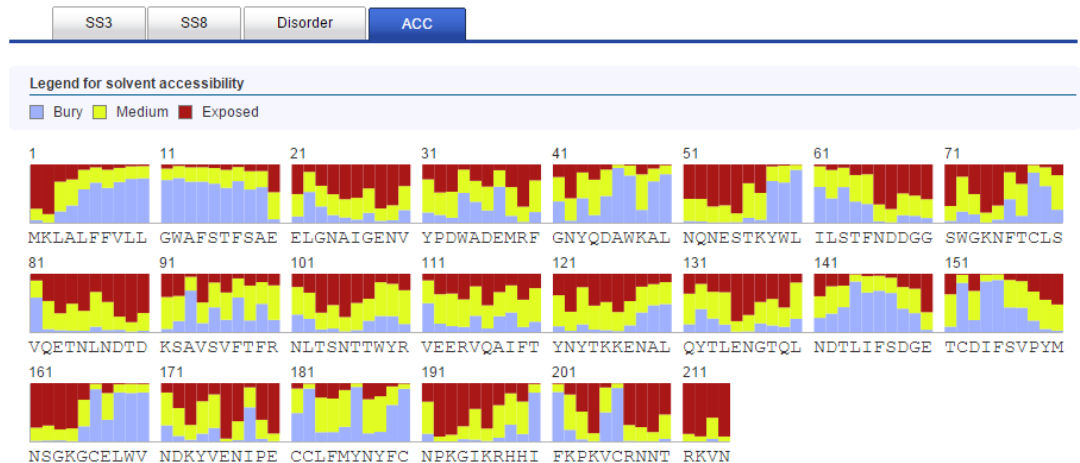


a)

•Acesso ao solvente: 35%E, 36%M, 28%B



b)



c)

Figura 16: Distribuição gráfica da participação dos resíduos de aminoácidos da Lipocalina 24 na predição dos parâmetros de desordem e acessibilidade do solvente preditas pelo aplicativo *RaptorX*. Em (a) modelo 3D predito pelo aplicativo *RaptorX* com as porcentagens nas condições de exposto ao solvente, exposição média ao solvente e pouco acesso ao solvente representadas pelas letras E, M e B, respectivamente. Em (b) ocorre a distribuição gráfica da predição de desordem dos resíduos de aminoácidos da proteína LIP24. Em (c) ocorre a distribuição gráfica da acessibilidade ao solvente dos resíduos de aminoácidos da proteína LIP24. Os números em (b) e (c) representam o número de resíduos de aminoácidos (aa) e as letras representam os resíduos de aa da sequência de lipocalina 24. E: *exposed*; M: *medium*; B: *buried*.

A figura 16 (a) apresenta o modelo 3D predito pelo aplicativo RaptorX contendo as percentagens correspondentes a acesso ao solvente. A LIP24 apresenta 35% de sua estrutura exposta ao solvente, 36% parcialmente exposta e 28% não exposta ao solvente. Portanto, a LIP24 possui uma acessibilidade intermediária (faixa entre 10 a 42 %). A comparação da distribuição gráfica dos resíduos da LIP24 das figuras 16 (b) e (c) resulta em um modelo apresentando desordem em diversas regiões, a começar no resíduo de metionina (M) na posição 1, da região correspondente ao resíduo de fenilalanina (F) na posição 14 até o resíduo de aspartato (D) na posição 33, de valina (V) na posição 81 até o resíduo de leucina (L) na posição 140 e lisina (k) na posição 202 até a asparagina (N) na posição 214 (correspondendo à cor marrom da figura 16 (b)) e ainda a figura 16 (c) revela que a maioria dos resíduos do modelo 3D situam-se em regiões de acessibilidade média (cor verde da figura 15 (c)) o que corrobora com as percentagens entre a faixa de 10 a 42% apresentadas no modelo 3D da figura 16 (a).

Após a avaliação das predições analisou-se os valores dos parâmetros preditos pelo aplicativo *RaptorX* para as proteínas LIP4, 7 e 24 (Tabela 8).

Tabela 8: Resultados dos parâmetros do aplicativo *RaptorX* na predição dos modelos estruturais das lipocalinas LIP4, 7 e 24.

Proteína	N	<i>p-value</i>	uGDT	GDT	RM	D %	S
LIP4	220	8,7E-12	132	60	220	0	133
LIP7	220	6,07E-12	133	60	220	2	136
LIP24	214	5,45E-12	131	61	214	5	135

N: número de resíduos de aminoácidos das proteínas; uGDT: *un-normalized global distance test* (uGDT); GDT: *global distance test* (GDT); RM: número de resíduos modelados; D: índice de desordem; S: Score do modelo predito.

A análise dos modelos preditos pelo aplicativo RaptorX resumidos na tabela 8 demonstra que todos os resíduos presentes nas proteínas foram modelados (N=RM), os modelos preditos não foram ao acaso pois os *p-value* de todas as lipocalinas preditas obtiveram um valor superior a ordem de 10^{-4} (considerado um *cut off* para um bom modelo de predição) e finalmente tanto o uGDT quanto o GDT foram acima de 50 corroborando com a qualidade do modelo predito (JIANZHU et al., 2013).

Além do aplicativo *RaptorX*, as LIP4 7 e 24 foram submetidas ao aplicativo *Phyre2*, utilizado para avaliar a qualidade dos modelos preditos. A análise das sequências proteicas referentes aos transcritos *lip 4, 7 e 24* no *Phyre2* produziu um conjunto de modelos 3D do *input*, baseadas em *templates* depositados, feito pela detecção de sequências homólogas com o PSI-BLAST, seguida de uma previsão da

Através da análise da proteína LIP7 pelo aplicativo PSI-PRED (figura 18) foram identificadas a presença de 3 α hélices e 9 folhas β . A hélice 1 é composta por vinte resíduos de aa (LALFFVLLGWAFSTFSAEEL), a 2 possui 5 resíduos (AWKAL), a 3 possui 5 resíduos (LFMYN). As 9 folhas β possuem a seguinte composição: a folha 1 (KYWLVLSTF), folha 2 (FTCLSVQETK), folha 3 (AVSVFTFRN), folha 4 (TTWYTVEERVQAIFTY), folha 5 (AIQYTL), folha 6 (TLIFS), folha 7 (CDIFSVP), folha 8 (CELWVN) e folha 9 (RHH).

A figura 19 representa a predição das estruturas secundárias da LIP24 através do aplicativo PSI-PRED.



Figura 19: Predição das estruturas secundárias da LIP24.

Através da análise da proteína LIP24 pelo aplicativo PSI-PRED (figura 18) foram identificadas a presença de 3 α hélices e 9 folhas β . A hélice 1 é composta por quinze resíduos de aa (LALFFVLLGWAFSTF), a 2 possui 5 resíduos (AWKAL), a 3 possui 7 resíduos (CLFMVNY). As 9 folhas β possuem a seguinte composição: a folha 1 (KYWLILSTF), folha 2 (FTCLSVQET), folha 3 (AVSVFTFR), folha 4 (TTWYRVEERVQAIFTY), folha 5 (ALQYTL), folha 6 (TLIFS), folha 7 (CDIFSVP), folha 8 (CELWVN) e folha 9 (RHHIF).

Para a predição de possível domínio transmembrana e presença de peptídeo sinal as lipocalinas 4, 7 e 24 foram submetidas ao aplicativo memsat-svm presente no PSI-PRED. A figura 20 representa o resultado da predição transmembrana dessas proteínas.

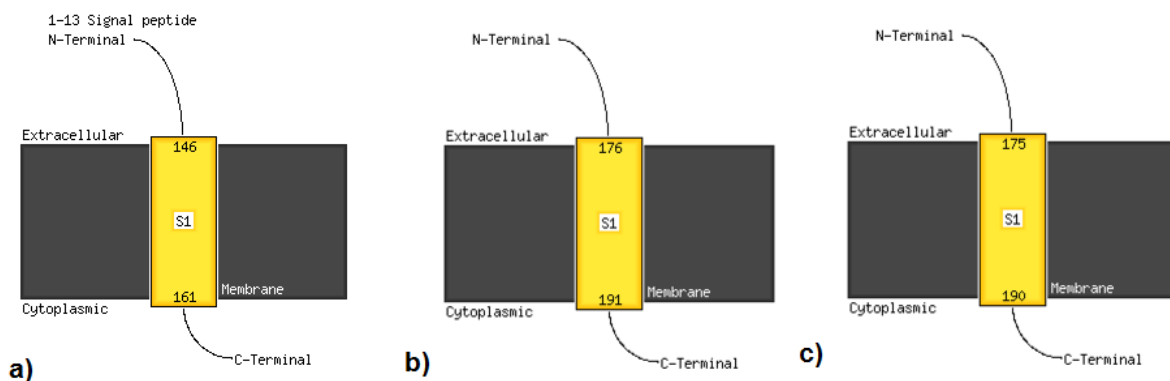


Figura 20: Análise de domínio transmembrana e presença de peptídeo sinal das lipocalinas 4 (a), 7(b) e 24(c) predita pelo memsat-svm do PSI-PRED.

A lipocalina 4 (figura 20 (a)) possui um peptídeo sinal composto de 13 resíduos de aminoácidos (MKLALFFVLLGWA) e um domínio transmembrana cuja a porção extracelular inicia-se no resíduo de fenilalanina (F) na posição 14 e finaliza em outro resíduo de fenilalanina (F) na posição 146 a sua porção intracelular inicia-se no resíduo de asparagina (N) na posição 161 até atingir a extremidade C-terminal. A análise pelo memsat-svm não identificou a presença de peptídeo sinal nas estruturas das lipocalina 7 (figura 20 (b)) e 24 (figura 20 (c)). A LIP7 possui um domínio transmembrana tendo como início a porção extracelular os aminoácidos próximos a extremidade N-terminal até o resíduo de aspartato (D) na posição 176 e intracelular a partir do resíduo de asparagina (N) na posição 191 em direção a extremidade C-terminal. A LIP24 possui um domínio transmembrana tendo como início a porção extracelular os aminoácidos próximos a extremidade N-terminal até o resíduo de valina (V) na posição 175 e a

porção intracelular tendo incício no resíduo de cisteína (C) na posição 190 até a extremidade N-terminal.

Para avaliar a qualidade das estruturas terciárias previstas no modelo 3D das proteínas LIP4, LIP7 e LIP24 a análise estrutural das lipocalinas 4, 7 e 24 foi realizada uma análise os modelos foram analisados pelo aplicativo *Ramachandran Plot Analysis* (RAMPAGE) (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>).

A figura 21 representa o diagrama de Ramachandram resultante da análise da lipocalina 4 no aplicativo RAMPAGE.

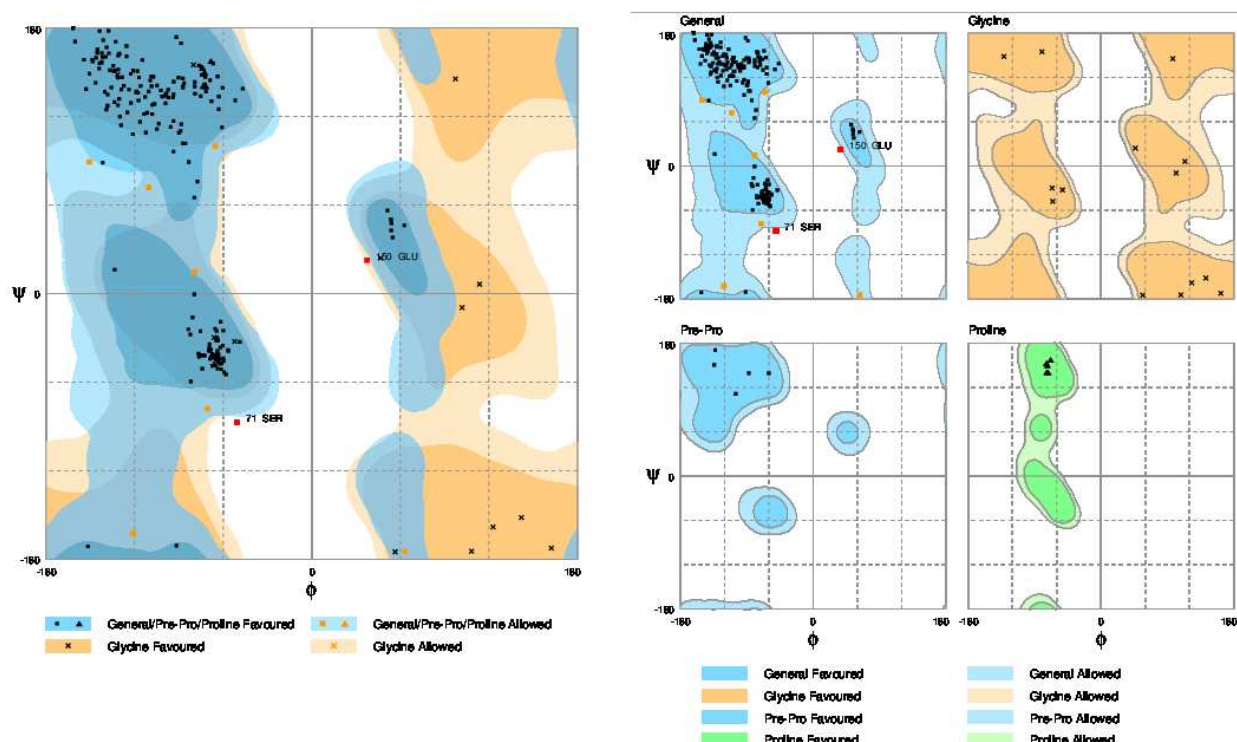


Figura 21: Diagrama de ramachandram construído pelo aplicativo RAMPAGE demonstrando a densidade de valores do par (ϕ, ψ) dos aminoácidos da lipocalina 4.

O diagrama de Ramachandram descreve os valores prováveis que os ângulos diedrais (ϕ, ψ) podem assumir e os compara a uma biblioteca de valores de ângulos diedrais. O mesmo restringe os ângulos aos valores correspondentes a dobramentos factíveis para estruturas secundárias (BERG et al., 2002). Através do aplicativo RAMPAGE foram avaliados 9 resíduos de aminoácidos (correspondentes aos 220 que compõe a proteína LIP4). Desses 7 foram identificados em regiões favoráveis correspondendo a aproximadamente 95,9% (209 resíduos), 3,2% (7 resíduos) encontram-se em regiões permitidas e 0,9% (cerca de 2 resíduos) encontram-se em regiões proibidas. O que se conclui que o modelo 3D previsto para a lipocalina 4 possui

alta qualidade. A tabela 9 demonstra o valor do par (ϕ, ψ) calculado pelo RAMPAGE para a composição do diagrama de Ramachandram.

Tabela 9: análise do diagrama de Ramachandram apresentando o cálculo dos ângulos diedrais do par (ϕ, ψ) para a lipocalina 4.

NR	aa	Par $(\phi; \psi)$	Região
45	N	-63,34; 100,04	Permitida
104	S	-70,63; -77,99	Permitida
159	S	-79,25; 14,66	Permitida
161	N	-110,43; 71,92	Permitida
174	Y	63,20; -174,63	Permitida
177	N	-150,71; 89,21	Permitida
215	D	-120,92; -162,52	Permitida
71	S	-50,59; -87,19	Proibida
150	E	37,45; 22,62	Proibida

NR: número do resíduo; aa: aminoácido; ângulo diedral $\phi; \psi$.

A figura 22 representa o diagrama de Ramachandram resultante da análise da lipocalina 7 no aplicativo RAMPAGE.

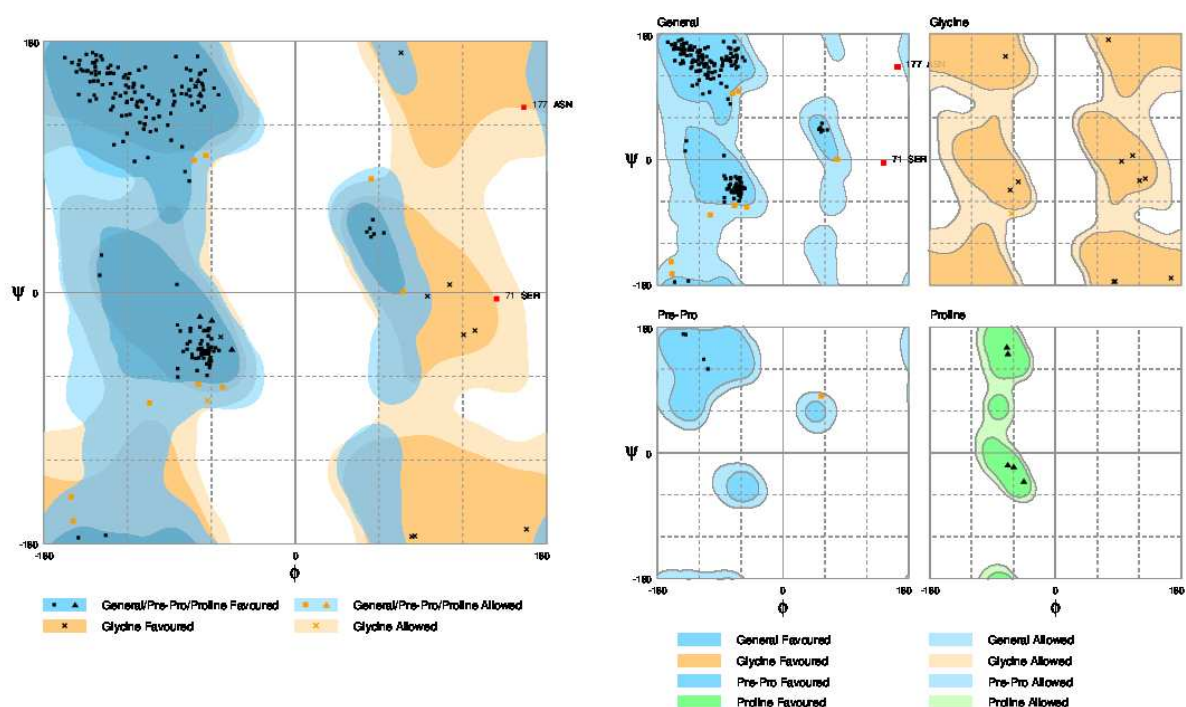


Figura 22: Diagrama de ramachandram construído pelo aplicativo RAMPAGE demonstrando a densidade de valores do par (ϕ, ψ) dos aminoácidos da lipocalina 7.

Através do aplicativo RAMPAGE foram avaliados 12 resíduos de aminoácidos (correspondentes aos 220 que compõe a proteína). Desses 10 foram identificados em regiões favoráveis correspondendo a aproximadamente 94,5% (206 resíduos), 4,6% (10 resíduos) encontram-se em regiões permitidas e 0,9 % (cerca de 2 resíduos) encontram-se em regiões proibidas. O que se conclui que o modelo 3D previsto para a

lipocalina 7 possui alta qualidade. A tabela 10 demonstra o valor do par (ϕ, ψ) calculado pelo RAMPAGE para a composição do diagrama de Ramachandram.

Tabela 10: análise do diagrama de Ramachandram apresentando o cálculo dos ângulos diedrais do par (ϕ, ψ) para a lipocalina 7.

NR	aa	Par $(\phi; \psi)$	Região
4	D	-63,9; 98,23	Permitida
68	D	-104,19; -79,14	Permitida
104	S	-69,35; -65,80	Permitida
127	Q	-72,28; 94,81	Permitida
149	G	-62,67; -77,30	Permitida
159	Y	77,37; 0,71	Permitida
162	S	-160,28; -146,27	Permitida
174	Y	-51,96; -67,79	Permitida
191	N	54,52; 81,40	Permitida
211	R	-58,98; -163,48	Permitida
71	S	144,26; -4,52	Proibida
177	N	163,84; 132,56	Proibida

NR: número do resíduo; aa: aminoácido; ângulo diedral $\phi; \psi$.

A figura 23 representa o diagrama de Ramachandram resultante da análise da lipocalina 24 no aplicativo RAMPAGE.

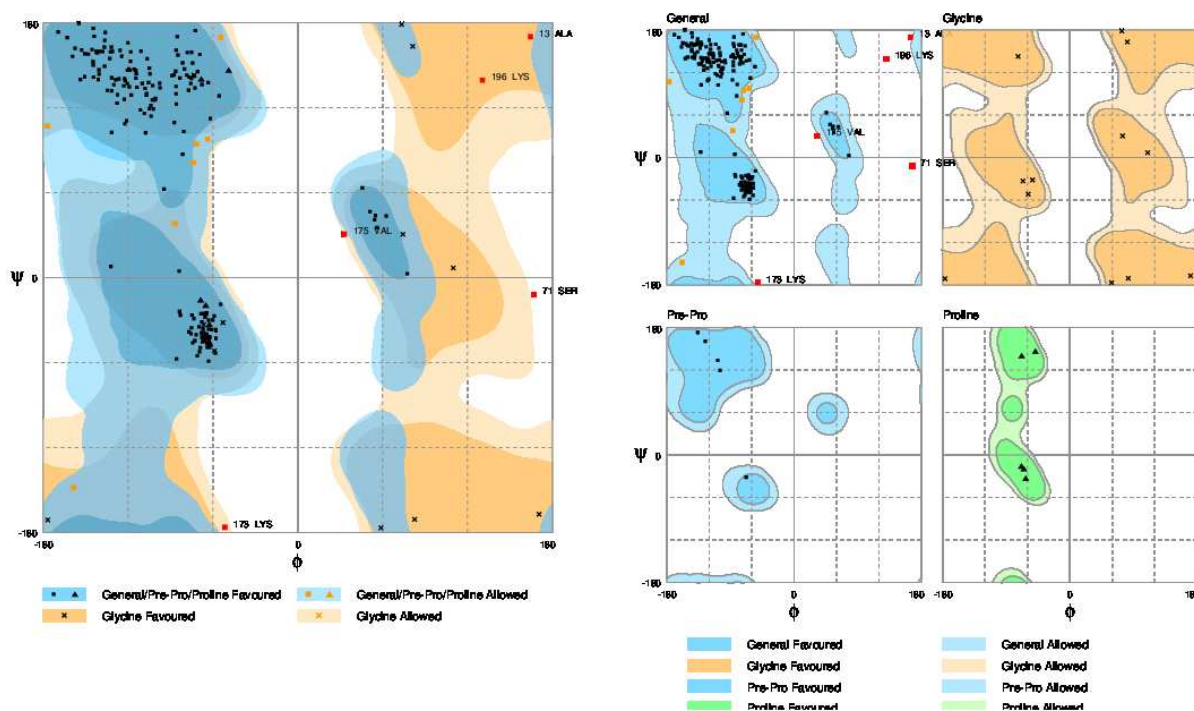


Figura 23: Diagrama de ramachandram construído pelo aplicativo RAMPAGE demonstrando a densidade de valores do par (ϕ, ψ) dos aminoácidos da lipocalina 24.

Através do aplicativo RAMPAGE foram avaliados 12 resíduos de aminoácidos (correspondentes aos 214 que compõe a proteína). Desses 7 foram identificados em regiões favoráveis correspondendo a aproximadamente 94,3% (200 resíduos), 3,3% (7 resíduos) encontram-se em regiões permitidas e 2,4% (cerca de 5 resíduos)

encontram-se em regiões proibidas. O que se conclui que o modelo 3D previsto para a lipocalina 24 possui alta qualidade. A tabela 11 demonstra o valor do par (ϕ, ψ) calculado pelo RAMPAGE para a composição do diagrama de Ramachandram.

Tabela 11: análise do diagrama de Ramachandram apresentando o cálculo dos ângulos diedrais do par (ϕ, ψ) para a lipocalina 24.

NR	aa	Par $(\phi; \psi)$	Região
35	A	-73,69 ; 81,20	Permitda
45	D	-63,61; 97,77	Permitda
103	T	-54,28 ; 169,66	Permitda
105	N	-71,59 ; 94,33	Permitda
126	K	-86,86 ; 38,04	Permitda
162	S	-158,40 ; -148,2	Permitda
177	N	-176,90 ; 106,99	Permitda
13	A	164,58 ; 170,06	Proibida
71	S	167,00 ; -11,82	Proibida
173	K	-51,48 ; 176,28	Proibida
175	V	33,00 ; 30,67	Proibida
196	K	130,60 ; 139,39	Proibida

NR: número do resíduo; aa: aminoácido; ângulo diedral $\phi; \psi$.

7.5.3 Validação do teste de concentração e eficiência dos primers de lipocalinas e dos candidatos a genes de referência para uso em RTq-PCR

Foram realizados teste de concentração e eficiência para os *primers* dos genes de lipocalinas *lip4*, *lip7* e *lip24* esses obtiveram uma curva de amplificação com um Ct na faixa entre 26-30. A literatura recomenda que não seja utilizado *primers* que amplifiquem em grande concentração (Ct < 15) ou em baixas concentrações (Ct > 30) (WANG et al., 2010). Os 3 *primers* de lipocalinas obtiveram sucesso nessa etapa em detrimentos dos demais. A curva de dissociação está representa na figura 24.

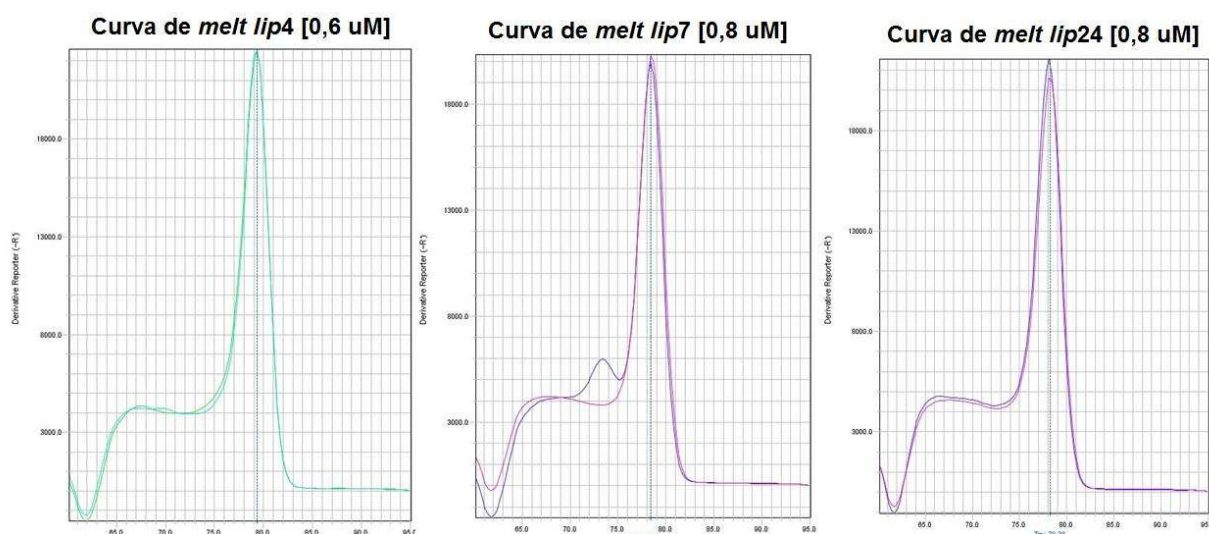


Figura 24: Curva de dissociação (curva de *melt*) dos *primers* de lipocalinas desenhados, as concentrações ideais para cada primer encontra-se entre os colchetes [].

A curva de dissociação apresentada pelos *lip 4*, *lip 7* e *lip 24* confirmam uma amplificação numa temperatura de *melt* em torno de 80 °C (figura 24). E ainda, a presença de picos coincidentes demonstra a especificidade dos *primers* testados para a obtenção do *amplicon*.

A figura 25, representa a curva de *melt* obtida para os genes constitutivos candidatos a genes de referência *hk2*, *hk4* e *hk6*.

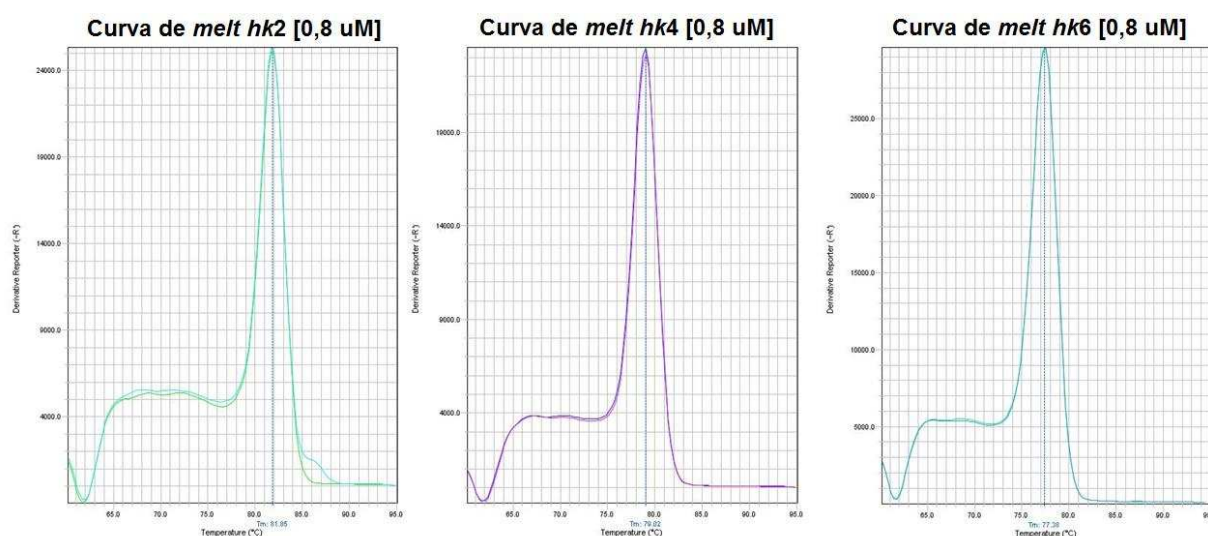


Figura 25: Curva de dissociação (curva de *melt*) dos *primers* dos genes constitutivos desenhados, as concentrações ideais para cada primer encontra-se entre os colchetes [].

A figura 25 demonstra que os *primers* dos genes constitutivos candidatos a gene de referência *hk2*, *hk4* e *hk6* obtiveram uma curva de amplificação com um Cp na faixa entre 24-28. A literatura recomenda que não seja utilizado *primers* que amplifiquem em grande concentração ($C_p < 15$) ou em baixas concentrações ($C_p > 30$) (WANG et al., 2010; FURQUIM et al., 2011). A curva de dissociação dos genes *hk2*, *hk4* e *hk6* resultaram na amplificação específica dos alvos reforçando os resultados do teste de eficiência e concentração de *primers*. Estudos semelhantes de validação de genes constitutivos candidatos a gene de referência resultaram em níveis inconstantes da expressão em diferentes tecidos de organismos avaliados em situação de infecção e não infecção suportando a hipótese de que caso ocorra níveis de expressão diferenciado tal situação pode não se relacionar às condições experimentais e sim às condições biológicas do organismo estudado, mais uma vez a realização do teste de concentração e eficiência dos *primers* desenhados torna mais segura o estabelecimento de potenciais candidatos a gene de referência no tecido estudado (KOZERA & RAPACZ, 2013).

7.6 Análise da expressão gênica de lipocalinas com SHBP *domain* e genes constitutivos por quantificação relativa utilizando RTq-PCR

Após definir a estratégia de amplificação e a melhor concentração dos *primers* a serem utilizados: 2 etapas sendo *Holding* 95 °C, 10 min., 40 ciclos de desnaturação a 95 °C por 15 seg., anelamento e extensão a 60 °C por 1 min. seguida da amplificação de 40 ciclos e desnaturação gradual para realização da curva de dissociação (curva de *melt*) utilizando o aumento de 0,3 °C por 1 min partindo de 60 °C até o limite de 95 °C, procedeu-se à quantificação da expressão relativa através de RTq-PCR.

Para essa análise utilizou-se o método da curva de calibração para cada *primer* avaliado, pois as eficiências dos *primers* usados e o gene de referência não foram comparáveis. As análises dos experimentos foram realizadas em placas de 48 poços com amostras em replicata, analisando os níveis de expressão de lipocalinas e genes constitutivos potenciais candidatos a genes de referência para glândulas salivares, intestinos e ovários de fêmeas de *A. sculptum* todos em relação a condição biológica de infectado e não infectado por *R. amblyommii*. Após as análises foram gerados os gráficos do log₂ da *Relative Quantity* (RQ) de maneira a avaliar a expressão gênica dos genes estudados normalizados para o gene *gapdh* através do aplicativo *StepOne* (*Life Technologies*[®]).

Sendo assim, para analisarmos a importância da expressão dos genes em situação de tecido/órgão infectado e não infectado por *R. amblyommii* e o quanto esse tratamento contribui em nível de expressão sugerindo uma possível regulação positiva ou negativa na expressão gênica, construiu-se os gráficos de log₂ RQ. O resultado reflete a razão entre a condição de órgão/tecido infectado pelo órgão/tecido não infectado. Assim se o log₂ RQ for maior que zero significa que a infecção no órgão/tecido estudado contribui na expressão diferencial do gene estudado, pois a razão citada (infectado/não infectado) será maior que 1, sugerindo uma regulação positiva do gene o que leva a uma barra com valores positivos. Enquanto que, se o log₂ RQ for menor que zero, significa que a infecção do tecido/órgão leva a diminuição da expressão do gene no tecido/órgão estudado sugerindo que a infecção contribui para uma regulação negativa da expressão, o que leva ao aparecimento de uma barra com valores negativos, pois o log₂ RQ em valores menores que 1 é negativo (figura 26).

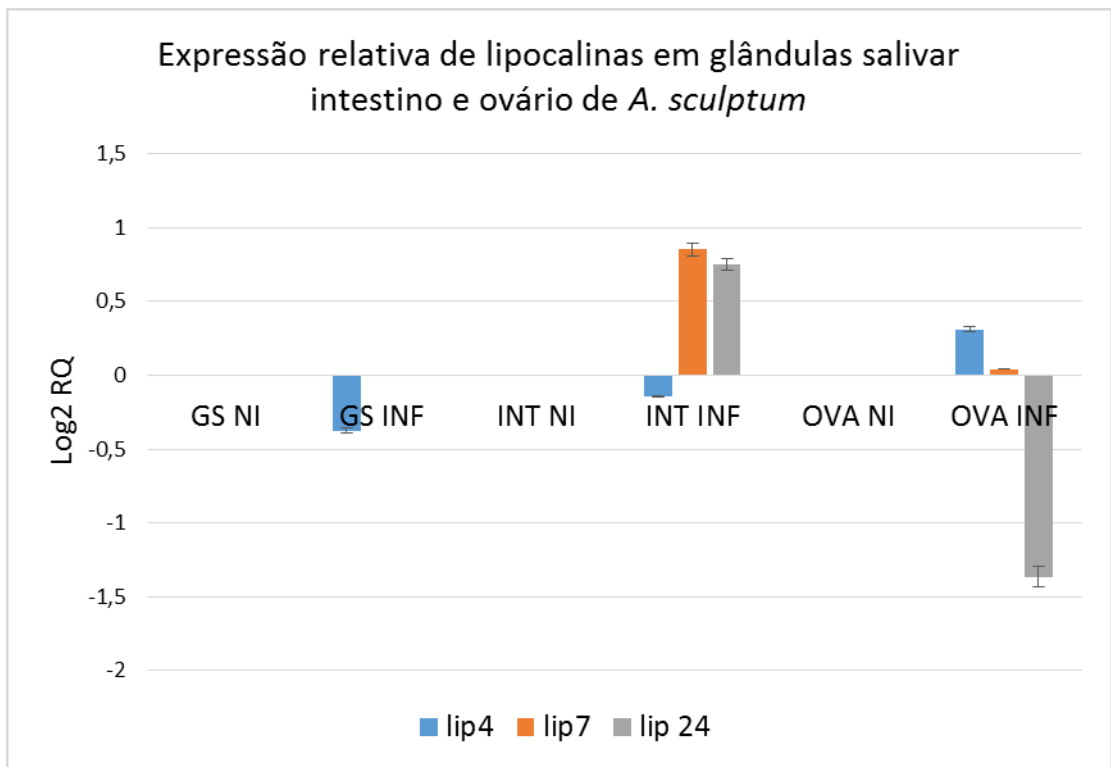


Figura 26: Expressão relativa do log2 RQ (*Relative Quantity*, RQ) dos genes *lip4*, *lip7* e *lip24* de lipocalinas com *SHBP domain* em glândulas salivares, intestinos e ovários de fêmeas do carrapato *A. sculptum* normalizados por *gapdh* em situação de não infecção/infecção com *R. amblyommii*.

A análise da figura 8 apresenta a expressão relativa dos genes de lipocalina *lip4*, *lip7* e *lip24* diferentemente expressos tanto em glândulas salivares quanto intestinos e ovários. A barra positiva observada no gráfico da figura 26 mostra uma regulação positiva (*up regulation*) na expressão do gene observado, a barra negativa mostra a diminuição da expressão (*down regulation*) a ausência da barra indica uma supressão gênica. A figura 26 demonstra que a infecção foi capaz de *down* regular a expressão do gene *lip4* tanto na glândula salivar quanto no intestino e superexpressá-lo no ovário, superexpressar o gene *lip7* no intestino e superexpressar o gene *lip24* em intestino porém *down* regulá-lo no ovário. Esses resultados levam à hipótese de um papel multifuncional das proteínas LIP4, LIP7 e LIP24 provenientes dos genes *lip4*, *lip7* e *lip24*, respectivamente, possivelmente relacionado ao combate da infecção no intestino e controle da reprodução no ovário da fêmea do carrapato *A. sculptum* em situação de infecção por *R. amblyommii*.

Após a verificação global da expressão, realizou-se um estudo individual da expressão desses genes em cada tecido infectado. A análise da figura 26 apresenta o gene de lipocalina *lip4* diferentemente expresso tanto em glândulas salivares quanto intestinos e ovários.

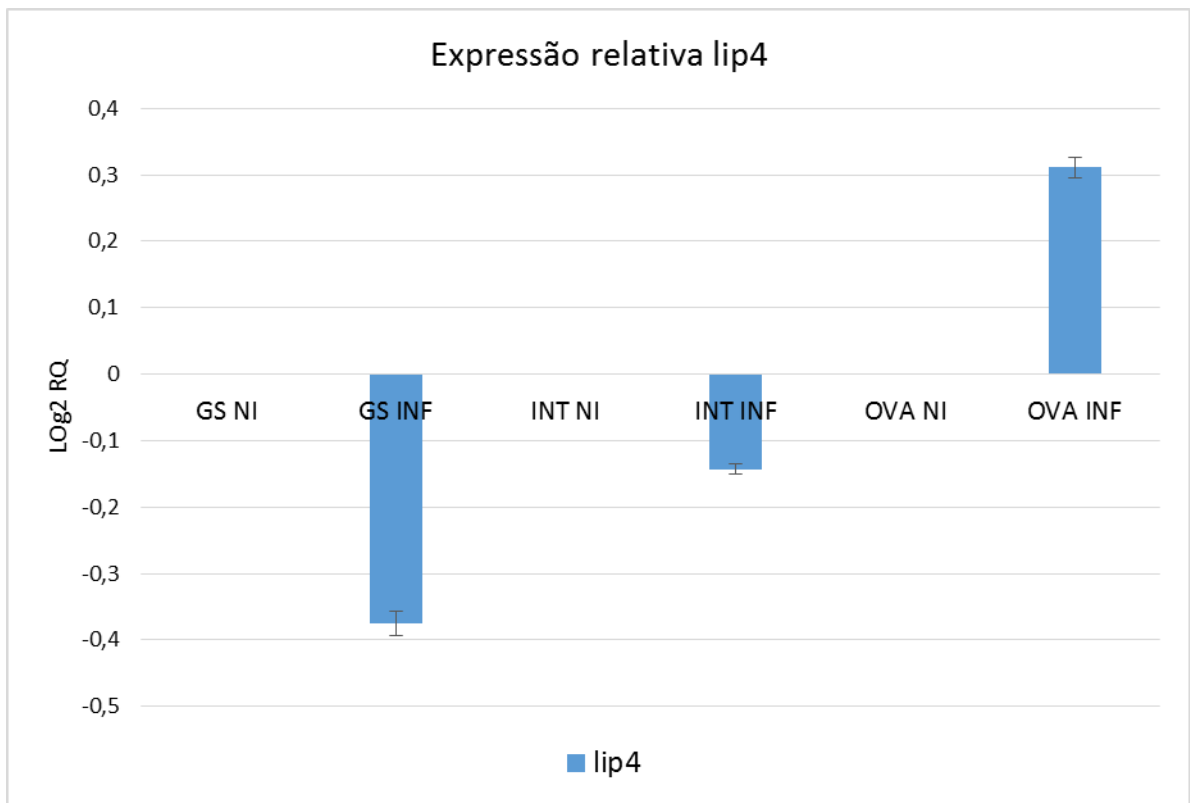


Figura 27: Expressão relativa do log₂ RQ (*relative quantity*, RQ) do gene *lip4* de lipocalinas com *SHBP domain* em glândulas salivares, intestinos e ovários de fêmeas do carrapato *A. sculptum* normalizados por *gapdh* em situação de não infecção/infecção com *R. amblyommi*.

A barra positiva observada no gráfico da figura 27 demonstra uma regulação positiva na expressão de *lip 4*, no ovário e sua *down* regulação tanto em glândulas salivares quanto nos intestinos infectados indicando que a infecção foi capaz de suprimir a expressão do gene tanto na glândula salivar quanto no intestino porém a expressão ovariana foi mantida o que leva à hipótese de um papel multifuncional da proteína LIP4 possivelmente relacionado a reprodução do carrapato *A. sculptum*.

O *lip4* teve sua expressão suprimida em glândulas salivares infectadas e no intestino infectado (figura 27). A supressão desse gene tanto na glândula salivar infectada quanto no intestino infectado leva à hipótese de que o *lip4* pode facilitar a infecção tanto na glândula salivar quanto no intestino, por isso sua *down* regulação por parte do carrapato. Porém no ovário, o aumento da expressão deste gene indica que a proteína LIP4 tem função distinta quando presente em glândula salivar e intestino. Na tentativa de verificar a possível superexpressão do gene *lip4* no ovário, e verificação da hipótese de que a infecção provoca *down* regulação de *lip4* em glândula salivar e intestino e *up* regulação de *lip4* em ovário realizou-se novo experimento utilizando outras amostras de glândula salivar, intestino e ovário não infectados retirados do mesmo tratamento de infecção da fêmea do carrapato *A. sculptum* (figura 28).

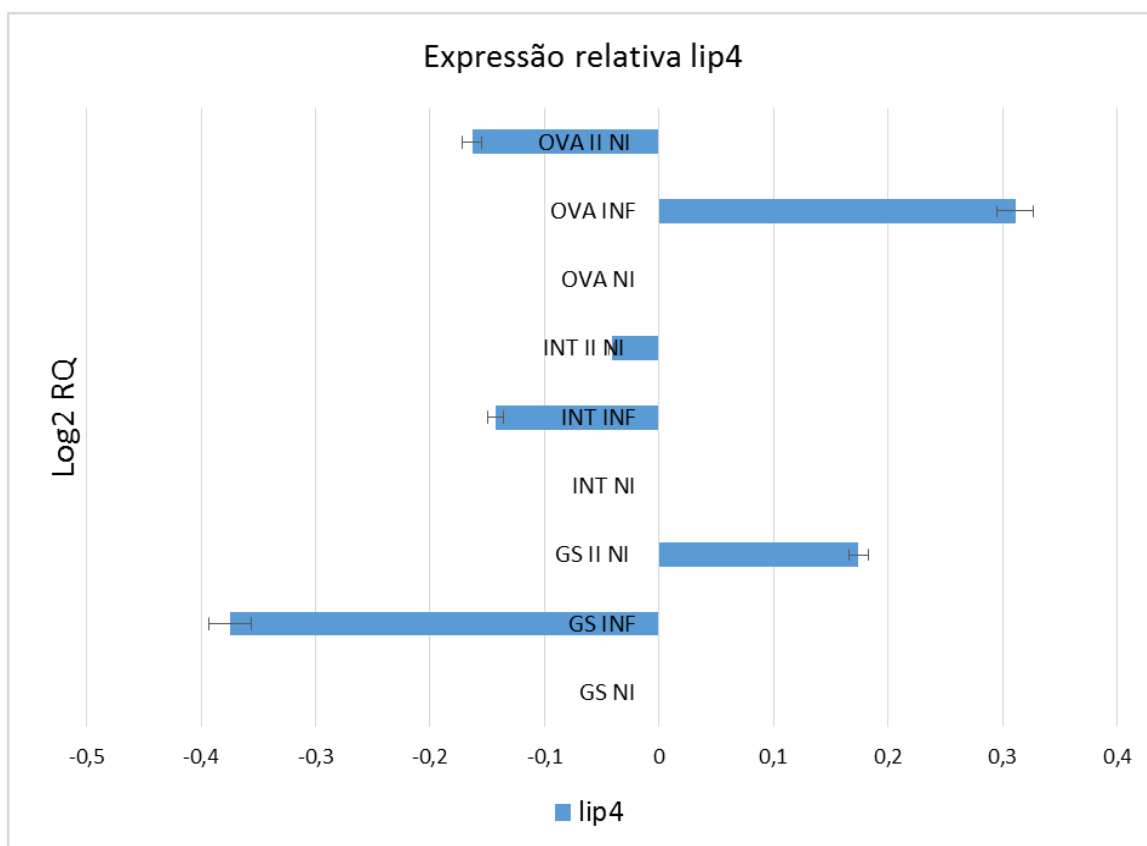


Figura 28: Expressão relativa do log2 RQ (*relative quantity*, RQ) do gene *lip4* de lipocalinas com *SHBP domain* em ovários de fêmeas do carrapato *A. sculptum*, confrontada com novas amostras de glândula salivar, intestino e ovário não infectado (GS II NI, INT II NI e OVA II NI), normalizados por *gapdh* em situação de não infecção/infecção com *R. amblyommii*.

A nova análise da expressão relativa do gene *lip4* no ovário utilizando o resultado de amostras não infectada resultou em *down* regulação de *lip4* no ovário (barra negativa de OVA II NI), diminuição da *down* regulação gênica na amostra de intestino não infectado (INT II NI) e *up* regulação de *lip4* na glândula salivar não infectada (barra positiva de GS II NI) (figura 28). Confrontando com as amostras de glândula salivar não infectada (GS II NI) e intestino não infectado (INT II NI) que tiveram sua expressão aumentada, a amostra de ovário não infectado sofreu *down* regulação. Assim através da figura 28, demonstra-se que somente no ovário o *lip4* encontra-se *up* regulado durante a infecção o que reforça a hipótese levantada pela análise da figura 27 de um possível papel multifuncional da proteína LIP4, relacionado à reprodução frente a infecção da fêmea do carrapato com *R. amblyommii*.

A figura 29 representa a expressão do gene *lip7* em glândulas salivares intestino e ovários de *A. sculptum* sob infecção de *R. amblyommii*.

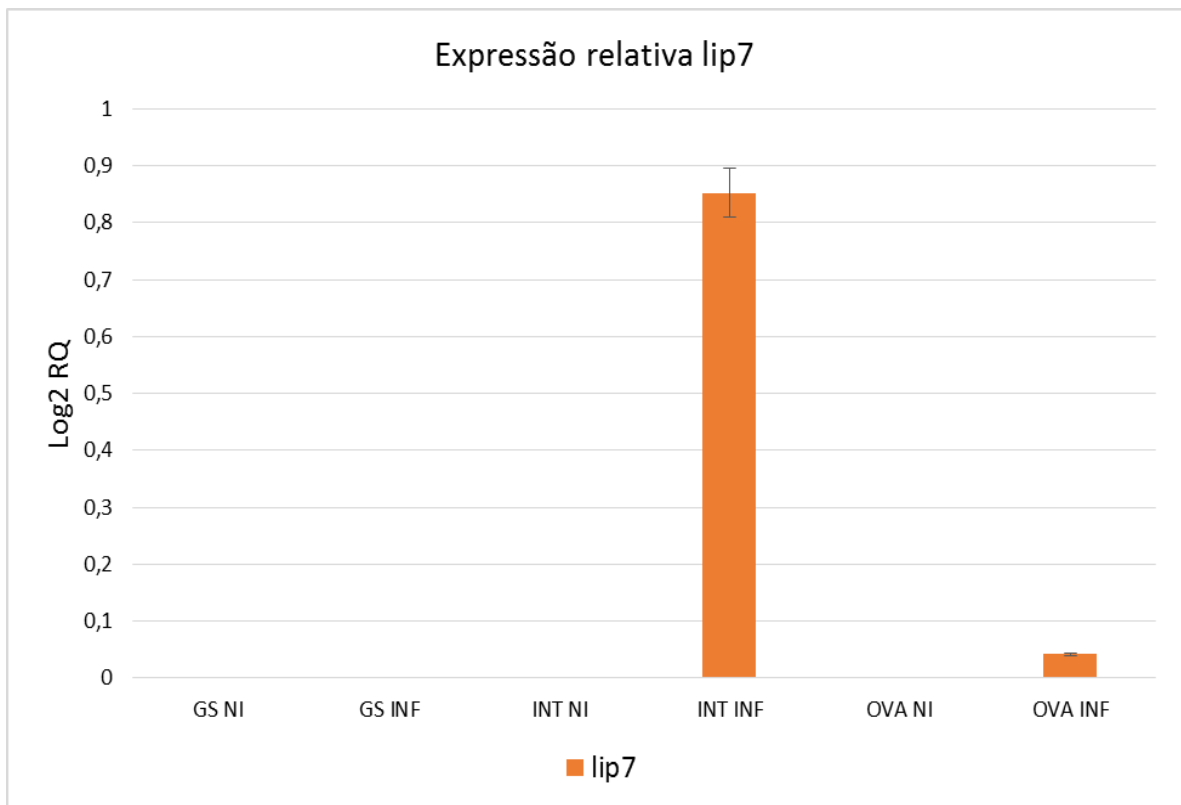


Figura 29: Expressão relativa do log2 RQ (*relative quantity*,RQ) do gene *lip7* de lipocalinas com *SHBP domain* em glândulas salivares, intestinos e ovários de fêmeas do carrapato *A. sculptum* normalizados por *gapdh* em situação de não infecção/infecção com *R. amblyommii*.

A análise da figura 29 apresenta o gene de lipocalina *lip7* diferentemente expresso tanto em glândulas salivares quanto em intestinos e ovários. A barra positiva observada no gráfico da figura 29 demonstra uma regulação positiva na expressão de *lip7*, no intestino e um pequeno aumento da expressão nos ovários infectados e a ausência do *lip7* em glândulas salivares infectadas e não infectadas. Sendo assim, pode-se inferir que a infecção superexpressa *lip7* no intestino. Para a verificação dessa hipótese realizou-se novo experimento utilizando-se novas amostras de intestino e ovário não infectado obtidas no mesmo tratamento de infecção da fêmea do carrapato *A. sculptum* (figura 30).

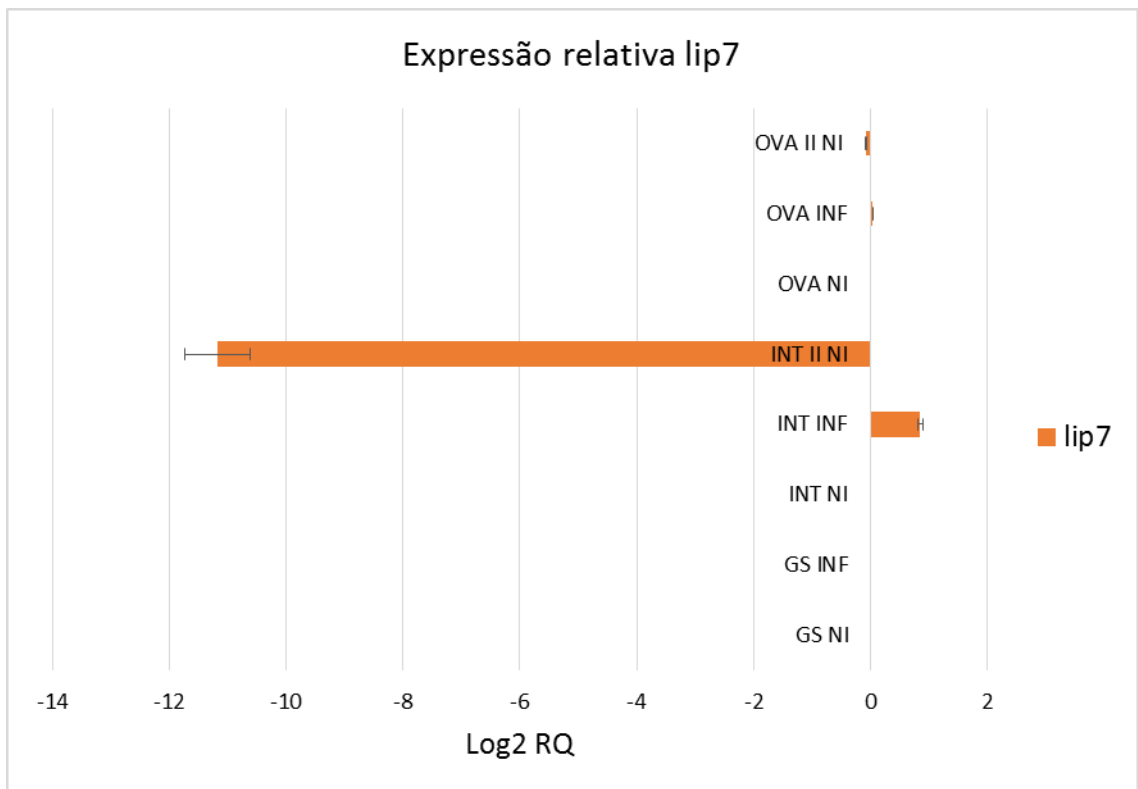


Figura 30: Expressão relativa do log₂ RQ (*Relative Quantity*,RQ) do gene *lip7* de lipocalinas com *SHBP domain* em ovários de fêmeas do carrapato *A. sculptum*, confrontada com novas amostras de intestino e ovário não infectado (INT II NI e OVA II NI), normalizados por *gapdh* em situação de não infecção/infecção com *R. amblyommii*.

Os resultados da expressão relativa das novas amostras não infectadas demonstraram que a infecção altera os níveis de expressão do gene *lip7* apenas no intestino infectado (figura 30) uma vez que a expressão gênica da nova amostra de intestino não infectado (INT II NI) resultou em *down* regulação e no ovário os níveis se mantiveram inalterados. Esse resultado reforça a hipótese de que a infecção por *R. amblyommii* superexpressa *lip7* no intestino devido a uma função protetora da proteína LIP7 sugerindo um papel multifuncional das lipocalinas no intestino.

A figura 31 representa a expressão do gene *lip24* em glândulas salivares intestino e ovários de *A. sculptum* sob infecção de *R. amblyommii*.

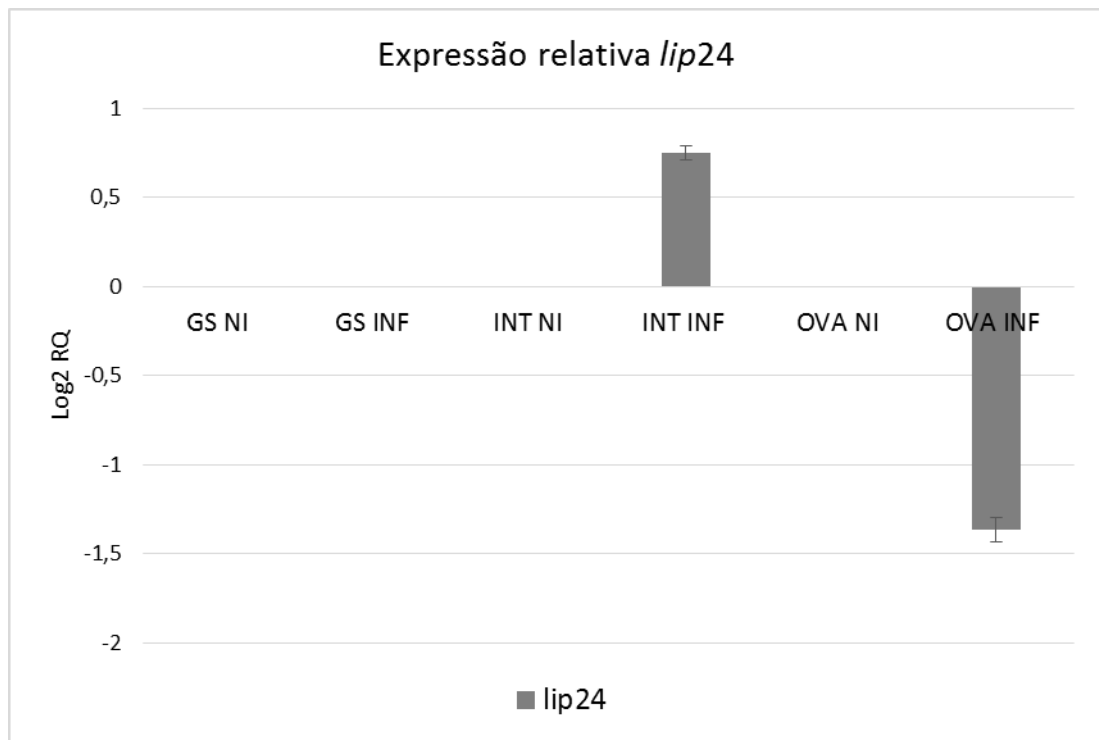


Figura 31: Expressão relativa do log2 RQ (*Relative Quantity*, RQ) do gene *lip24* de lipocalinas com *SHBP domain* em glândulas salivares, intestinos e ovários de fêmeas do carrapato *A. sculptum* normalizados por *gapdh* em situação de não infecção/infecção com *R. amblyommii*.

O resultado da expressão relativa demonstra que o gene *lip24* não está presente nas glândulas salivares tanto na ausência quanto na presença da infecção. Em relação ao intestino, a infecção provoca *up* regulação na expressão gênica de *lip24* e *down* regulação no ovário. A *up* regulação de *lip24* no intestino leva a hipótese de uma possível função protetora desse gene como função adicional das lipocalinas nesse órgão, uma vez que o carrapato não possui enzimas digestivas no intestino fazendo desse órgão um bom alvo para infecções bacterianas. Desse modo a captação de lipídeos no intestino infectado pode contribuir para a defesa da infecção do carrapato frente à bactéria *R. amblyommii*. E ainda, a *down* regulação de *lip24* no ovário sugere que esse gene pode estar envolvido com a reprodução do carrapato (figura 31). Para a verificação dessas hipóteses realizou-se novo experimento utilizando-se novas amostras de intestino e ovário não infectado obtidas no mesmo tratamento de infecção da fêmea do carrapato *A. sculptum* (figura 32).

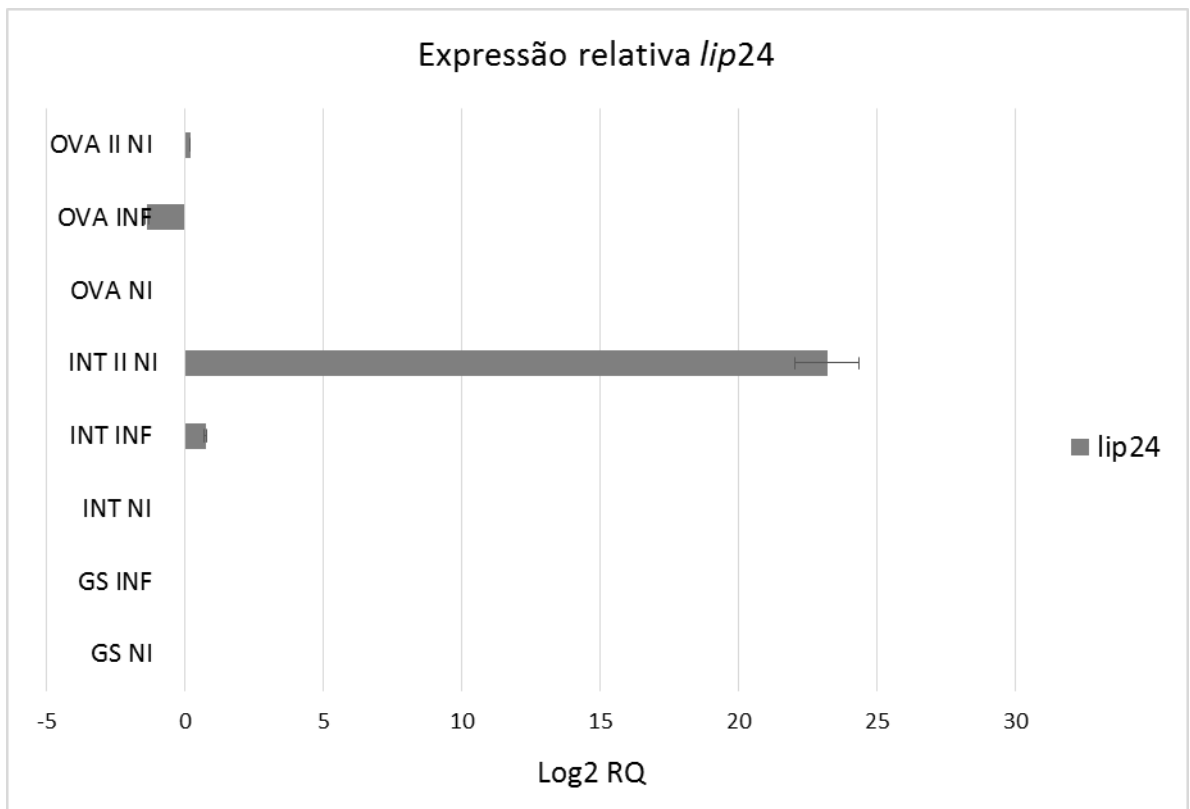


Figura 32: Expressão relativa do log2 RQ (*Relative Quantity*, RQ) do gene *lip24* de lipocalinas com *SHBP domain* em glândulas salivares intestinos e ovários de fêmeas do carrapato *A. sculptum*, confrontada com novas amostras de intestino e ovário não infectado (INT II NI e OVA II NI), normalizados por *gapdh* em situação de não infecção/infecção com *R. amblyommii*.

A nova análise da expressão relativa do gene *lip24* em intestinos e ovários utilizando amostras não infectadas resultou em *up* regulação de *lip24* no intestino e diminuição da *down* regulação desse gene no ovário, o que se verifica com o aumento da expressão gênica na amostra de intestino não infectado (INT II NI) e redução da supressão gênica de *lip24* no ovário (redução da barra negativa de OVA II NI) (figura 32). Assim através da figura 32, demonstra-se que a infecção provoca uma *down* regulação no gene *lip24* durante a infecção o que reforça a hipótese da função protetora desse gene no intestino frente a infecção com *R. amblyommii*. E ainda, amostra não infectada de ovário não apresenta *down* regulação reforçando a hipótese da participação deste no processo de reprodução do carrapato.

De maneira geral, sabe-se que as glândulas salivares dos Ixodidas contribuem na fixação bem como na modulação da resposta imuno-inflamatória do hospedeiro, sendo necessário um arsenal de moléculas que contribuam no prolongamento do repasto e na modulação negativa da resposta imune. De fato, em análise morfológica da ultraestrutura de glândulas salivares de *A. cajennense* detectou-se a secreção de grânulos contendo lipídios livres e complexados com proteínas em células acinares do

tipo II (DENARDI et al., 2011), estas localizam-se no final do sistema de ramificação de ductos e respondem pela secreção de biomoléculas multifuncionais que agem na fixação de carrapato no hospedeiro (WALKER, 1985), digestão de tecidos (LAVOPIERRE & RIEK, 1955; WALKER, 1985), inibição dos mecanismos de defesa da cascata de coagulação e reações inflamatórias (RIBEIRO & MATHER, 1998; PAESEN et al., 1999).

A presença de lipocalinas no intestino pode se relacionar às funções de captação de lipídeos pois essas proteínas podem abrigá-los nos cálices hidrofóbicos formados pelas estruturas do β -barril. A ultra estrutura das células do intestino da fêmea de *A. cajennense* apresenta um complexo contendo diferentes organelas e inclusões citoplasmáticas responsáveis por funções adicionais neste órgão além da digestão, como estoque de reservas funcionais decorrentes da digestão de hemoglobinas e armazenamento de lipídios (BALASHOV, 1967; CAPERUCCI et al., 2010). Estudo do mialoma (transcriptoma de intestino) conduzido em *Ixodes ricinus*, identificaram um unigene (MPAAD93TF) participante do metabolismo de lipídeos envolvendo a formação da vitelina (proteína resultante da conversão da heme lipoproteína vitalogenina) no oócito (HEEKIN et al., 2013a; KOTSYFAKIS et al., 2015), nosso experimento levanta a hipótese da participação das lipocalinas com a presença de domínio SHBP no combate a infecção intestinal e na reprodução do carrapato *A. sculptum*.

Em relação ao intestino, os lipídeos se internalizam pela ação da vitalogenina, uma proteína homóloga a Apolipoproteína B100 (ApoB) de mamífero, sendo a mesma um dos transcritos mais identificado em mialoma de carrapatos (SHELNESS & LEDFORD, 2005; HEEKIN et al., 2013a). A vitalogenina apresenta expressão diferencial em bovinos infestados por *R. microplus* com *Babesia bovis* fato associado a maior necessidade de metabolismo de lipídeos dos animais infectados (HEEKIN et al., 2013b). Esse tipo de expressão diferencial também ocorreu no *lip7* (*up* regulado pela infecção) e no *lip24* (*down* regulado) em decorrência da infecção por *R. amblyommii* sugerindo correlação entre o aporte lipídico e o combate à infecção uma vez que a ausência de enzimas digestivas no lúmen intestinal corrobora para o desenvolvimento de infecções nesse órgão.

Em relação a expressão relativa no ovário serão necessários estudos adicionais da influência da expressão do gene *lip4* em ovários infectados, que venham a contribuir no melhor entendimento da interação microrganismo-hospedeiro visto que dos genes de lipocalinas estudados, somente o *lip4* apresentou um sugestivo aumento da

expressão gênica quando ocorre a infecção ovariana. Devido à transmissão transovariana apresentada pelos carrapatos, uma das estratégias de contenção da mesma consiste na redução da população do artrópode por diminuição das taxas de reprodução, sendo assim nossos dados apresentam o gene *lip4* como possível participante da redução da prole, uma vez que, a infecção provoca a redução da reprodução como estratégia de defesa à infecção transovariana.

Estudos realizados em ovários de carrapatos *R. microplus* infectado por *B. bovis* apresentaram diminuição dos processos reprodutivos de oogênese e embriogênese relacionados e expressão diferencial das proteínas ovarianas dentre elas inibidores de serino-proteases, calreticulina e isômeros *cis* e *trans* da peptidil-prolil isomerase, além do aumento da glutathione s-transferase (GST) que auxiliam na digestão tanto na redução de proteínas e lipídeos quanto detoxificação das espécies reativas de oxigênio (ROS) durante o repasto sanguíneo (HEEKIN et al., 2013b). Salienta-se o fato que o gene *lip4* foi o único dentre os estudados que apresentou *up* regulação na presença de infecção ovariana, sendo assim, a presença o aumento da expressão gênica de *lip24* no ovário pode sugerir uma adaptação evolutiva no controle da infecção para a espécie *Amblyomma sculptum*.

O estudo do papel da *lip4* nos diferentes tecidos infectados e adicionalmente da proteína LIP4 recombinante, pode tornar-se uma estratégia para o desenvolvimento de vacinas ou abordagens de controle desse artrópode uma vez que nosso trabalho sugere uma participação desse gene no processo de reprodução visto que as fêmeas utilizadas estavam todas ingurgitadas. Quanto a presença da *up* regulação de *lip7* e *lip24* no intestino infectado apresentado em nosso trabalho, levanta-se a hipótese da participação destes genes na proteção frente à infecção bacteriana. Uma abordagem semelhante em mialoma levou à descoberta da Bm86, uma molécula da qual se desenvolveu a vacina bovina contra *R. microplus* estas proteínas agem como antígeno levando a imunoproteção do hospedeiro pois os carrapatos têm sua digestão dificultada quando ocorre o repasto uma vez que anticorpos no sangue do hospedeiro presentes no sangue ingerido atacam a Bm86 no intestino dos artrópodes tornando-os incapazes de digerir o sangue ingerido (WILLADSEN, 2004). Caso o papel dos genes *lip7* e *lip24* sejam semelhantes a Bm86, a mesma estratégia pode ser utilizada.

7.6.1 Análise dos genes constitutivos candidatos a genes de referência

Em nosso trabalho foram analisados a expressão gênica de 3 genes constitutivos *g6pdh*, *3ohcoadh* e *gapdh* (dois pares de *primers* para cada gene),

potenciais candidatos a genes de referência em glândulas salivares, intestinos e ovários do carrapato *A. sculptum* em situação de não infecção/infecção com *R. amblyommii*, uma vez que até o momento não existe trabalhos validando genes de referência para nenhum tecido do carrapato *A. sculptum*.

O estudo da expressão relativa dos genes constitutivos *g6pdh*, *3ohcoadh* e *gapdh* em glândulas salivares intestinos e ovários de fêmeas de *A. sculptum* resultaram na construção de gráficos dos valores de Cp analisados pelos softwares estatísticos *Normfinder* e *Bestkeeper* de maneira a determinar o melhor gene candidato a controle endógeno (gene de referência) (tabela 12)

Tabela 12: Estatística descritiva dos 3 genes candidatos a genes de referência calculados pelo *Bestkeeper* baseados nos valores de *crossing point* (Cp).

Gene	<i>g6pdh</i>	<i>g6pdh</i>	<i>3-ohcoadh</i>	<i>3-ohcoadh</i>	<i>gapdh</i>	<i>gapdh</i>	<i>Bestkeeper</i>
primer	<i>hk1</i>	<i>hk 2</i>	<i>hk 3</i>	<i>hk 4</i>	<i>hk 5</i>	<i>hk6</i>	
n	9	9	9	9	9	9	3
geo mean [Cp]	26,04	22,71	29,31	26,06	25,74	25,68	25,86
ar mean [Cp]	26,06	22,83	29,76	26,14	25,75	25,7	25,88
min [Cp]	24,59	20,69	22,15	23,40	24,33	23,66	23,95
max [Cp]	26,95	26,96	33,50	27,95	26,38	26,52	26,74
std dev [± Cp]	0,89	2,22	4,64	1,74	0,31	0,85	0,91
CV [% Cp]	3,43	9,73	15,57	6,66	1,22	3,31	3,51
min [x-fold]	-2,73	-4,04	-142,70	-6,38	-2,66	-4,05	3,74
max [x-fold]	1,87	18,98	18,27	3,68	1,56	1,79	1,85
std dev [± x-fold]	1,86	4,66	24,85	3,34	1,24	1,8	1,88
<i>p-value</i>	ns	ns	0,001*	0,001*	0,210	0,006*	na

Abreviações: n: número de amostras; Cp: *Crossing point*; *: valores significativos; ns: não significativo; na: não aplicável; *geo mean* [Cp]: média geométrica do *crossing point*; *ar mean* [Cp]: média aritmética do Cp; Min [Cp]: valor mínimo de Cp; Max [Cp]: valor máximo de Cp; std dev [± Cp]: desvio padrão do Cp; CV [% Cp]: coeficiente de variação do Cp expresso em porcentagem; min [x-fold]: e max [x-fold]: valores extremos dos níveis de expressão (alto e baixa) representado pelo coeficiente de regulação absoluto; std dev [± x-fold]: desvio padrão do coeficiente de regulação absoluto.

A análise do Cp dos genes candidatos pelo aplicativo *bestkeeper* resultou na escolha do *3ohcoadh* e *gapdh* como ideais para os alvos de lipocalinas (tabela 12). Os primers *hk3* e *hk4* (pertencentes ao gene *3-ohcoadh*) e *hk6* (pertencente ao *gapdh*) apresentaram um *p-value* significativo ($p < 0,01$) representando um nível de significância de 99%. Os primers *hk3* e *hk4*, apesar de serem significativos, apresentaram uma variação maior que *hk5* e *hk6* (CV % e std dev [± x-fold]) Sendo assim as análises foram normalizadas para o *gapdh* (primer *hk6*) pois o mesmo apresentou um melhor resultado, tanto na análise estatística quanto na otimização do experimento (teste de concentração e eficiência de primers).

Após a escolha do melhor gene de referência (*gapdh*) buscou-se estabelecer se a variação da expressão dos genes candidatos nos diferentes tecidos e tratamento era

significativamente diferente do *gapdh*, assim, aplicou-se aos valores de Cp obtidos no experimento, os testes de *Two-way* ANOVA de maneira a comparar os Cp dos genes *g6pdh* e *3-ohcoadh* com o *gapdh* para verificar a hipótese de uma variação significativa da expressão gênica (Figura 33).

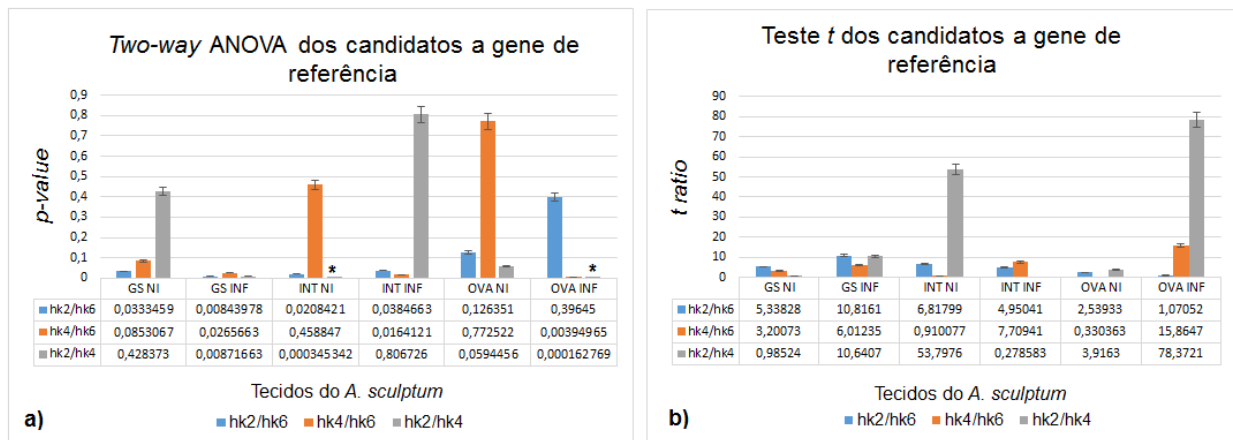


Figura 33: Resultado da análise de *Two-way* ANOVA (a) e teste *t* (b) dos valores de Cp dos genes candidatos a genes de referência.

*: indica nível de significância < 0,001 (*p-value*).

A análise estatística da figura 33 representa os genes constitutivos *g6pdh*, *3ohcoadh* e *gapdh* candidatos a gene de referência para estudos de expressão gênica em glândulas salivares, intestinos e ovários de fêmeas alimentadas e ingurgitadas do carrapato *A. sculptum* no contexto de infecção e não infecção com *R. amblyommii*.

A figura 33 (a) revela diferenças significativas na expressão gênica dos genes candidatos a gene de referência quando expressos em intestino e ovário em carrapatos fêmeas de *A. sculptum*. Essa diferença ocorre entre os níveis de expressão dos genes *hk2* quando comparado com *hk4* em intestino não infectado e ovário infectado (figura 33 (a)). Há de se considerar que apesar do teste *t* (Figura 33 (b)) não apresentou diferenças significativas em nenhum dos tecidos analisados, porém o teste de *two-way* ANOVA resultou na rejeição da hipótese nula em intestino não infectado (INT NI, figura 33 (a)) e ovário infectado (OVA INF, figura 33 (a)). Portanto, a análise estatística da figura 33 (a) revela que o primer *hk2* e *hk4* possuem níveis de expressão diferentes nos tecidos estudados apresentando um nível de significância (*p-value* < 0,0001) quando comparados entre si. Assim, apesar do teste *t* validar a hipótese nula, ou seja todos os genes candidatos poderiam ser utilizados pois não há diferença significativa da expressão dos mesmos nos tecidos e contexto analisado (infecção e não infecção), o teste de *Two-way* ANOVA que possui um nível descritivo maior que o teste *t* apresentou valores significativos em relação a expressão no intestino não infectado (INT NI, *p*-

value = 0,00034532, figura 33 (a)) e no ovário infectado (OVA INF *p-value* = 0,000162769, figura 33 (a)). Conclui-se que para estudos da expressão gênica dos genes candidatos a gene de referência quando expressos em glândula salivar, intestino e ovário em carrapatos fêmeas alimentados no contexto de infecção e não infecção com *R. amblyommii* pode-se utilizar se necessário os genes *g6pdh* e *3-ohcoadh* como um segundo controle endógeno quando o experimento ocorrer em glândula salivar, intestino e ovário desde que o primeiro controle seja o *gapdh*.

A fim de validar os *primers* como controle endógeno primário ou secundário para estudos em ausência de infecção foi realizado outra análise de *Two-way* ANOVA de maneira a comparar a expressão de cada primer em diferentes tecidos somente no contexto não infectado e dessa maneira validar o uso desses *primers* para a amplificação dos genes de lipocalinas nessa situação (figura 34).

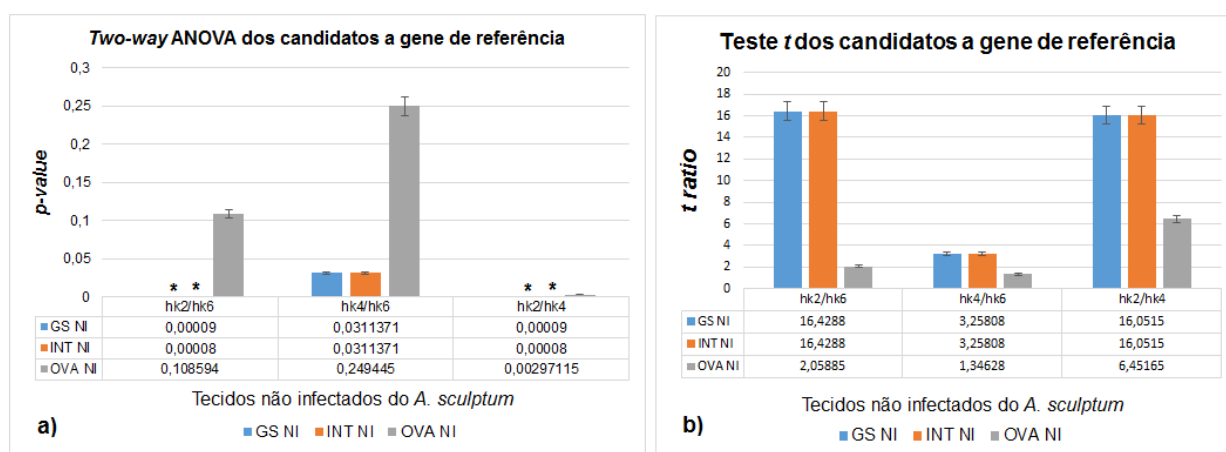


Figura 34: Resultado da análise de *Two-way* ANOVA (a) e teste *t* (b) dos valores de Cp dos genes candidatos a genes de referência em tecido não infectado. *: indica nível de significância < 0,001 (*p-value*).

A análise estatística da figura 34 (a) revelou diferenças significativas na expressão gênica dos genes candidatos a gene de referência quando expressos em glândula salivar não infectada (GS NI, figura 34 (a)) e ovário não infectado (OVA NI, figura 34 (a)) da fêmea do carrapato *A. sculptum* infectada com *R. amblyommii*. Cabe ressaltar que tal diferença novamente não apareceu pelo teste *t* (Figura 34 (b)) em nenhum dos tecidos não infectados analisados. Novamente o teste de *two-way* ANOVA serviu para avaliar a aceitação da hipótese nula resultante da análise dos resultados de Cp pelo teste *t*. Sendo assim, os *primers* *hk2* e *hk6* em glândula salivar possuem níveis de expressão diferentes nos tecidos estudados apresentando um nível de significância (*p-value* < 0,0001) quando comparados ao primer *hk6* correspondente ao *gapdh*, gene de referência de escolha em nosso estudo. Assim como ocorreu na figura

33 (b) o teste *t* não detectou diferenças significativas na expressão relativa entre os genes *hk2* e *hk6* nos tecidos não infectados analisados o que poderia levar à uma conclusão equivocada quanto ao uso do mesmo em detrimento do *gapdh*, em experimentos nesses tecidos. Porém ao executarmos a *Two-way* ANOVA (valores de *p-value* da figura 34 (b)), o mesmo resulta na discriminação significativa dos valores de Cp em relação a expressão em glândula salivar não infectado (GS NI, *p-value* < 0,0001, figura 34 (a)) na comparação entre os genes *hk2* e *hk6* e no ovário infectado (OVA INF *p-value* = 0,000162769, figura 34 (a)). E ainda, quando comparamos o primer *hk4* com o *hk6* a mesma resulta em diferenças significativas tanto em GS NI quanto em INT NI portanto apesar do teste *t* validar a hipótese nula (figura 34 (b)) pois nenhum valor de *t ratio* (figura 34 (b)) foi menor que 0,05 (indicando significância para os parâmetros analisados), assim, a hipótese nula foi rejeitada devido aos resultados do teste de *two-way* ANOVA. Portanto, os primers *hk2* e *hk4* que correspondem a *g6pdh* e *3-ohcoadh* respectivamente, só devem ser usados como genes de referência em tecidos não infectados no caso de experimento a se realizar no intestino do carrapato *A. sculptum* (INT NI, figura 34 (a) e (b)) pois nesse tecido não há diferença significativa na expressão dos 3 genes estudados.

Em relação aos resultados apresentados nesse tópico observa-se que o gene de escolha para estudos de nível de expressão gênica nos tecidos e contexto estudados quer por uma regulação positiva ou negativa foi o *gapdh*. Até o momento poucos são os genes analisados e validados como genes de referência para o gênero *Amblyomma*. Em estudo recente foi analisado cinco genes candidatos a gene de referência gliceraldeído-3-fosfato desidrogenase (*gapdh*), β -actina, α -tubulina, fator de alongação 1 e glutatona s-transferase (GST) esse estudo apontou a viabilidade do uso da β -actina e do *gapdh* (BROWNING, 2012). Até o momento, os genes *3ohcoadh* e *g6pdh* não foram analisados e nem validados como genes de referência para o gênero *Amblyomma* e apesar de serem constitutivos não são utilizados para essa finalidade. O *g6pdh* foi analisado em estudo com *Ixodes scapularis* porém nessa espécie os resultados foram menores que os demais genes estudados. O *gapdh* é um dos genes de referência mais utilizados para o estudo da expressão pela grande maioria dos trabalhos publicados das mais variadas espécies (De JONGE et al., 2007). Cabe ressaltar em relação a esse gene que o mesmo sofre variabilidade na expressão frente a algumas condições tais como hipóxia e condrogênese (modulação positiva), osteogênese e adipogênese quebrando o mito desse gene “clássico” e demonstrando assim a necessidade da validação do mesmo em cada tecido e condição estudada.

De modo geral, observa-se que as ações anti-histamínica e anticoagulante apresentadas pelas lipocalinas, associadas à sua capacidade de abrigar substâncias hidrofóbicas, no caso ocorrendo o sequestro tanto da serotonina quanto da histamina, contribuem na supressão da resposta do hospedeiro. Na literatura existe a caracterização de domínios de ligação de amina biogênica (BAB) presentes tanto em carrapatos Ixodidaeos quanto Argasidaeos (BEAUFAYS et al., 2008), isso demonstra que essas proteínas desempenham papel importante para a alimentação do artrópode e conseqüentemente para o desenvolvimento do seu ciclo de vida.

Quanto ao estudo da expressão de lipocalinas sugere-se a necessidade de trabalhos envolvendo os genes *lip 4*, *7* e *24* em situação de *Knock out* ou *Knock down* de maneira a esclarecer os possíveis papéis de cada gene estudado na função anti-histamínica exercida durante a alimentação do artrópode sobre o hospedeiro, bem como, relacionar o quanto a infecção corrobora na expressão diferencial desses genes nos tecidos estudados.

Quanto aos genes de referência, para a validação destes deve-se observar algumas recomendações, o Cp não deve ser muito alto (<15) ou baixo (>30) (WAN e al., 2010) e deve se estabelecer a teste de concentração e curva de eficiência (KOZERA & RAPACZ, 2013). Estes testes foram realizados em nosso trabalho sugerindo que, para estudo da expressão em glândulas salivar e intestinos e ovários o gene constitutivo analisado e validado como gene de referência em *Amblyomma sculptum* é o *gapdh*, pois o mesmo manteve melhores resultados e foi o gene de escolha apontado pelo programa bestkeeper para controle endógeno dos tecidos estudados em situação de infecção e não infecção por *R. amblyommii*, no caso da utilização de um segundo controle endógeno recomenda-se o *3ohcoadh* devido à menor variação quando comparado com o *g6pdh*.

8. CONCLUSÃO

Foram identificados em nosso trabalho nove transcritos de lipocalinas no transcriptoma de glândulas salivares de *A. sculptum* possuidoras de domínio de ligação a histamina e serotonina (*lip1*, *lip4*, *lip7*, *lip10*, *lip14*, *lip17*, *lip21*, *lip24* e *lip27*). Estas quando confrontadas com as lipocalinas depositadas no *Swiss-prot* relacionadas à sua sinonímia, *A. cajennense*, encontram-se depositadas, à exceção do *lip17* (As_6240), recomenda-se o depósito deste gene identificado no transcriptoma de Moreira 2010.

Foram analisados e validados 3 *primers* de genes constitutivos (*g6pdh*, *3ohcoadh* e *gapdh*) candidatos potenciais a genes de referência em glândulas salivares, intestinos e ovários do carrapato *A. sculptum* em situação de não infecção/infecção com *R. amblyommii*. Para gene de referência em glândulas salivares, intestino e ovários recomenda-se o *gapdh* e se necessário um segundo controle endógeno o gene de escolha seria o *3-ohcoadh*.

Foram validados 3 modelos estruturais correspondentes às proteínas lipocalina 4, lipocalina 7 e lipocalina 24. A partir dessa validação *in silico* torna-se necessário uma obtenção da proteína recombinante das proteínas estudadas seguida da análise através de ressonância magnética nuclear e cristalografia por raioX de maneira a comprovação experimental da estrutura modelada.

Quanto ao estudo da expressão relativa, o *lip4* teve sua expressão *down* regulada em glândulas salivares infectadas e no intestino infectado. Cabe ressaltar que a proteína LIP4 tem função distinta quando presente em glândula salivar e intestino uma vez que a infecção provocou *down* regulação de *lip4* em glândula salivar e intestino e *up* regulação de *lip4* em ovário. Somente no ovário o *lip4* encontra-se *up* regulado durante a infecção o que se infere que existe um papel multifuncional da proteína LIP4, relacionado à reprodução frente a infecção da fêmea do carrapato com *R. amblyommii*.

O gene de lipocalina *lip7* encontra-se diferentemente expresso tanto em glândulas salivares quanto em intestinos e ovários. O mesmo apresentou uma regulação positiva na sua expressão no intestino e um pequeno aumento da expressão nos ovários infectados, em glândulas salivares tanto infectada como não infectada, o esse gene encontra-se suprimido. Conclui-se que a infecção superexpressa *lip7* no intestino infectado reforçando a hipótese de que a superexpressão frente à infecção por *R. amblyommii* no intestino deve-se a uma função protetora da proteína LIP7 sugerindo um papel multifuncional dessa lipocalina no intestino.

O resultado da expressão relativa do gene *lip24* demonstra que o mesmo não está presente nas glândulas salivares tanto na ausência quanto na presença da infecção. Em relação ao intestino, a infecção provoca *up* regulação na expressão gênica de *lip24* e *down* regulação no ovário. No intestino aventa-se uma possível função protetora desse gene como função adicional das lipocalinas nesse órgão, uma vez que o carrapato não possui enzimas digestivas no intestino fazendo desse órgão um bom alvo para infecções bacterianas. A captação de lipídeos no intestino realizada pela função principal do *lip24* pode contribuir para a defesa da infecção do carrapato frente à bactéria *R. amblyommii*. No ovário, sua *down* regulação sugere possível envolvimento desse gene na reprodução do carrapato pois o *lip24* encontra-se *down* regulado somente em caso de infecção.

Portanto, nosso trabalho com as lipocalinas sugere a necessidade de estudos posteriores envolvendo os genes *lip4*, *lip7* e *lip24* em situação de *Knock out* ou *Knock down* de maneira a esclarecer os possíveis papéis de cada gene estudado na função anti-histamínica exercida durante a alimentação do artrópode sobre o hospedeiro, bem como, relacionar o quanto e porquê a infecção corrobora na expressão diferencial desses genes nos tecidos estudados pois foi demonstrado que pode haver papel multifuncional das lipocalinas com domínio SHBP nos tecidos estudados. Conclui-se que os dados obtidos na análise de expressão gênica por *real time* RT-qPCR suportam a hipótese de que os três genes de lipocalinas interferem na interação microrganismo-vetor sendo necessários trabalhos adicionais para verificação da validade desta hipótese, uma vez que dentre as 121 lipocalinas preditas no transcriptoma de glândulas salivares, somente nove apresentam domínios SHBP sugerindo um papel especial na função destas nos tecidos estudados.

Finalmente, através do nosso trabalho de identificação de lipocalinas com domínio de ligação a histamina e serotonina e validação de genes de referência em transcriptomas de glândula salivar, intestino e ovário do carrapato *Amblyomma sculptum* reforçamos a hipótese de uma conservação evolutiva da função anti-histamínica por meio da participação de lipocalinas na modulação da resposta do hospedeiro. Estas, durante o repasto sanguíneo do carrapato, apesar da existência de receptores cutâneos de histamina no hospedeiro vertebrado, atuam reduzindo a reação cutânea inflamatória, prolongando o período de alimentação do artrópode.

Estudos experimentais para determinar a atividade anti-histamínica apresentada pela saliva do carrapato *A. sculptum*, bem como, a investigação de uma possível interação desta atividade com receptores celulares associados com o padrão de

conservação de sua estrutura terciária, virão constituir uma estratégia potencialmente sólida como base para o desenvolvimento de biofármacos. Neste contexto, o potencial da ação anti-histamínica apresentada pelas lipocalinas do *A. sculptum*, em especial *lip4* nos ovários infectados e *lip7* e *lip24* nos intestinos infectados associado à sua capacidade de abrigar substâncias hidrofóbicas no interior de sua estrutura terciária, torna estas moléculas candidatas a produto de uso farmacêutico como estratégia de controle dos índices de reprodução ou na resistência à infecção do carrapato *A. sculptum* frente a bactérias em geral em especial às riquetsias.

9. REFERÊNCIAS BIBLIOGRÁFICAS

- ABASCAL, F.; ZARDOYA, R.; POSADA, D. 2005. ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics.*, 21(9):2104-2105.
- ALJAMALI, M.N.; HERN, L.; KUPFER, D.; DOWNARD, S.; SO, S. ROEB, B.A.; SAUER, J.R.; ESSENBERG, R.C. 2009. Transcriptome analysis of the salivary glands of the female tick *Amblyomma americanum* (Acari: Ixodidae). *Insect. Mol. Biol.*, 18:129-154.
- ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. 1990. Basic local alignment search tool. *J Mol Biol*, 215(3):403-410.
- ALTSCHUL, S.F.; MADDEN, T.L.; SCHÄFFER, A.A. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, 1997;25:3389–3402.
- ANATRIELLO, E.; FERREIRA, B. R.; BRANDÃO, L. G.; VALENZUELA, J. G.; RIBEIRO, J. M.; SILVA, J. S. D.; MIRANDA SANTOS, I. K. F. 2009. Comparative analysis of transcriptomes of salivary glands from ticks, *Amblyomma cajennense* and *Rhipicephalus sanguineus*. *Vet. Immunol. and immunopath.*, 128(1-3):223, 2009.
- ANDERSEN, J.F.; GUDDERRA, N.P.; FRANCISCHETTI, I.M.; RIBEIRO, J.M. 2005. The role of salivary lipocalins in blood feeding by *Rhodnius prolixus*. *Arch. Insect. Biochem. Physiol.*, 58:97-105.
- ARAGÃO, H.B.; FONSECA, F. 1961. Notas de Ixodologia. VIII. Lista e chave para os representantes da fauna ixodológica brasileira. *Mem. Inst. Oswaldo Cruz.*, 59:115-129.
- ASTIMO, P.; JONNALAGEDA, M.; ARNOLD, K.; BARATIN, D.; CSARDI, G.; DE CASTRO, E.; DUVAUD, S.; FLEGEL, V.; FORTIER, A.; GASTEIGER, E.; GROSDIDIER, A.; HERNANDEZ, C.; IONNIDIS, V.; KUZNETSOV, D.; LIECHTI, R.; MORETTI, S.; MOSNTAGUIR, K.; REDASCHI, N.; ROSSIER, G.; XENARIOS, I.; STOCKINGER, H. 2012. ExPASy: SIB bioinformatics resource portal, *Nucleic Acids Res.*, 40(W1):W597-W603.
- BALASHOV, Y.S. 1972. Bloodsucking ticks (Ixodoidea) - vectors of disease of man and animals. *Miscel. Pub. Entomol. Society of America*, 8:163-376.
- BATISTA, I. F.; CHUDZINSKI-TAVASSI, A. M.; FARIA, F.; SIMONS, S. M.; BARROS-BATESTTI, D. M.; LABRUNA, M. B.; LEAO, L. I.; HO, P. L.; JUNQUEIRA-DE-AZEVEDO, I. L. 2008. Expressed sequence tags (ESTs) from the salivary glands of the tick *Amblyomma cajennense* (Acari: Ixodidae). *Toxicon.*, 51(5):823-34.
- BEATI, L.; NAVA, S.; BURKMAN, E. J.; BARROS-BATTESTI, D. M.; LABRUNA, M. B.; GUGLIELMONE, A. A.; CACERES, A. G.; GUZMAN-CORNEJO, C. M.; LEON, R.; DURDEN, L. A.; FACCINI, J. L. 2013. *Amblyomma cajennense* (Fabricius, 1787) (Acari: Ixodidae), the Cayenne tick: phylogeography and evidence for allopatric speciation. *BMC evolutionary biology*, 13:267-272.
- BEAUFAYS, J.; ADAM, B.; DECREM, Y.; PREVOT, P.P.; SANTINI, S.; BRASSEUR, R.; BROSSARD, M.; LINS, L.; VANHAMME, L.; GODFROID, E. 2008. *Ixodes ricinus* tick

lipocalins identification, cloning, phylogenetic analysis and biochemical characterization. *PLoS One*, 3(12):e3941.

BERG, J.; TYMOCZKO, J.; STRYER, L. 2002. Biochemistry 5th ed. W.H. Freeman, 894p.

BROWNING, R.; ADAMSOM, S.; KARIM, S. 2012. Choice of a stable set of reference genes for qRT-PCR analysis in *Amblyomma maculatum* (Acari: Ixodidae). *J. Med. Entol.*, 49(6):1939-1346.

BRUNNER, A.M.; YAKOVLEV, I.A.; STRAUSS S.H. 2004. Validating internal controls for quantitative plant gene expression studies., *BMC Plant Biology*, 4:14.

BUSK, K.P. 2014. A tool for design of primers for microRNA-specific quantitative RT-qPCR. *BMC Bioinformatics* 15(29):1-9.

BUSTIN, S.A.; BENES, V.; GARSON, J.A.; HELLEMANS, J.; HUGGETT, J.; KUBISTA, M.; MUELLER, R.; NOLAN, T.; PFAFFL, M.W.; SHIPLEY, G.L.; VANDESOMPELE, J.; WITTEWER, C.T. 2009. The MIQE guidelines: minimum information for publication of quantitative real time PCR experiments. *Clin. Chem.*, 55(4):611-622.

CAPERUCCI, D.; BECHARA, G.H.; MATHIAS, M.I.C. 2010. Ultrastructure features of the midgut of the female adult *Amblyomma cejennense* ticks Fabricius, 1787 (Acari: Ixodidae) in several feeding stages and subjected to three infestations. *Micron*, 41(7):710-721.

CARDOSO, L.D. 2004. Detecção e caracterização de *Rickettsia* ssp. Circulante em foco inativo peri-urbano do município de Caratinga, MG. *Dissertação de Mestrado*, Universidade Federal de Ouro Preto. 127p.

CHERVONEVA, I.; LI, Y.; SCHULZ, S.; CROKER, S.; WILSON, C.; WALDMAN, S.A.; HYSLOP, T. 2010. Selection of optimal reference genes for normalization in quantitative RT-PCR. *BMC Bioinf.* 11:253.

CHMELAR, J.; ANDERSON, J.M.; MU J.; JOCHIM, R.C.; VALENZUELA, J.G.; KOPECKÝ, J. 2008. Insight into the sialome of the castor bean tick *Ixodes ricinus*. *BMC Genomics*, 9:233.

CONLON, P.J.; PROCOP, G.W.; FOWLER, V.; ELOUBEIDI, M.A.; SMITH, S.R.; SEXTON, D.J. 1996. Predictors of prognosis and risk of acute renal failure in patients with Rocky Mountain spotted fever. *Am. J. Med.*, 101(6):621-626.

COULLIN, I.; VARGAFTIG, B.B.; JACOBS, M.; PAESEN, G.C.; NUTTALL, P.; MAILLET, I.; LEFORT, J.; MOSER, R.; WESTONDAVIES, W.; RYFFEL, B. 2004. Arthropod derived histamine binding protein prevents murine allergic asthma *J. Immun.*, 173(5):3281-3286.

CSERMELY, P.; KORCSMÁROS, T.; HUBA, J.M.K.; LONDON, G.; NUSSINOV, R. 2013. Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review. *Pharmacol. Ther.*, 138(3):333-408.

- DANTAS-TORRES, F.; ONOFRIO, V. C.; BARROS-BATTESTI, D. M. 2009. The ticks (Acari: Ixodida: Argasidae, Ixodidae) of Brazil. *System. & Appl. Acarol.*, 14:30-46.
- DAVIDSON, M.G.; BREITSCHWERDT, E.B.; WALKER, D.H.; LEVY, M.G.; CARLSON, C.S.; HARDIE, E.M.; GRINDEM, C.A.; NASSISSE, M.P. 1990. Vascular permeability and coagulation during *Rickettsia rickettsii* infection in dogs. *Am. J. Vet. Res.*, 51:165-170.
- DE JONGE, H.J.M.; FEHRMANN, R.S.N.; DE BONT, E.S.J.M.; HOFSTRA, R.M.W.; GERBENS, F.; KAMPSWA, F.; DE VRIES, E.G.E.; VAN DER ZEE, A.G.J.; TE MEERMAN G.J.; TER, E.L.S.T.A. 2007. Evidence based selection of housekeeping genes. *PLoS One*, 2(9):e898.
- DENARDI, S.E.; FURQUIM, K.C.S.; BECHARA, G.H.; CAMARGO-MATHIAS, M.I. 2011. Salivary glands of *Amblyomma cajennense* (Acari: Ixodidae): a historical and ultrastructural overview. *Exp. Appl. Acarol.*, 54:9421-9432.
- DHAR, A.K.; BOWERS, R.M.; LICON, K.S.; VEAZEY, G.; READ, B. 2009. Validation of reference genes for quantitative measurement of immune gene expression in shrimp. *Mol. Immunol.*, 46:1688-1695.
- DHEDA, K.; HUGGETT, J.F.; BUSTIN, S.A.; JOHNSONMA, F.; ROOK G.; ZUMLA, A. 2004. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *BioTech.*, 37:112-119.
- DIAMANT, G.; STRICKLAND, R.D. 1965. *Manual of livestock ticks for animal disease eradication division personnel*. Washington: USDA, Res. Serv., 141p.
- DIAS, E. L.; MARTINS, A. V. 1939. Spotted fever in Brazil. *Am. J. of Trop. Med.*, 19:103-108.
- DÍAZ-MARTIN, V.; MANZANO-ROMAN, R.; SILES-LUCAS, M.; OLEAGA, A.; PEREZ-SANCHEZ, R. 2011. Cloning characterization and diagnostic performance of the salivary lipocalin protein TSGP1 from *Ornithodoros moubata*. *Vet. Parasitol.*, 178(1-2):163-172.
- DONG, L.; SUI, C.; LIU, Y.; YANG, Y.; WEI, J. 2010. Validation and application of reference genes for quantitative gene expression analysis in various tissues of *Bupleurum chinense*. *Mol. Bio. Reports*, 38(8):5017-5023.
- EDDY, S.R. 1998. Profile hidden markov models. *Bioinformatics*, 14:755-763.
- FALUS, A. 1994. Histamine modulates the acute phase response at multiple points, *Immunol. Today*, 15(12):596.
- FENG, H.M.; POPOV, V.L.; WALKER, D.H. 1994. Depletion of gamma interferon and tumor necrosis factor alpha in mice with *Rickettsia conorii*-infected endothelium: impairment of rickettsicidal nitric oxide production resulting in fatal, overwhelming rickettsial disease. *Infect. Immun.*, 62:1952-1960.
- FIN, R.D.; BATEMAN, A.; CLEMENTS, J.; COGGILL, P.; EBERHARDT, R.Y.; EDDY, S.R.; HEGER, A.; HETHERINGTON, K.; HOLM, L.; MISTRY, J.; SONNHAMMER,

- E.L.L.; TATE, J., PUNTA, M. 2014. The Pfam protein families database. *Nucl. Acids. Res.*, 42:D222-D230.
- FLOWER, D.R. 1993. Structural relationship of streptavidin to the calycin protein superfamily. *FEBS*, 233(2):99-102.
- FLOWER, D.R. 1996. The lipocalin protein family: structure and function. *Biochem J* 318(Pt1):1-14.
- FLOWER, D.R. 2000. Beyond the superfamily the lipocalin receptors. *Biochim Biophys Acta*, 1482(1-2):327-336.
- FLOWER, D.R.; NORTH, A.C.T.; ATTWOOD, T.K. 1993. Structure and sequence relationships in the lipocalins and related proteins. *Prot. Sci.* 2:753-761.
- FLOWER, D.R.; NORTH, A.C.T.; SANSOM, C.E. 2000. The lipocalin protein family: structural and sequence overview. *Bioch. Biophys. Acta.*, 1482:9-24.
- FONSECA, L.M.G. & MARTINS, A.V. 2007. Febre Maculosa: Revisão de literatura. *Saúde & Amb. Rev.*, 2(1):1-20.
- FRANCISCHETTI, I.M.; MATHER, T.N.; RIBEIRO, J.M. 2003. Cloning of a salivary gland metalloprotease and characterization of gelatinase and fibrin (ogen) lytic activities in the saliva of the Lyme disease tick vector *Ixodes scapularis*. *Biochem Biophys Res Commun*, 305(4):869-875.
- GANFORNINA, M.D.; GUTIERREZ, G.; BASTIANI, M.; SANCHEZ, D. 2000. A phylogenetic analysis of the lipocalin protein family. *Mol Biol Evol*, 17(1):114-126.
- GARCIA, G. R.; GARDINASSI, L. G.; RIBEIRO, J. M.; ANATRIELLO, E.; FERREIRA, B. R.; MOREIRA, H. N.; MAFRA, C.; MARTINS, M. M.; SZABO, M. P.; DE MIRANDA-SANTOS, I. K.; MARUYAMA, S. R. 2014. The sialotranscriptome of *Amblyomma triste*, *Amblyomma parvum* and *Amblyomma cajennense* ticks, uncovered by 454-based RNA-seq. *Parasit. & vect.*, 7:430.
- GILLESPIE, J.J.; REIER, M.S.; RAHMAN, M.S.; AMMERMAN, N.C.; SHALLOM, J.M.; PURKAYASTHA, A.; SOBRAL, B.S.; AZAD, A.F. 2007. Plasmids and rickettsial evolution: insight from *Rickettsia felis*. *PLoS ONE* 2:e266.
- GIULIETTI, A.; OVERBERGH, L.; VALCKX, D.; DECALLONNE, B.; BOUILLON, R.; MATHIEU, C. 2001. An overview of real time quantitative PCR: Applications to quantify cytokine gene expression. *Methods*. 25:386-401.
- GRZYB, J.; LATOWSKI, D.; STRZALCKA, K. 2006. Lipocalins - a family portrait. *J Plant Physiol.*, 163(9):895-915.
- GUGLIELMONE, A.A.; ROBBINS, R.G.; APANASKEVICH, D.A.; PETNEY, T.N.; ESTRADA-PEÑA, A.; HORAK, I.G.; SHAO, R.F.; BAKER, S.C. 2010. The Argasidae, Ixodidae and Nuttalliellidae (Acari: Ixodida) of the world a list of valid species names. *Zootaxa*, 2528:1-28.

GUGLIEMONE, A.A.; BEATI, L.; BARROS-BATESTI, D.M.; LABRUNA, M.B.; NAVA, S.; VENZAL, J.M.; SZABÓ, M.P.J.; MARTINS, J.R.; GONZÁLEZ-ACUÑA, D.; ESTRADA-PEÑA, A. 2006. Ticks (*Ixodidae*) on humans in South America. *Exp. Appl. Acarol.*, 40:83-100.

HARRELL, G.T. & AIKAWA, J.K. 1949. Pathogenesis of circulatory failure in Rocky Mountain spotted fever. Alteration in the blood volume and the thiocyanate space at various stages of the disease. *Arch. Intern. Med.*, 83:331–347.

HAAS J.; ROTH, S.; ARNOLD, K.; SCHMIDT, T. 2013. The protein model portal – a comprehensive resource for protein structure and model information. Database bat031.

HASEGAWA M & KISHINO H. 1994. Acuracies of the simple methods for estimating the bootstrap probability of a maximum-likelihood tree. *Mol. Bio. and Evolution*, 11(1):42-145.

HEEKIN, A. M.; GUERRERO, F. D.; BENDELE, K. G.; SALDIVAR, L.; SCOLES, G. A.; DOWD, S. E.; GONDRO, C.; NENE, V.; DJIKENG, A.; BRAYTON, K. A. 2013a. Gut transcriptome of replete adult female cattle ticks, *Rhipicephalus (Boophilus) microplus*, feeding upon a *Babesia bovis*-infected bovine host. *Parasit. Res.*, 112(9):3075-90.

HEEKIN, A. M.; GUERRERO, F. D.; BENDELE, K. G.; SALDIVAR, L.; SCOLES, G. A.; DOWD, S. E.; GONDRO, C.; NENE, V.; DJIKENG, A.; BRAYTON, K. A. 2013b The ovarian transcriptome of the cattle tick, *Rhipicephalus (Boophilus) microplus*, feeding upon a bovine host infected with *Babesia bovis*. *Parasit. & vect.*, 6:276.

HEID, C.A.; STEVENS. J.; LIVAK, K.J., WILLIAMS, P.M. 1996. Real time quantitative PCR. *Genome Res.* 6:986-994.

HIGUCHI, R.; DOLLINGER, G.; WALSH, P.S.; GRIFFITH, R. 1992. Simultaneous amplification and selection of specific DNA sequences. *Bio Technology.*, 10(4):413-417.

HUELSENBECK, J.P.; RONQUIST, F. 2001. MRBAYES Bayesian inference of phylogeny. *Bioinformatics*, 17(54):755-9.

JIANZHU, M.A.; WANG, F.Z.; JIMBO, X. 2013. Protein threading using context-specific alignment potential. *Bioinformatics* (Proceedings of ISMB), 29(13):i257-i265.

JONES, D.T.; TAYLOR, W.R.; THORNTON, J.M. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*, 8:275–282.

JULIAN, G.S.; OLIVEIRA, R.W.; PERRY, J.C.; TUFIK, S.; CHAGAS, J.R. 2014. Validation of housekeeping genes in the brains of rats submitted to chronic intermittent hypoxia, a sleep apnea model. *Plos One*, 9(10):e109902.

KÄLLBERG, M.; WANG, H.; WANG, S.; PENG, J.; WANG, Z.; LU, H.; XU, J. XU. 2012. Template-based protein structure modeling using the RaptorX web server. *Nat. Protocols* 7:1511–1522.

- KARPLUS, K.; KARCHIN, R.; BARRETT, C.; TU, S.; CLINE, M.; DIEKHANS, M.; GRATE, L.; CASPER, J.; HUGHEY, R. 2001. What is the value added by human intervention in protein structure prediction. *Proteins* 45(S5):86–91.
- KEIRANS, J.E. 1992. *Systematic of the Ixodidae (Argasidae, Ixodidae, Nuttalliellidae): An overview and some problems*. In: *Tick vector Biology. Medical and Veterinary aspects*. Berlin, spring Verlag. 1-21p.
- KELLER, P.M.; WAXMAN, L.; ARNOLD, B.A.; SCHULTZ, L.D.; CONDRA, C., CONNOLLY, T.M. 1993. Cloning of the cDNA and expression of moubatin, an inhibitor of platelet aggregation. *J. Biol. Chem.*, 268(8):5450-5456.
- KELLEY, L.A.; MEZULIS, S.; YATES, C.M.; WASS, M.N.; STERNBERG, M.J.E. 2015. The Phyre2 web portal for protein modeling – prediction and analysis. *Nat. Protocols*, 10:845-858.
- KEMP, D.H.; BOURNE A. 1980. *Boophilus microplus* the effect of histamine on the attachment of cattle-tick larvae—studies *in vivo* and *in vitro*. *Parasitology*, 80:487–496.
- KONTOPIDIS, G.; HOLT, C.; SAWYER, L. 2002. The ligand-binding site of bovine beta-lactoglobulin: evidence for a function?. *J. Mol. Biol.*, 318(4):1043-1055.
- KOTSYFAKIS M.; SA-NUNES, A.; FRANCISCHETTI I.M.; MATHER, T.N.; ANDERSEN, J.F.; RIBEIRO, J.M. 2006. Antiinflammatory and immunosuppressive activity of sialostatin L a salivary cystatin from the tick *Ixodes scapularis*. *J. Biol. Chem.*, 281(36):26298-26307.
- KOVAR, L. 2004. Tick Saliva in Anti-Tick Immunity and Pathogen Transmission. *Folia Microbiol.* 49(3):327-336.
- KOZERA, B. & RAPACZ, M. 2013. Reference genes in real-time PCR. *J. Appl. Genet.*, 54:391–406.
- KROGH, A.; BROWN, M.; MIAN, I.S.; SJÖLANDER, K.; HAUSSLER, D. 1994. Hidden markov models in computational biology. Applications to protein modeling. *J. Mol. Biol.*, 235:1501-1531.
- KUBISTA, M.; ANDRADE, J.M.; BENGTSSON, M.; FOROOTAN, A.; JONAK, J.; LIND, K.; SINDELKA, R.; SJOBACK, R.; SJOGREEN, B.; STROMBOM, L.; STAHLBERG, A.; ZORIC, N. 2006. *The real-time polymerase chain reaction. Molecular Aspects of Medicine*, Elmsford, NY, 27:95-125.
- KUMAR, S.; NEI, M.; DUDLEY, J.; TAMURA, K. 2008. MEGA a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.*, 9(4):299-306.
- LABRUNA, M.B., PINTER, A., TEIXEIRA, R.H.F. 2004. Life-cycle of *Amblyomma dubitatum* (Acari: Ixodidae) using capybaras (*Hydrochaeris hydrochaeris*) as hosts. *Experiment. Appl. Acarol.*, 32:79-88.

- LAVOIEPIERRE, M.M.J.; RIEK, R.F. 1955. Observations on the feeding habits of argasid ticks and on the effect of their bites on laboratory animals, together with a note on the production of coxal fluid by several of the species studied. *Ann. Trop. Med. Parasit.*, 49:96–113.
- LEBOULLE, G.; ROCHEZ, C.; LOUAHED, J.; RUTI, B.; BROSSARD, M.; BOLLEN, A.; GODFROID, E. 2002. Isolation of *Ixodes ricinus* salivary gland mRNA encoding factors induced during blood feeding. *Am. J. Trop. Med. Hyg.*, 66(3):225-233.
- LEINONEN, R.; DIEZ, F.G.; BINNS, D.; FLEISCHMANN, W.; LOPEZ, R.; APWEILER, R. 2004. UniProt archive. *Bioinformatics*, 20:3236–3237.
- LIVAK, K.J.; SCHMITTGEN, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25(4):402–408.
- LORD, J.C.; HARTZER, K.; TOUTGES, M.; OPPERT, B. 2010. Evaluation of quantitative PCR reference genes for gene expression studies in *Tribolium castaneum* after fungal challenge. *J. Microb. Methods* 80:219-221.
- LOVELL, S.C.; DAVIS, I.W.; ARENDALL, W.B.; de BAKKER, P.I., WORD, J.M.; PRISANT, M.G.; RICHARDSON, J.S.; RICHARDSON, D.C. 2003. Structure validation by calpha geometry phi, psi, and cbeta deviation. *Proteins*, 15,50(3):437-450.
- MANS, B. J. 2005. Tick histamine-binding proteins and related lipocalins: potential as therapeutic agents. *Curr. Opin. Investig. Drugs*, 6(11):1131-1135.
- MANS, B. J.; GASPER, A. R.; LOUW, A.L.; NEITZ, A.W. 1998. Purification and characterization of apyrase from the tick, *Ornithodoros savignyi*. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.*, 120(3):617-624.
- MANS, B.J, & RIBEIRO J.M. 2008b. A novel clade of cysteinyl leukotriene scavengers in soft ticks. *Insect. Biochem. Mol. Biol.*, 38(9):862-870.
- MANS, B.J. & RIBEIRO J.M. 2008a. Function mechanism and evolution of the moubatin-clade of soft tick lipocalins. *Insect. Biochem. Mol. Biol.*, 38(9):841-852.
- MANS, B.J.; RIBEIRO, J.M.; ANDERSEN, J.A. 2008. Structure function and evolution of biogenic amine-binding proteins in soft ticks. *J. Biol. Chem.*, 283(27):18721-18733.
- MANSUETO, P.; VITALE, G.; CASCIO, A.; SEIDITA, A.; PEPE, I.; CORROCCIO, A.; di ROSA, S.; RINI, G.B.; CILLARI, E.; WALKER, D.H. 2012. New insight into immunity and immunopathology of Ricksttsial diseases. *Clin. Dev. Immunol.*, 12:1-26.
- MARTELL, M.; GÓMEZ, J.; ESTEBAN, J.I.; SAULEDA, S.; QUER, J.; CABOT, B.; ESTEBAN, R.; GUARDIA, J. 1999. High-throughput real-time reverse transcription PCR quantitation of hepatitis C virus RNA. *J. Clin. Microbiol.*, Feb., 37(2):327-332.
- MISHRA, J.; MORI, K.; QING, M.; KELLY, C.; BARASCH, J.; DEVARAJAN, P. 2004. Neutrophil gelatinase-associated lipocalin: a novel early urinary biomarker for cisplatin nephrotoxicity. *Am. J. Nephrol.*, 24(3):307-315.

MOODY, A.; SELLERS, S.; BUMSTEAD, N. 2000. Measuring infectious bursal diseases virus RNA in blood by multiplex real-time quantitative RT-PCR. *J. Virol. Methods*; Mar 85(1-2):55-64.

MOREIRA, H.N. 2012. Genoma funcional de glândulas salivares do carrapato *Amblyomma cajennense* em condições de infecção e não infecção com *Rickettsia amblyommii* [Dissertação de Mestrado] Viçosa Universidade Federal de Viçosa. 120p.

NARASIMHAN, S.; KOSKI, R.A.; BEAULIEU, B.; ANDERSEN, J.F.; RAMAMOORTHY, N.; KANTOR, F.; CAPPELLO, M.; FIKRIG, E. 2002. A novel family of anticoagulants from the saliva of *Ixodes scapularis*. *Insect. Mol. Biol.*, 11(6):641-650.

NAVA, S.; BEATI, L.; LABRUNA, M.B.; CÁCERES, A.G.; MANGOLD, A.J.; GUGLIEMONE, A.A. 2014. Reassessment of the taxonomic status of *Amblyomma cajennense* (Fabricius, 1787) with the description of three new species, *Amblyomma tonelliae* n. sp., *Amblyomma interandinum* n. sp. and *Amblyomma mixtum* (Koch, 1884), and *Amblyomma sculptum* (Berlese, 1888) (Ixodida: Ixodidae). *Ticks and Tick-borne Dis.*, 5:252-276.

NUGENT, T. & JONES, D.T. 2009. Transmembrane protein topology prediction using support vector machine. *BMC Bioinformatics*, 10:159-170.

NUNN, M.A.; SHARMA, A.; PAESEN, G.C.; ADAMSON, S.; LISSINA, O.; WILLIS, A.C.; NUTTALL, P.A. 2005. Complement inhibitor of C5 activation from the soft tick *Ornithodoros moubata*. *J. Immunol.*, 174(4):2084-2091.

PAESEN, G.C.; ADAMS, P.L.; HARLOS, K.; NUTTALL, P.A.; STUART, D.I. 1999. Tick histamine-binding proteins isolation cloning and three-dimensional structure. *Mol. Cell.*, 3(5):661-671.

PAESEN, G.C.; ADAMS, P.L.; NUTTALL, P.A.; STUART, D.L. 2000. Tick histamine-binding proteins: lipocalins with a second binding cavity. *Biochim. Biophys. Acta.*, 1482(1-2):92-101.

PEI D.S.; SUN, Y.H.; CHEN, S.P.; WANG, Y.P.; HU, W.; ZHU, Z.Y. 2007. Zebrafish GAPDH can be used as a reference gene for expression analysis in cross-subfamily cloned embryos. *Analytic. Biochem.* 363:291-293.

PENG, J.; XU, J. 2011a. A multiple-template approach to protein threading. *Proteins*, Jun, 79(6):1930-9.

PENG, J.; XU, J. 2011b. RaptorX: exploiting structure information for protein alignment by statistical inference. *Proteins*, 79(S10):161-171.

PFÄFFLE, M.; LITWIN, N.; MUDERS, S.V.; PETNEY, T.N. 2013. The ecology of tick-borne diseases. *Int. J. Parasitol.* 1016:1-9.

POUX, S.; MAGRANE, M.; ARIGHI, C.N.; BRIDGE, A.; O'DONOVAN, C.; LAIHO, K. 2014. Expert curation in UniProtKB: a case study on dealing with conflicting and erroneous data. Database. doi:10.1093/database/bau016.

- RAIBLE, D.G.; LENAHAN, T.; FAYVILEVICH, Y.; KOSINSKI, R.; SCHULMAN, E.S. 1994. Pharmacologic characterization of a novel histamine receptor on human eosinophils. *Am. J. Respir. Crit. Care Med.*, 149(6):1506-151.
- RANDALL, M.B. & WALKER, D.H. 1984. Rocky Mountain spotted fever: gastrointestinal and pancreatic lesions and rickettsial infection. *Arch. Pathol. Lab. Med.*, 108:963–967.
- RIBEIRO, J.M.; MAKOUL, G.T.; LEVINE, J.; ROBINSON, D.R.; SPIELMAN, A. 1985. Antihemostatic, antiinflammatory, and immunosuppressive properties of the saliva of a tick, *Ixodes dammini*. *J. Exp. Med.* 161(2):332-344.
- RIBEIRO, J.M.; MAKOUL, G.T.; ROBINSON, D.R. 1988. *Ixodes dammini*: evidence for salivary prostacyclin secretion. *J. Parasitol.* 74(6):1068-1069.
- RONQUIST, F.; HUELSENBECK, J.P. 2003. MRBAYES 3 Bayesian phylogenetic Inference under mixed models. *Bioinformatics*, 19:1572–1574.
- RYDKINA, E.; SILVERMAN, D.J.; SAHNI, S.K. 2005. Similarities and differences in host cell signaling following infection with different Rickettsia species. *Ann.N.Y. Acad. Sci.* 1063:203-206.
- SAMBROOK, J., RUSSELL, D.W. 2001. *Molecular cloning: a laboratory manual*. 3ed. New York, Cold Spring Harbor Laboratory.
- SANGAMNATDEJ, S.; PAESEN, G.C.; SLOVAK, M.; NUTTALL, P.A. 2002. A high affinity serotonin- and histamine-binding lipocalin from tick saliva. *Insect Mol Biol*, 1(11):79-86.
- SAWIER, L. & RICHARDSON, J.S. 1991. Using appropriate nomenclature. *Trends Biochem. Sci.*, 16:11.
- SAYLER, K.A.; WAMSLEY, H.L.; PATE, M.; BARBET, A.F.; ALLEMAN, R. 2014. Cultivation of *Rickettsia amblyommii* in tick cells, prevalence in Florida lone Star tick (*Amblyomma americanum*). *Parasit. & vect.* 7:270.
- SCHLEHUBER, S & SKERRA, A. 2005. Lipocalins in drug discovery from natural ligand-binding proteins to anticalins. *Drug Discov Today*, 10(1):23-33.
- SELVEY, S.; THOMPSON, E.W.; MATTHAEI, K.; LEA, R.A.; IRVING, M.G.; GRIFFITHS, L.R. 2001. Beta actin an unsuitable internal control for RT-PCR. *Mol. and Cell. Probes* 15:307-311.
- SHELNESS, G.S.; LEDFORD, A.S. 2005. Evolution and mechanism of apolipoprotein B-containing lipoprotein assembly. *Curr. Opin. Lipidol.* 16:325–332.
- SKERRA, A. 2000. Lipocalins as a scaffold. *Biochim. Biophys. Acta*, 1482(1-2):337-350.
- SUZEK, B.E.; HUANG, H.; MCGARVEY, P.; MAZUMDER, R.; WU, C.H. 2007. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*, 23:1282–1288.

- TEGONI, M.; RAMONI, R.; BIGNETTI, E.; SPINELLI, S.; CABBILLAU, C. 1996. Domain swapping creates a third putative combining site in bovine odorant binding protein dimer. *Nat. Struct. Biol.* 3:863–867.
- TENG, X.; ZHANG, Z.; YANG, L.; LI, F. 2012. Validation of reference genes for a quantitative expression analysis by real time RT-PCR in four lepidopteran insects. *J. Insect. Sci.* 12(60):1-17.
- THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic. Acids Res.*, 22(22):4673-4680.
- TRIVEDI, S.; ARASU, P. 2005. Evaluation of endogenous reference genes for real-time PCR quantification of gene expression in *Ancylostoma caninum*. *Mol. and Biochem. Parasitol.*, 143:241-4.
- TYBURSKI, J.; STUDZIŃSKA, A.; DACA, P.; TRETYN, A. 2008. PCR in real time: methods of data analysis. *Biotech.*, 1(80):86–96.
- VALBUENA, G.; FENG, H.M.; WALKER, D.H. 2002. Mechanisms of immunity against rickettsiae. New perspectives and opportunities offered by unusual intracellular parasites. *Microb. Infect.*, 4:625-633.
- VALENZUELA, J. G. 2004. Exploring tick saliva: from biochemistry to 'sialomes' and functional genomics. *Parasitol.*, 129(S):83-94.
- VITALE, S.; MANSUETO, S.; GAMBINO, G.; MOCCIARO, C.; SPINELLI, A.; RINI, G.B.; AFFRONTI, M.; CHIFARI, N.; La RUSSA, C.; Di ROSA, S.; COLLETTI, P.; BARBERA, C.; La ROSA, M.; Di BELLA, G.; FERLAZZO, V.; MILANO, S.; D'AGOSTINO, P.; CILLARI, E. 2001. The acute phase response in Sicilian patients with boutonneuse fever admitted to hospitals in Palermo, 1992-1997. *J. Infect.*, 42:33-39.
- WALKER, A.R.; FLETCHER, J.D.; GILL, H. 1985. Structural and histochemical changes in the salivary glands of *Rhipicephalus appendiculatus* during feeding. *Int. J. Parasitol.* 15(1):81–100.
- WALKER, D.H. & ISMAIL, N. 2008. Emerging and re-emerging rickettsioses: endothelial cell infection and early disease events. *Nat. Rev. Microb.*, 6:375-386.
- WALKER, D.H. & MATTERN, W.D. 1980. Renal failure in Rocky Mountain spotted fever. *Arch. Intern. Med.*, 140:867–867.
- WALKER, D.H.; VALBUENA, G.A.; OLANO, J.P. 2003. Pathogenic mechanisms of diseases caused by Rickettsia. *Ann. N. Y. Acad. Sci.*, 990:1-11.
- WANG, Y.; HAN, Z.; YAN, S.; MAO, A.; WANG, B, REN H, CHI Y. 2010. Evaluation of suitable reference gene for real-time PCR in human umbilical cord mesenchymal stem cells with long-term in vitro expansion. *In Vitro Cell. & Develop. Bio. Animal*, 46:595-599.

- WAXMAN L. & CONNOLLY TM. 1993. Isolation of an inhibitor selective for collagen-stimulated platelet aggregation from the soft tick *Ornithodoros moubata*. *J. Biol. Chem.* 268(8):5445-5449.
- WEISS, E.; STRAUSS, B. 1991. The life and career of Howard Taylor Ricketts. *Rev. of Infect. Dis.*, 13:1241-1242.
- WIKEL SK. 1982. Histamine content of tick attachment sites and the effects of H1 and H2 histamine antagonists on the expression of resistance. *Ann. of Trop. Med. and Parasitol.*, 76:179–185.
- WILLADSEN, P. 2004. Anti-tick vaccines. *Parasitol.*, 129:S367–S387.
- WONG, M.L.; MEDRANO, J.F. 2005. Real time PCR for mRNA quantitation *Biotech.* 39:75-85.
- ZHANG, Y.; RIBEIRO, J.M.; GUIMARAES, P.N. 1998. Nitrophorin-2: a novel mixed-type reversible specific inhibitor of the intrinsic factor-X activating complex. *Biochem.* 37(30):10681-10690.

ANEXOS

Tabela S1: Sequência de aminoácidos dos transcritos de lipocalinas identificadas no RNA-seq de glândulas salivares de *Amblyomma sculptum*.

Proteína	Cod. do transcrito RNA-seq	Sequência
As_LIP1	As_10488	MKLVLFFVLLGSAFSTFSTGTPGDATERKRPYWANERRFGNYQDAWKALNQSYSTKYWLVLSTFN DDGGSWGKNFTCLSVQESELNETDKSAVSVFTFRNLTSNTTWFTVKEKQVAVFTYNYTKKQNAFQY TLENGTQINDTLIFSDGETCDIFSVPSMNGGKGCELWVNEKYVDNVPKCCLFIYNYFCTTQGMKRHYI YKRKACQNAQIFPDQPQYTC
As_LIP2	As_30706	MDLQPWYLLYLIADVVMATCPGEAKKVCSTGTRNGGARCISPWQFMSSSNIYLIKSNSYKIGTICIKAR TLEKDPKQTLSHVVSFYVNASQTWLAMNATYNPLKTHHTTGRAKAFYSVYNGETNYTFSLGLTYYAI VRKEKGSSSEDLRNSCELWVNDKYFNRNCTAKKERCEKKFRTICGPNPEISYDCYNCNQPKRGDN
As_LIP3	As_30883	MDLWPWYLLYFITVVMPTCPGEAKKLTDRPKNNGSNEDSPWKFLSSTNIYLNKTNISLPMSCIKANT TERNQCEETLSHVVSFKNASSGLWGHLNATYAPFNKSEKVATAFISVDPTFNVTIYTFLDSDYKYCA VVHKWKHPITITGFGNICELWVNDYFQSMDSSELQSSCHKKQFQEHCRQCPTFYSYDIHECDC
As_LIP4	As_10489	MKLVLFFVLLGSAFSTFSTGTPGDATERKRPDANERRFGNYQDAWKALNQSYSTKYWLVLSTFN NDDGGSWGKNFTCLSVQETELNETDKSAVSVFTFRNLTSNTTWFTVKEKQVAVFTYNYTKKQNAFQY TLENGTQINDTLIFSDGETCDIFSVPSMNGGKGCELWVNEKYVDNVPKCCLFIYNYFCTTQGMKRHYI YIKRACRNTGQMIFDMPYTC
As_LIP5	As_32787	MNGFSIALLLALGITTGAQLLKKDLYAINTTGRVWTVQRTFVRSSEKGEHTCVYAWQVSLEGDKY QFEQHYMEGGKRIMEPLHGRLSDGDSGPLVTSKEEGGVGIPYTLVYWDNDRKCGILKFPDQKTKKI ECELHVWEDYLLKSPSTPCDHDYDSICGSPKHAVYHSGCLTDAK
As_LIP6	As_40720	MTTAVFTAFAVFTALLGLPCVQSQGVSRMAMYEYDQFQDIKAFNSSKFYWFYASNYDTLNIEDRR CINRISELTNDGMVFETGYMQGNTSKETSYYGNFYVSWPVKNNKNVQRNINSLNHLVVEPGQAEV PPTNFSLVSDYKDCIILLVLNEDIKSHFAPEYMETYATQCIVLLSDSEARRGINNTYHHKNICQNQYY NACLKYKGQENSHRQYYDGSQSPENDVTPTS
As_LIP7	As_61324	MKLALFFVLLGSAFSTFSAEELGNATGEIVPEWADEMRFNGNYQDAWKALNQNESKYWLVLSTFN DDGGSWGKNFTCLSVQETKLNDDTKSAVSVFTFRNLTSNTTWYTVVEERVQAIIFTYNYTKKENAIQYT LENGTQINDTLIFSDGETCDIFSVPYMNSGKGCELWVNDKYVDNIPECCLFMYNYFCNPKGKIRHHIF KPKMCRNNTKRVNENQISC
As_LIP8	As_21225	MTRLAVSCAFAAIVTLQLCTEVMSSGSRDDKPYEERVANFRHQNVREVADIRGKVVYKQNYRKGSE SRCLYFKRINRYSKTKYRFELGSTPPYSKMYLVKFSRDLDKTGKHDMMYNAVYKFLPNDKYKVRKQ MYADKKRTCMFVDDRDSEHERARCQLMRPARYANERLPEQCKRVYNQNLGKSIKLYRPPWCRSL PERWYTVYKRRKRSRSD
As_LIP9	As_23634	MNLFALALFFALAVGTSTAQTMEQLREALDTERVWTVQRTFERIDDKGKHTCVYALRGTLEGNEY QFEQWYKHGERWVRHPLHGRLSQKGNNAVLTVSQEKGANGLAYKLLFWDKTRSCGILQFTDKDDD RIKCELHVWEPALRNGLSTYPCETEYEQFCGDQKYPVYTPACMYN
As_LIP10	As_6242	MQRSELPLFLLSLCLLNSGDEEPREATSNFIMELDIRKAFNTSEVIWLYKQNYNNTFNNGTEDGVL KYHTYCLSDKRISLTEGAYNFTQNVRIDGAEQVFEYGEFVRNATRTENPPNSLLIYDTGTPDHRHF ELMTLYNNGSYKCMVFFVYSLYQMLTDAEVLCEMYIPENAIKNSPPMDCLKFFNRKCAIPPQKEYQI YFENCTAGAGTEVKPSHENNSVE
As_LIP11	As_7608	MRTLFTGVLFITADFAHSISRYEYEPNLENEADYVQQDDIYKAFNISPFDYWLFGFSYHSNHTENKSC VYFDIENISESGMNVSSNFIKHKRGKINYTGEFYSTNFKNMSDQAREKYNLSYARLRENGTWPM NYTLIYSDYQNCISIFRVLQIHEGYGCMVLLNATAYTGMPEACQSVFNDTCAKCNHPTFEQVFNTSC HKTPPELLTLGNEQFPSC
As_LIP12	As_53994	MAHSIITMPSLISKSAVFLVMAAFIGIGAHAGRRIDDEPWYEQYQDVRRALNISTDYWLHTVNFPRAY TEGRRCTIFHIKQWHEDRMNYSHYLVNDTLQGMVEYVGTFFQTPLNGPVPERKTNNGFYASLTSEK WHPQNYTLIYSDYENCLILRVRDFYKGYACMVLVNEPPTNNSIPWKCQLVYRNACNGSGIAEQIYER SCNKSQLLGRWS
As_LIP13	As_55838	MRKSRSCCTMGRWRALWIAAHVAICCHVGYGNHITDNGENIDIKKAVNTNQTLWLLEQTYKNSFTLC SENFCEKENEICIRNKMVNISDTEYFNFQTVRVENDDATLYLGKFVETNPLKSMVEVTDTTGPADPELV TLQYQDPENGPVMSFIEPLDPELRKSFSGTCEMYFRGRPTSSDPSQSCKNFFTRCNCSTAVYKPYR DTCSDPTLADSTASNDIPGN
As_LIP14	As_6241	MKSTMKLALFFVLLGSAISAFSAGGLDARGGSPYPDWADEMRFNGNYQDAWKALNQSETTKYWLVL LSNFNNDGGSWGENFTCLSVQETQLNETDKSVSVFTFRNASNATTEFKVVEEKVAVFNYNMTKE NAIEYTLANGTHITDLIFSDGETCDIFSVPIYINGGKGCELWVNDKYVDKIPKCLFVYNYFCTPKGKT RYIYKMKMCNIGKKNENSS
As_LIP15	As_7785	MRTLFTGVLFITADFAHSISRYEYEPNFENEANYVQQDDIYKAFNISPFDYWLFGFSYHSNHTENKSC VYFDIENISESGMNVSSNFIKYQQRGKINYTGEFYSTNFKNMSDQAREKYNLSYARLRENGTWPM NYTLIYSDYTNCSIFRVLQIHEGYGCMVLLTNAATAYTGMPEACQSVFNETCAKCNHPTFEQVFNNSC HKTPESS
As_LIP16	As_5909	MAHSIITMPSLIVKAVALLAITLSFMEIETHGDIRIDEPTYFEGQDIYKAFNISDGFVLHTQNFERKMTG GRKCTYFHMGRIDQNGMNYTSYYWFPNGTGMTPYHGIFYKTPPVSTLERKKSNGLNVSKTSEPW HPRNFRVYADYKSWLILRVLDFFRPMNGYACMVLVSDPPSNISMPDCKNAFRIACNGSGKAEQIYE QSCNKPALQL

As_LIP17	As_6240	MKLTMKLALFFVLLGSAISAFSAGRPGDAPGQSPYPDWADEMRFGNYQDAWKALNQSNTTKYWL LSNFNNDGGTWFGENFTCLSVQETQLNETDKSAVSVFTFRNASNATTEFKVEEKVAVFNYNVTYKE NAFEYTLANGTNITDITLIFSDGETCDIFSVPIYINGGKGCCELWVNDKYVDKIPKCCFLVYNYFCTPKGT RYIYKMKMCRGPTNQKKNETS
As_LIP18	As_86866	MNSFVVMALLGATVISAKPSKEDLYKALNTDRIWTVRRSYERGAPNAKHSCVYAKKEDLSDDDNY QFVQGFKNQSKWQNETLYGKLSQDQAGEAVLTVSKIPNPGEGNGVPTLQYWEKENHCGVLTFTGK DRELKCELHVVEQKLSPATNWLRCDELAYETICTNMHKFPVYTDDECLREVRN
As_LIP19	As_31297	MRTLLLTAFLVITAHLGHSDEYYEYETDIYNETYYVQHQNIYAFNDIPDFYWLFGFNYSKHTENKSC VYFDVENITQSAMNWSNFIRYQGRGKINYTGEFYSTKFKNMSLDQGRQVHNSLYARLRETQWP MNYTLIYSDYQNCVFRVLQIDGGYGCMVLLTNATAYTGMPEECQRFVNDCTCLCNHTTFEQVFN SCLRTPGELTFKNEEFSSC
As_LIP20	As_6502	MEDENLFAHSTITMYWLILKAAVFLAIAVAFMGIDAYEDISSHDVKNEVDYRAFNFSSDYWLLAVNFP RVYTQGRRTFFHIKELREDSMNYSSHVFNNSWQRMDYVGTFFTTPLENGSVVEEPQKNNSFYAS RTSEKWHQNYTVVSSDYNLSLILRVQDFYNGYACMLVSEPPANTTSLPWECQLENTTACNETCIS EQVYEKNCCKPENVTR
As_LIP21	As_6244	MKLTMKLALFFVLLGSAISAFSAGRPGDAPGQSPYPDWADEMRFGNYQDAWKALNQSNTTKYWL LSNFNNDGGTWFGENFTCLSVQETQLNETDKSAVSVFTFRNASNATTEFKVEEKVAVFNYNVTMKE NAIEYTLANGTNITDITLIFSDGETCDIFSVPIYINGGKGCCELWVNDKYVDKIPKCCFLVYNYFCTPRGK RYIYKMKMCKNKK
As_LIP22	As_113729	MPCLIAKTAFLAITLPPFMEIGAQEERRIDEPEGYFNHQDVYKAFNFSGGFWLHTQNFQRMKTSGRIC TYFQIHGIDENGMNYSHYLVNGTKQNMTYYGQFRTPAVGNETRNNSNAFNASKTSEKWHPSDY RVIYSDYKSCVILRVDFPRMNGYACMLVSDPPANTSMESECEERYTLACNITCITEQIYDDSCNKS EVV
As_LIP23	As_55591	MPWLILKAAVFLGLAVPFMGIDAHVDTLIGEETIDQKNQDVSRAFNFSSDYWLYAVNFPRIYTKGRN CTLFHIKELREDRMYSSHYLENNRQSMYVGNFYKTPLEGGSGDRERVKYNSFRATLTSEQWHP RNYTLIYSDYENCLILRVLDYFYGYACMLVSDRAANTTSLPWYQKLNNTACNETCISEQIYEDSCK KSALLEGSS
As_LIP24	As_61325	MKLALFFVLLGWAFFSTFAEELGNAIGENVYPDWADEMRFGNYQDAWKALNQNESTKYWLILSTFN DDGGSWGKNFTCLSVQETNLNDTKSAVSVFTFRNLTSNTTWYRVEERVAIFTYNYTKKENALQY TLENGTQLNDTLIFSDGETCDIFSVPIYINGGKGCCELWVNDKYVENIPECCFLMNYNYFCNPKGIKRHHI FKPKVCRNTRKVN
As_LIP25	As_30884	MDLWPWYLLYLITVVMPTCPGEAGEGYQVTTNGEEGCVSPWQFMSSSNIYLIKSNMSYGNITCIKA RTSAKDQNTQTLSHVVVYMLNSQKWFAMNATYNPLTKDANGTATAFTSVDGNTTYTFFALLDITY AIVRKEKGSSEDLRESCCELWVNDYFRPDCTAKKLRCEENFCMKCEPKREMSYDCYNCTKPKQ
As_LIP26	As_30885	MDLWPWYLLYLITVVMPTCPGEAGEGYQVTTNGEEGCVSPWQFMSSSNIYLIKSNMSYGNITCIKA TVESSEREETLSHVSVKFTSGSWYEINATYAPVKNKTRRKPATAFISVDPGSNVTIYTLHTNKYCA VVNKWKRAEEGLGDICELWVNDYLRPGRKVPSCSKKFKQHKAGFSYDKHECDC
As_LIP27	As_6243	MKSTMKLALFFVLLGSAISAFSAGTHGDAPGGAPYPDWADEMRFGNYQDAWKALNQSNTTKYWL LSNFNNDGGTWFGENFTCLSVQETQLNETDKSAVSVFTFRNASNATTEFKVEEKVAVFNYNVTYKE NAFEYTLANGTNITDITLIFSDGETCDIFSVPIYINGGKGCCELWVNDKYVDKIPKCCFLVYNYFCTPKGT RYIYKMKMCKNKK
As_LIP28	As_89672	MAVFAALFFCACIFKVAQETDAAAPDSETLDGYHLLGLRKTFRVLTQTTFFHQDNTHRCITVTVTG KDDDNHIVTLEVDYNTVNDQAWKKYSQRFKFINEGSQYNKMRIYDGGSPPGTYNFLRRDLGCTVIE VEDSQHPGQDIVGLAREEHIVPTIKERGNMMLWVENGKKEYPDQPCYDEYNRLC
As_LIP29	As_76441	MKSTMKLALFFVLLGSAISAFSAGRPLGNAPGGSPYPDWADEMRFGNYQDAWKALNQNESTKYWL LSNFNNDGGTWFGENFTCLSVQETQLNETDKSAVSVFTFRNASNATTEFKVEEKVAVFNYNVTMKE NAIEYTLANGTHITDITLIFSDGETCDIFSVPIYINGGKGCCELWVNDKYVDKIPKCCFLVYNYFCTPKGT RYIYKMKMCKNKNENS
As_LIP30	As_30880	MDLWPWYLLYLITVVMPTCPGEAGEGYQVTTNGEEGCVSPWQFMSSSNIYLIKSNMSYGNITCIKA TERNQCEETLSHVSVFKNASSGLWGHLNATYAPFNKSEKVAATAFISVDPTFNVTIYTLFSDYKYCA VVHKWKHPTITGFGNICELWVNDYFQSMDSSELQSSCNKFKQEHCRCTPFSYDIHECDC
As_LIP31	As_75895	MNGLGAFLVIGVTACAADSTTKDDLRRALNTEQKIWVVKRSYTRSTGEKEHKCVYATKDSLEGDN YEFHQGYKVGEEWKKELKYLIGLSEGDGFATLKVSKRGEQGITYLKAWNHKEKCGILTFQKQSNF ECELHAWDEVLETEGADPCAGEYQQICGGKAAHQVFSHECLTGS
As_LIP32	As_89125	MEMLWAVAFATISLGGIANGKLGPPGGPLKLRDVSDFSKAMSQFEYLVAISDSNNDTVFECLAAR TNIDQEAKTATYTFLLPTAGMELSFHIKPGDEPATGILTSNVDPSPREGIIYTDYQNCVVANLEIHGD GNQCTLWARLEVDKNVPQVCIDHFVDVCGVTPAHSRDLCPDGEDY
As_LIP33	As_30881	MDLWPWYLLYLITVVMPTCPGEAGEGYQVTTNGEEGCVSPWQFMSSSNIYLIKSNMSYGNITCIKA ERNQSEETLSHVSVFKNASSGLWGHLNATYTPFNKTDKVATAFISVDPKYNTIYTLFNSDYKYCAV VNKWKHPTITGFGNICELWVNDYFQSMDSSELQSSCNKFKQEHCRCTPFSYDIHECDC
As_LIP34	As_65847	MNTRAVSVAVATLCVTIADAETRRGPHKLQRDVADISEIVAFFGDSVAVMDIDDDVLDCLTTIKTQY DPEGWSATYLSVQSGTGERFQPLHLWPKDKPDYFNFSSTRPGEVLVGHSDYFDHESCVAIEM PAVGDQCTLWVTRERKDSIPDHCLQNFEDICGTGVDLHTKTICKDVKSAFSKTPETSARAILMTTAW
As_LIP35	As_50908	MPHSIITMPSLSVKAAVLLAITLFFREIETHGDIRVDEEPIYFKDQDIYQAFNISKGFWLHTQNFERQKT GRRCTYFQIQSIDENGMNYSYLLFADNTKGMKPYHGIFYQTPPVSTSERNKRNLNSETSEPW HPRNFRLTYSDYKSCILRVLDFERMNGYACMLVSDSQAHISMPSTCQHAYRNACNGSGIAEQIYE KSCKKPLQV

As_LIP36	As_88490	MAQKESILAIIVFLVATIESRAEEGSSSSQPYRNIWKFLNTTDAIYSYRISREGVECKVDYSISENATD VSFQRYIPGSASALLPVKGRQSTRPVVTVGTFTFYAQLPNEAPDGMCIKYEKDGASYTEAIDRQSS NDLCALFSVAGRTDGIRLPPAVQ
As_LIP37	As_83441	MAGKITSALLITLVMMVNSRAETNSAPKQANYANLLKFLNTEAVLIFSLKTRSDLQCGVDNYRDT TSTNTFISRSYLAYSRGQENHRANVRSSGRSGMQWHPERLEGQFTSYKGFDMNTFKGEQAYTTE VLDSQQEKCAIFTRATSGSDRQDRYEVVRWSAIKDDGQLECCSSFVEKFKITLKNGLMIPAQPLIS KCKQMCLNGPCRPEGAEVH
As_LIP38	As_80624	MAILQTLVCLFALTFAIVEPSKATETNDDIYKFLNANPKIWWYNTSEANERNIVCRHDDRVDITQEGIS FYRYNNTRELLKGFVNWVYPESKTYDAMYVNDTGDKPVSDVEVIEVSRDNLCAVVSVTLMRTDQ PVLREIRVREEALQTGPDEKCKKFKDEILKEANKTAKSQSPACLTPHSL
As_LIP39	As_117390	MAISKTFILICLTVLVTLAEPREATQERSNTDSEIVEFLKAHDKTWVYNTTEENKENITCRFDFKHNITE KDVFFKRYNNSSEELLKGEFLNWNPDGNNTYDAMYVSDMQGEPVTVYEVFVSSDKSCAVVSVTL MRTVPTYVLRVREIRVSEDALKKGPDKVCKREKFEKILKVAKKKAKSHYSDDCKSKLNVY
As_LIP40	As_17860	MYALFVLAMFFTTGAPAMLDRLNNTSIEQFYNTSEPIWTTLTRADANFLNSCIVDVTFNNTLSNLYFERS YYWNQTRVKSPLRRLNKRRTDRSSRTLVDARLPILKPKQPLFEEELLYQSDSTCGVFRYFKHPD VTRYDLRVRNCSLATKPHENCTMFFNGTSLKLRPSIREKLQVREYVNPQCQAILWKTGDCQ
As_LIP41	As_115808	MALSQCFLLSFVVALFALLIANAEEKITKYEDKDIYKFLNNTDKIWWYNTTEKRRGRRTICKYDLKCNISS NGTFFHRVSKERPEELLRGEFLNWSATAPTEYNAMLVFNKSKRAVGLLEVYASKAYCAVFTVNLF RRGSPVYRELVRWDRALKKGPGSGCTNKFDQLQRNLNQSASKSHYSRTCQSLPSSMRC
As_LIP42	As_96613	MAALAIYVRLVSGLLLVTAFAQAVKAEESTEDFKQMLVKGSVFWVNTSEPGEIACRKDEISNVQEK NVSFARYSRKETTSENLYGILFNWDDDDSRGYDGIITNSKDELLSEEYLYLNPSTKCAVVKVMKF GDSKDANKNTWRDLRVSNAAGVVSAGPQDCAEQFENALKVTRKPTRQPYTDQCK
As_LIP43	As_54516	MSELWSCVLLLFAGLVVRVNTLFSNDELFGYGHKDAKTFVNTTDDVIWTVGTSDDPSDQTPVCQVDK WQCNQPSHCQLTRSSYIDRRKVVQGFVTFRASFTHRLNGWLQDILEIRNETGKLLRRETVMFLSR DSRCGVVYVYVPGVQGGWVELRARNSAIKEGLDFNCTRYFLELIGDQARRYDKKITVSLCIYLVVKIFY QISDCL
As_LIP44	As_40237	MRAMTQKQSIHAIIFALLATFILSNADDVSNPQIEPDIVKFSNESVIWVSYSLGFDQRNPCKVDYDVV TIPPHTSFRFRFTSDQPTRPYVQVQPNRRFVRNTYSVEGNFSYASDKSKNLDAMNLQYRETRARKG ERSGGGERSGQSDKSASTHREAIIFEAQGGKCAVFSVSRSGVLTAMKGSVDLRIKGSATDTPAVN KCFEEFKESSVYTDVTRRIQLVQPIASIAQCLERCRSNACTQGRA
As_LIP45	As_33701	MEKRNTPSIFIVLLVAMRSTAAADAESAPANNADFNKFLASEDPVWTIESTARNPPTCRLDVIDRQ NGRDALLHRYEKNMFRGSRSGQGVQSYDFIATFPTTDSMSLRQANDAPYSHEKLELISRDKGC ALISWKALRNAGSPLQKVEVTEEVARNASIDSEGNASCLNDLKQIMEEQHKGSQRRLLVYSPSCQK IHERRLSRPE
As_LIP46	As_78707	MALKQSIQTIAFVLLATVIHSNAEELRSSQGSEPDIEKFYSNGSVIWSYSISTDGHLCCKVDFAVVSNQ PDISFKRYSSDLHTRIPPARFRGQTFRIRGRFTNTRAHPPSGKFDMDIEYEDRQSDSTCGVFRYF YLEAIVFESENGDCAVFAVRPSPNTEAVRAHGHRRGPARAQVKQTLDLRIKSSSEDAEKTKRCFQEF KSEKYTKLHNGRVAQNIIMSLGKCKTGCLSNQECAQRLLEEQ
As_LIP47	As_28374	MEGVNYFTIAFILAAGISITGSTPTTGHLYNALNTGDRIWVYKRSYELHDENNDLHTCVAARQTYLTS SYQFIQNYLAGKTRHTEELFATIAQGRPLPILDVSKKPEHGHQVYVYTLHEHWNEEQHCAVLTVELFGKT QCELHVWEQDIKRYTPSTCDTVYASICYEQHTVFSDDCLNPAY
As_LIP48	As_17866	MFAILLTMMFFTTASIFTWYQLRHGFNTSIEQFYNTSEPIWTTLTRADASFNLSCIVDVTFNNTLSNLY FERSYYWNQTRVKSQLRRALNKRGTDRSSRTLVDARLPILKPKQPLFEEELLYQSDSTCGVFRYF KHPDVTRYDLRVRNCSLATVPHENCTMFFNGTSLKLRPSIREKLQERKVVYEPKQAILWSTGDCQ
As_LIP49	As_10888	MSCFTLALFLALGVTAFGSTNNPTRKDLLEEFATNEKIWLFRRTYEEQRAGKNYRCVYVLKTSSEGS KYEFEQYFRHAGIWESRHFYTEIRDQGGGSPVLIVRNSTESNGTPFTLRFDWSSSHCAILAWRNAKG EDEYEMHLRNNDLQSSYLEKCHQEFKLADGRDKTIYANNCKPKQAS
As_LIP50	As_56564	MPYRIFLVVVSALLYATESSHGGRKGGKTDIKEFVNTNDIIVTYNTTAHKRLVCLMDFKKEVTENDT QFERHHFTRQKRWAAKLLIGKFGIWSPEDRNGKPYDYMEVITFSQRKVVGAIELEYQHDDNTCGV FTVMMSGNAIPPTFSHELVRKNSSIQTGPKKECKKQFRKVVGEGKKIKGAYDHRCHKALKYNYENFLH LP
As_LIP51	As_7195	MAKLATFGRLLFALLVATAVTRIEAEVDNTDYDFREFLQQSKVIWVRDTPSEPVTTCRRDTISKINERS VSFERQFKNNSQTFQVSLKGLFNWQHWEDETRTPYDSMAVSNNEVPQTDEILEVLDTRNNCAV VKVIVFGLTAEGLTTVWRELRSQDALTVSPSEKCLNYFNVAVVGESRPSYSPSCQAH
As_LIP52	As_1555	MNSPMIFLITLGVTFIGADAKPRDLYNALNTDYPLWFMRSYPRDNECVYSKKIDLNDNYEFYQI YKAGGKRETLHLIASIDPRDEPVLDSVQTKGQPGIEHTLRYWNTNTHCGILTLEHEGKQHCEMYVW NENVRTAPSYDCEQEFELKCGQLGKYRVYKDYCTE
As_LIP53	As_2043	MAQLCILQALVFTLFVVAALSDKNGNQSDDFTEFLKDKTIWVFNTEPEGEVTCRRDTILNIEDSSVSF TRQFKNTEISVENLEGLFNWEHESKQRPYDSMTISKSENGIEEFLFVDTTGNCAVVKVMNI VDFNTDPSRIWRELRLTNPREGIDPTSECWKTFDNAVNTGKNWRESYNKDCQ
As_LIP54	As_15657	MNCFPLAVFVALITFAAHRKWTDLRDLANTRQPIWVMVRSYREVEKDGREHKCVFSEMKSIVGS DYEFYQQYQVQSETRRYLLYGALTDGQHAGDEPVLDSVSEYHQTRQYTLRHWDPTHECGILTQ RDGETECEMYVWQENLQSTHSYACENAYRGLCSKKYPVYTTDCQR
As_LIP55	As_55154	MALNLKQLVVALVARSSAQDDSSNSEQRRRTIDISEFLSTVNETPVVFGVVISPSQPCKVDI VSHTNKTGTTDFRYTPGASRVQRHARMSMIKAHIQQGFIRTGGQGETKLDTMELSSSESTPYGEEKI LYQSENDTCAIFRARLLYGRDKGSYVDFRVKATAITDESARKCWDNIEEVVRSMMNLPNTDLKFPAAEI AGCQAQCQNQHVTCKALSTANP
As_LIP56	As_44527	MAYAAVSIFLALYVAGSGATPTMGDLKQALRTNDAIWIVKRTFNPTSGSDQHQCIVYEIGTLMEDRY

DFTQHYYWVAGQWHSNRLYAKLKDRQGTGQPGFLVNNKKEGQETGRLYLQYWDGNNRRCGILKVAL
GGEDHYELHVWNERVAVVPRSCHEEYKLTTPGRKEFKLYSHDCANKRVAE

As_LIP57 As_97593 MKVVMENQQVWIMILLNLLIFTVKAENNTTEEDIRKFFDQSERIWTSHSTAGGEISCKVDEVYSKT
ETSIEFNRYMPLPYSNWGHQKLFGRFRKWKDTDTTKGFYDTMDLSISGRKPAGLERIHYNPNKNC
AVVWILKHSGSNLKAAYDIRVKESFLKGVRQSDCWKELKRVADRRPIRRVYKKGCEVDI

As_LIP58 As_10703 MSCFTLALFLALGVTAFGSTNKLTRKELEEFATNEKIWLFRRTYEEETAGQKYRCIYALKRSASGSE
YMFQYFRHGETWESRYFYTEIRDEPSSPVLIVKNRTESQGTFFSLLFWDSKSHCAIFTWRNAKGE
NELEMHVRNSDLQSPDLGDCRREFVNLAKGRDKTIYEYDCKR

As_LIP59 As_6776 MAKLSTFGGLLFPLLVATAVTRIEAEADNTDYDFREFLQQGKVIWVRDTPSEPVTTCRRDTISTIDKRS
VSFERQFKNNSQTFQESLKGKLFNWQHWEDETRTPYDSMAVSNNEVPETDEILELDTRNKCAVV
KVIVFGLSAEGLTTVWRELVRSEDALTMSPSAKCLNYFNEVVGKESRSPSYSPSCQ

As_LIP60 As_17867 MYAIFLLTMFFTTEASDFTLQQLRHGFNTSIEQFYNTSEPIWTTFFRADASFNLSCIVDVTFTNTLSNLY
FERSYYWNQTRVKIPLRRALNKRGTDRSSRTLVDARRPIYEPQQPLFEEELLYQSDSSCGVFRYF
KHPDEPGAPAAGYWTFSFSSSVRTSSFPQDPCVARDISPFAVEALQLLGMTSG

As_LIP61 As_50138 MLCLRGRNMRITFLVFLGVSHCVNALNLDLLMALNTTQPIWLKSTCDHFPHEPKSICVYSLKEYL
NETDYSFDQTYYYGPTRNSSHFYGKFFDLGESTDPMVTVSLEQGAEGINYLRYINTTEHCGVLTYE
DGDREVEGLHAWNDSVRNELHGCENAFKFCNGTAYDLFFDNCTASSENIMEHL

As_LIP62 As_84148 MGEKRSIAALLIVLATAVQSGADDASGLADNRDIVKFLNTSGAIIYSITTTQSTVLCKVDFAIKEEGT
EVTFKRYTPDYSSRAKSLAVGAPKAAKFSGLGTTTVKGTFCYQVDCGGKFDAMDIARSGVEYKEA
MEFQDTEGECGIFAVTAPMHLKAAVTLDLRVRSSGENKVKGEECLRQFKESDRSKKYENIAESLPEE
SVCKEKCCEMHQCMGPHSL

As_LIP63 As_45076 MRLTGSLILIGSVLSFLSAAENEVDTVKAPQAALNIVKVFNTSRNLWLYRRTESDEFEVVDGNWKIGFLE
RCIFFKINISMSNYFWHRTEWNGSRLSSLWMGTFVENGEKQPLSMLVSDLEEHPTPFMMNMLR
YTERGNHSCSVFVRRFDCEIEEENCEMYIRNNGSIEPTKGCTESFNSICRNRTSYLIYEESCQRTG
KNQIKGVLKRRKQ

As_LIP64 As_17859 MHAIFVLAMFITTGAPAMLDRLKTSIEQFYNTSEPIWTTFFRADVSFNLSCIVDVTFTNTLSNLYFERSY
YWNQTRVKSPLRRALNKRGTDRSSRTLVDARRPIYEPQQPLFEEELLYQSDSSCGVFRYFKHPG
VTRYDLRVRNCSLAMKPHENCTMFFKGTSKLRPSILEKLQEREVYNTCQAILWKTGDCQ

As_LIP65 As_56562 MPHRIFFLAVVSALLYATESHGRKGGKTDIKEFVNTNDIWTYNTTAHKRLPCLMDFKKEVTEKDIQ
FERHHFTRQKRWAAKHLIGKFIWSPEDKNKSKPYDYMEVLTFSQRKWVDAEILEYQHEDNTCGVF
TVMMSGNIPPTFSHELVRKNSSIQTGPKKECKQFLKVVGEGKKIKGAYDHRCHKALKYYNENFLHL
P

As_LIP66 As_17864 MYAIFLLTMFFTTEASDFTFQQLRRGFNTSIEQFYNTSEPIWTTLTRADASFNLSCIVDVTFTNTLSNLY
FERSYYWNQTRVKIPLRRALNKRGTDRSSRTLVDARRPIYEPQQPLFEEELLYQSDSSCGVFRYF
KHPDVTRYDLRVRNCSLATEPHENCTMFFNGTSKLRPSIREKLQERKVVYEPKCQAILWSTGDCQ

As_LIP67 As_114525 MERSRFPGILLALIAHVPLCTAENAERVTDIRKFLNTSEAIWTRTKTFDATDLGMRCKMDIRNSSTLY
TINFTRFLETRTRWKEVQLVGEFGIWGANASKPDDLHDEMEVKQVIQYRRTTGLRWISTEVLEHS
TEDMTCGVFTVISFLGVTQKFSLELRVKNSSVHAIPCEECSEVFKRKIQHRQRPRKVVYWPQCQDILKDP

As_LIP68 As_17873 MHAIFVLAMLFTTGAPAMLDRLNLSIEQFYNTSEPIWYTRADSRNSISCLVDVAVNTTLYLFDRS
SYLNKTRLLKALKMYLYNGRTGNTSRTLVEAEPIYEQKNPLFEEQLLYQSGDSSCGVFRYFKHPN
VNRVYDLRVRNCSLATGPDTCWLYFNKTAADRISKYERLKERKVVYDSCQAILWTTDGCQ

As_LIP69 As_115809 MALSQCFLLSLVVALFALLIANAEMNKLTKNNDIYKFLNTTDKIWWYNTTEKRRRQTICKYDLKCNISSN
GTFHHRVSKERPEELLRGEFLNWSATAPTEYNAMLVFNKRAVGLLEVLEYASKAYYCAVFTVNLFR
RGSPPVYRELVRVDRVLKIKVLAQDAPTSSTNCSEV

As_LIP70 As_113877 MAKSQCFVFLVLTFFFPNPTEQEQRHDDIYKFLSVSRKLWVYNTTEKQDAGHTICRYDVKVINI
TEEIIFFHRFSNRSSWTTELMRGDFFNANATTEKPYNAFFLYDANGTAVGIEALEYASRAYYCAIFTVS
LFQSGTPAHRELVRSENALMNGPGEACRKYDQLLRSVKQTAKAQYSSSCLGQLQPGRC

As_LIP71 As_8440 MAKLSTFRGLLLALLVATAVTRIEAEVDNIDYDFRDFLQEGKTIWVRNTSEPVTITCRKDTLKFVNDNRN
ASFERHFKNNSKSFQSGMKGEFFNWHHWPETHTPYDSMAVSNNTENVHKTDEILELLEDAKNDCAV
VKVILGDLNSNEGFTVWRELVRSEALTKSPSKECLDYFKLAAGGNPRPTYSPDCHQMP

As_LIP72 As_108939 MANYGIVIFALSTILVATVFTSQEPDFDIKFFNATEPIWTTNSTAGDWFPCKVDVKHSINETSATFYR
AYIDRWKRWKRNKTLVGQFGFWTNSSEGAYDYMNIERTIEGPGAGEEVIHQSDASQCAVFFVMGHV
GVTISPSYDLRVKQSFGLGRVDRDTCWKEFLNVTNGMKRKRLYHERCQKAFDINGDRILYS

As_LIP73 As_54515 MSELWSCVFLFAGLVVRVSTLFSNDELFGYGHKDAKTFVSTTDVIWTVGTSDPSDQTPVCQVDK
WQCNGPSSHQQLTRSSYIDRRKVVQCFKVTFRASFTHRLNGWPDILEIRNETDKLRRRETVMFLSR
DSRCGVVYVPGVQGGWVLRARNSAIQEKLDNFCTTYFLELIGDQARR

As_LIP74 As_53938 MSCFTLALFVALGVTAFGSTNKPTRKELEEFIRTNEKIWLFRRTYEEETDRQNHRCVYALKRNEGS
QYLFEQYFRHGGIWESEYFYTEIRDEQSPVLIVKNRTGSGQGTPTLRFWDSNSHCAVLTFFPNEKGEN
EFEMHIRNNDLESADFSDCTKEFETLAEGRDKTIYEYDCKR

As_LIP75 As_17862 MYAIFLLTMFFTTEASDFTLQQLRHGFNTSIEQFYNTSEPIWTTFFRADASFNLSCIVDVTFTNTLSNLY
FERSYYWNQTRVKSPLRRALNKRGTDRSSRTLVDARRPIYEPQQPLFEEELLYQSDSSCGVFRYF
FKHPDVTRYDLRVRNCSLATEPHENCTMFFNGTSKLRPSIREKLQERKVVYEPKCQAILWSTGDCQ

As_LIP76 As_17863 MYAIFLLTMFFTTEASDFTLQQLRHGFNTSIEQFYNTSEPIWTTLTRADASFNLSCIVDVTFTNTLSNLY
FERSYYWNQTRVKSPLRRALNKRGTDRSSRTLVDARRPIYEPQQPLFEEELLYQSDSSCGVFRYF
FKHPDVTRYDLRVRNCSLATEPHENCTMFFQRIQTQTKHTRKIARAKSVRT

As_LIP77 As_2042 MAQLWILQALVFTLVVAALSDRNNGQSDSDFTEFLKKGKTIWVFNTEPEVTCRRDTILNIEDSSVS
FTRQFKNNTSISVENLEGKLFNWEHWESKERKPYDSMTISKSKNGIEEFLFVDTTGDCAVVKVMM
IVDNTPSRIWRELVRNPLEGIDPTSECWKTFDNAVNITGKNWRESYNKDCQ

As_LIP78	As_108578	MVEARICSAALLFPLLLTAVLAFSNEEEQEPTADIEKFLRENKVVVYHTTEPSKVTCLRDEYINVNISH ARFSRYFVDNARKKEEKNLIGKFLDWEGGVPLQGGKYDAMEINDIEKPMYPLSYDVYEFQGGKDGKC AVVAVIDNSAGTTETWLDLRLVSDSVIEEGAQDCKAQFERILRIYPERKRVSRKSYESQCKNWF
As_LIP79	As_98549	MAMSCSQALWAVLIAMCSNIGNGNLQDNGNNIDIRKAVNTTDLKLLWQTYRNGFDVCTKDCIKV NETCIYNNMINITKEEYFNQITWTTERDDTTNYKGTFKDGIPEPKSMEVEYITDDPSDGSVSDDGQH QTWTLSYLDPDDYRCMVFAIGELDNEVQEVLTGCDMYIRGTPRKSDDPPSGCEEFFRQRNCVNTAIYKP YTDCCIKAKDSAVPIGGGNQLPQ
As_LIP80	As_8430	MAKLSTFPGLLLALLVATAVTRIEAEVDNIDYDFRDLQEGKAIWVRNTSEPGTITCRKDTISKVDNRN VSFERHFKNNKSFLESMKGEFFNWHHWEDPTHYPYDSMAVSNTEVNHKTDEILELLDAKNDCAVV KVIWDLNNEGFTVWRELRVSEAAITKSPSKECLDYFKLAAGGNRPRTYSPDCHQMP
As_LIP81	As_54075	MAAHMVLQAAAFGLLAAMVAVSSGANEENYDFKELLTEGKVLWVFNTTEPGSVSCRKDSILTVDVA VTFERSFKNNSIVRKENLTGKLINWYIEVSPKYDLEINDTQGELETQEILEYLNEEKTCAVVKMKV KEYSDGTTGIWRDLRLSNDAVEQAVPPECFQKFEQAIQFSKTRPRTAAYDSGCRGSA
As_LIP82	As_56070	MADARISSVLLFAFLLTAVSADNGVVQNSTVDVVKFVEENRKYVYRTTEPGNVTCRLDITYNTAT EALIFRSFLNGTRIEKNLKAFAVLKDELPHHEGPPFNAMEINDLEIQREYTLIEEVFEFQSDAGKCAVV MVLNDSGTATWLDLRLVSDTSVGTGPTKECTDEFEKILSYVTKPESRTSYSNECKRNS
As_LIP83	As_55222	MAALWVFRGVVFSLLVATVLTSDDWKDRNEKIDFRQMLKKSSVIWVNLNTEQGNVKCRKDLITDVNN THVSFDRIYEGEGTQOMQETLSGKLFNWNENPDNLNLPYDAIGVSKPGADPKIEEELDYLNKENS CAVVKVMNFATFNEQGMSSHVWRELRVATARLLDRETPAECLKYKDSVAASPKSSSRQVYFPDPCRSH
As_LIP84	As_56563	MPHRIFLAVVSALLYATESHGRKGGKTDIKEFVNTNDIWTYNTTAHKRLPCLMDFKKEVTEKDIQ FERHHFRQKRWATKLLIGKFGIWSPEKDKNSKPYDYMEVITFSQRKWVDAEILEYQHEDNTCGVFT VMSGNGIPPTFSHELVRVKNSSIQTGPKKECKKQFRKVVGEGKKIKGAYDHRCHALKYNNENFLHLP
As_LIP85	As_80499	MALFQTLVLSLFAFTAAVVEPREGTETSGDISKFLNANLKIWVYNTSEENKENITCRHVDVRYNITEGEI FFHRYNNTSTGELLKGNFVNWNDPKSKIYDAMYVTDARGIEVDEVIEYLSDDLCAVVSVTLMRD QPVVLEIRVSEKALPTGPDECKKKFDEILKEANKTAKSQSPACRTLHH
As_LIP86	As_17865	MYAIFLLTMFFTTEASDFTFQQLRRGFNTSIEQFYNTSEPIWTTFRADASFNLSCIVDTFNTLLNLY FERSYYWNQTRVKSRLRALNKKGTDRSSRTLVDARLSILIPQQPLFEEELLYQSDSSCGVFRYF KHPNVTRYDLRVNCSLAMKPHENCTMFFKGTSKLRPSILEKLGQERKVVYEPKCAILWSTDGCC
As_LIP87	As_58208	MADVRISMLFAFLLTAVSADDNDVQNSTVDVVKFLQSQPKLWVYNTTEPGNIKCRDLFYTTITDNO TSFSRSLNSTRKEEKLTKGKFFVYEGGNKPAKGYDAIEIKDDTTTEYPLSEEVFEFQSDDETC AIVAVLDNSSGITKTWLDLRLVSSAIESQKGGKCAERFEEILKIYHERDRKWRIYSYDSDCKTEQS
As_LIP88	As_120249	MEKPNIQVVAIAFLVGLCSSASQGTQERIDSGELDIYKFLSGADGILTLSSSKEGLRCRYDAIGVTTIN VTLTRYLLERSNREHLLRQNRPQRLLELGRFNKWNNSNFDVQEKYDVTVLTNKGGGYSTETIYHKN KCAVFSVSGSRNAIGSRTPKILEVVRWKYSGGSESDQLLCLQDAKNFLRQQPSTQPKLLGDQSCLPK TATPESA
As_LIP89	As_56677	MADARICSVLLFAFLLTVSSEGEDVEISNDIVKFLTESPKLWVYNTTEQSNITCRLDVYNSVTANET SFSRFSRNGTREEEKNLTGTFFLWEGGNLADGKYDAIKISDGIQREYPLSEEVFEFQSKDGTCAVV AVLDNSSGMTKTWLDLRLVNSAIESEPGKECSGKFEELQAYHVRDRTRWRIYSYDRECKNKL
As_LIP90	As_120181	MKTRMLLIVALLIAATESSEGVASETEESVRVTIQEFYSGTDPWIYGVSGRNSHPCKIDTNTIKSGT NTFFRRYSHDGSQWRWNSITAQLSGESDFTVSYDDGERHSEKMLAASTNGCALFEVAVRNTGAI NTPLQKTRLRNFRPTSEIDFRVKRTDGHQSQFRKICFSSVVKARKERKLPDVSQIEDFANIRCSQLC QGDACHPEMLREAFN
As_LIP91	As_56843	MSLLHTFLVCLFAFAVALAEPKEGTETNDIHKFLNANDKIWVYNTSEENEKIDTCRYDVRVTITQKGI SFHRYNNTSTPELLKGEFLNWNPKSNTYDAMYVTDAAANNSVTDEVIEYVSSDNCVAVSVSLMRT AHTETLREIRVPENALENEPEEGCKKKFEEILRLANKTARSHYSPSCKNKSTSTLLTQ
As_LIP92	As_7196	MAKLSTFGGLLALLVATAVTRIEAEADNTDYDFREFLQDQKVIWVRNTSEPGTTCRRDTISKINERS VSFERQFKNNSQTFRESLKGELFNWQHWEDETHYPYDSMAVSNTEADVHKTDEVLDAKNDCAVI KVIWDPKSVGFSTVWRELRVSAALKKNPSDKCWNYFKDAVKKETSRRPSYSPSC
As_LIP93	As_113700	MPYRIFLWVVFALLYATESHGRKGGKTDIKEFLNNTDIIWYNTTAKKRLACLMDFKKEVTEKDIK YRHHFLGQMRWATIDLIGRFNIWHPEDRKTSKPYDYMEVKTFKHMKWVDDEILEYQHHDNTCGVFT VMSYGGIRPSSISHELVRVKNSSIQMGPKHECKEQFRKVLGKGEKIKRVYDHRCHKALKNHNEHFLHLP
As_LIP94	As_99270	MTTNGSISSYPISPLSRVACKTDYGIEIKTPYISFYRYVSESARNQPSSKISSRPSVQHYQVTGLTY ALNVEEGKPFDAISIHAYESGARFLEAVEYQNDENTCALFSVRPLDTSVRVRTPTDRDLRIKTEALGKE QKKKCFMSALEYIRNTDKSRTVTELSRLSVCERSQSNNECSKSLENNMK
As_LIP95	As_54682	MVTLISAIMQSSAMMESSDGSKENAIDISKFLSTDGIKSLMVVFSISPTSLEPCKVDFHTHTAANGTFF VRYYLQEQSTNRRRTAGRPPVGRFSAGRKREVLVGOFTNIGKAQEPKSDTMNVFKYGDWEIQRP APGKLPQDVETIEHQSDDDACAFSTSRRTAGAPRPLLDLFRVRAEIKNEGLACLTSIEQHVKKMKA NHPEIIVPTIDNIISCCQKCLTASHCKLSDEAEL
As_LIP96	As_7788	MRTLFLTGVLFITADFAHSISRYEPIFDNEADYVQRQDIYKAFNISPDFYWLFGFSYHSNHTENKSCV YFDIEDISESGMNWSSNFIKHGVRSKLCIEPPVPLKARLTYCPAEGK
As_LIP97	As_80236	MNCFIVAILLTVCAAAGTKPTVETLQRALETREIRIWTVRSYDRYTNKKHRCVYAYKTFEQGYDYK FDQFYQYADKVVSNPIYGRISDGTGPKLQKISKELGGEGVPTLLHWNEQQNCGILVFNKSTSDRY ECELHVWEPLNLRNGINPCDHEYDEHCPVREKYSSYEPDCLPQLG
As_LIP98	As_6953	MNSFTIAVFLALGVTALGYTPSVDLLYKALNTTGRIWTVLRSYKRFGTGKEHTCVSVENPILTKTSYEF DQYKVDGTRIRHLYGEGTRGAYGPVLTVRNETGRQGIPYTLRYWNEDKHCGFLTFPNIENGEYEC ELHVWEQELVRDDANPRNFPCEKDYQDFCGHGRKHIVYARDCLN
As_LIP99	As_10928	MNFITISIFLTLGGTAIGAEPYTLQDLRDALGTPYPFWMNFRFTKRSKQGEHKCIFSQKINLNDYV

FTQGFKSGSGWHYKTLFAALVPGEGNGLEPILDVSEKQGKGRD

As_LIP100 As_50127 MSCFSFATLLALCAAALGTDAATLEDLKEALGTRDPIWVLRNFYSIAGTVDCVYDKVTSNSDNEYQ
LEHGYYRVEGHVKKVTLVGTAGQGRTESYLDVKEGTEGHEKNRHFLKYWNNDKKCFILSFHNRSTSGQT
QCQLHLWAQSVSYTASGVSQDKYDDICGKTKYEVYKSDCV

As_LIP101 As_31300 MIFPLAFCLGVIATAAAAADAAGTNEHPKDLDMMKMAEVNETLVVVKRKHNETKLRCQSAKKTRTIN
QTLYQYTLRARYMNGTEYQYAEVEEVQVKLVPLRLGGGYRSKYTDKHGVKFSRLRLRKMNDGNCFFV
MVNKNNGTRGCELLVPASKITQNIATECDSYNTTCKGDSLQLYESSCTYDQQPGC

As_LIP102 As_32490 MSFFSLALFLALGVTAFGYTSPTLEDLREAISTNQNIWVTRQTYERKTGGQNHRCVYVVRKTSAPDQ
PYEFDQYFRHGETWVNDNIHLADISQEPKDPVLTVRNGKGSRGIPYSFLFWRNDRSHCAILSFRNEKG
QNEFELHWVNDLDRHPGVVQQQELHRISEGRELSIYENDCKPN

As_LIP103 As_10679 MSCFTLALFVALGVTAAGSTNKPTLEDLKAIFIDTDEKIWLFRRTYEEQTDGRNHRCVYALKRPSGS
QYVFEQYFRHGVWESRYFYTDIREGTGGKPVLTVRNSTESKGTPTLLFWDSDRSRSHCAILTWPNAK
GKNEVEMHVRNNDLRSPDLDDCSREFVNLAEGRDKTIYEGDCKS

As_LIP104 As_12499 MNGFSIALLLALSVGTIDALLTKKDLNALSTPRRIFAVQRTFERSGDKGKHTCVYAIQTHLQDDDYQF
EQHYKEGPIGRANYLYGKLSGDKGKPVLTVSYEQGREGTPYTLTYWDPRRHCAILEFEKGETRCEL
HVWEDDFASTSTPCDHEYERYCGPVKYEVSQRTCLR

As_LIP105 As_102988 MIFSLALCLGLMAAATAVQTTEDGQESRDLDMNMVNEVNETLVVYKHKHTQNTKHRCLSAKKQHRE
SDTQYQYELRLRGLGGQYQAQTVNVTLDKDDGGVYKSRYTDRDGINYNLTKEMDPAGPCFVIFVVK
DDGNKGCELLVTASVNGTIPEKCKNYYSQHCEGKSSTLHRPRCIYN

As_LIP106 As_31301 MIFPLAFCLGVIATAAADTAGKNEHPKDLDMMKLAEVNETLVVVKRKHETQTLRCQSAKKTETINQ
TLYQYTLRARYMNGTQYQYAEQVQVKLVPLLSGGGYRSKYTDKDGKVFSLRLRKMDEAGNCFV
MVNKNNGTRGCELLVPASKITQNIAPVCDSSYNTTCRGDSLQLYDSNCTYDQQRGC

As_LIP107 As_51348 MPCFPVAVLLVCATALGTDPTYTLKDVKALSTSDPIWIALRNFDSVAGTVDCVYDKVFKTDNDY
TLEHGYNQSKMHKMTLYATAGQEGNQAYLDVKESSQSGQGNRHVLRKHWDDSKNCFILEYQNTTAR
RVQCQLHLWASSVKSRYGRDFECQEYKDEICGDTKYKVKYKTTCTKLV

As_LIP108 As_81142 MIYFITVFLALGVSAATGYHPTWENLKNALNTQDRIWVTHRSYERYTDDKKHDCVYALKQSLTGVNY
EFSQTYDYGTKRITRELHGEISQEGPDAVLTSEKAGQKIPYTLRYWDEHRHCGVLTFTDQDGVFK
CELHVWEGNLPVSGEYPCLELQYRQICPSLHYSQVHSPDCLSR

As_LIP109 As_44528 MAYAAVSIFLALCVAGSGATPTMDELKQALSTHDAWVNRFTFNPRSGSDQHQCYYVEIGTLMEDRY
DFTQHYVWVAGQWHSNRLYAKLKDRQGTGQPGFLVNVQKEGQETGRLYLQYWDGNNRGCGILKVAL
GGEDHYELHWVNEQVARVPRSCSEEEYKTLGRKQFKLYSYDCANKRVAG

As_LIP110 As_110797 MKCFTIIVILALGATSPAYSANVGDLYKALSTTGRIWTLRTRTYQRFTGTKEHRCISVEKTFLNKTYEF
DQYKVDGTRIGHLYGEFTRGPYGPVLTVRNETGRQGIPYTLRYWNEDEKHCGLTFPNIETREEEC
ELHVWEQELDRADANRQKFPCEEKYNNDYCGDHRKHKHYAEDCLY

As_LIP111 As_56745 MNSFAIALFVVLGSAIAAHLKPEDLYNALNTPHVSVMVARSYKRPEQGEQNKCVHSRRLTLTRNDY
TFYQYKHGSTQEQLLYAGLTTDRATGEPAMDVNETPGQSGQLYTLRFWDKVNQCGLTVDVKGK
PECEMYAWDANVRNPSALSECDFKSELCSRSGMRVEVFIRGEC

As_LIP112 As_10927 MNFITISIFLTLGGTAIGAEPYTLQDLRDLGTPYPFWMNFRTFKRSDKGQEHKCFISQKINLNDYV
FTQGFKSGSGWHYKTLFAALVPGEGNGLEPILDVSEKQGYSKGRFTLRFWDRNHNHCGVLTHTDNDG
TMECELYAWKESVGNQQRATAACELQFQEI

As_LIP113 As_98550 MTMSCSQALWAIAMCSNIGNGNELQDNGNIDIRKAVNTTEKLWLFWQTYRNGFDMCGNGICIK
VNETCIYNMKNINITEKEYFNQITWIGKDDTTNRYRGTFKDGDIDPPKSMVEITEITEDPSDASFSGDDGQ
HQTWTLGYLEEIDGRCMVFSIGELDNEIKVLTGCDMYIKGTPKESDPPPGRFFEKRCNVTKIFQP
YTKACIKATDSAVASGDSNHLQP

As_LIP114 As_40031 MKWQLQVFWLLQSALCSYSGNEVEFLQSPLKENGENDIRKVLITNESLWLYQETYENGFTACSRTR
CLYEREICIRILMTHSHTQYNTFLKMNLDNTAVENSYFGTFANYTNPCKSLVLRDPLGGEEVQLTL
GYTDELNRRCVFTINSLGPALTPFLGTCEMYIRGEYPLSPSSGCKRFFEEERCNMTRIYKPYTAECKN
AINTPTINDIPN

As_LIP115 As_80509 MLVLFVAVYLQAAAAGNVPTLADLVEALNTTEKVVILRLSYDLYWQEPYDRRNCVYVEKNSFTQDA
YNFTQHYITDGEKTLQLRAKLLQKNRPVMMVILKPRPNVYKTLRSWNKEHKCGLLMFKDNNEMH
CEMYAWDSKVVDLNVTSNSCEDEYDNICKKQRKYDGNAYQCQRALNPPEPPAPACR

As_LIP116 As_31881 MRLCGSFSVFFIILVLCQAGKGEKLLKFKARGRNNLNITKVFEKGRWLWYKQTCNIFYFPEKTD
PVEFVQKCFIKTINISNTDVYFREMMDVNGDKVEILYRGDFIPEEEGAPKSMNVTDLSVNDTSPFEV
MELGYSNQGCYVFFLTRLHEDDEEDSSGRRNCEMYIRNHKPPGPKCENFFKTCNGEQFYEPYEE
SCKIKKNIKVARKNSE

As_LIP117 As_5660 MNSFTIIVFLALGVTAALGYTPSVDLKYALNTTGRIWTVLRSYKRFRTGTKEHTCVSVENPILTKTSYEF
DQYKVDGPRIRHHLGEGFTRGAYGPVLTVRNETGRKGIPYTLRYWNEDEKHCGLTFPNIESGEYEC
ELHVWEQELVRDDANPRNFPENKYKDFCGHDKRHIVYTKDCLN

As_LIP118 As_106712 MIFAHALCLSLIAVAAAETAASKQPEPIDMMHMVNSERLVVFARKHTTDPFRCQSAMKKSQDQIK
GYDYSKARNGSEYLEDNVHVTLLPLPVGSGYKSTYIDRTGTNFTLTLWKKDESGRCFVIFVEKNTG
EQGCELLVTASTRNETVPRDQCSFYRNTENCENEGTSKLYKDKCKYD

As_LIP119 As_21521 MIFSLALCLGAMAAAATAAQVPERGVESGDLMMNLVNVSETLVVYKHKHTQNTTVRCLS
AKKQYNTTLYEYILRGRNASKIYGEHSVNTLQEREIADKSTYISQNTYTLTKAMDPTGTFCFVYVKKDDNN
EGCELLTASVSGKIPKCKQEYYSKNCEGRNIRLHNDCKYEN

As_LIP120 As_21522 MIFSLALCLGAMVAATAAQVPERGVESGDLMMNLVNVSETLVVYKHKHTQNTTVRCLS
AKKQYNTTLYEYILRGRNPSKIYGEHSVNTLQESEGIAYKSTYTSKNITYALTKEMDESGPCFVIFVKKDDG
NEGCELLVTAAPKVSQKIPKCKQEYYSKNCEGRNIRLHNDCKYEN

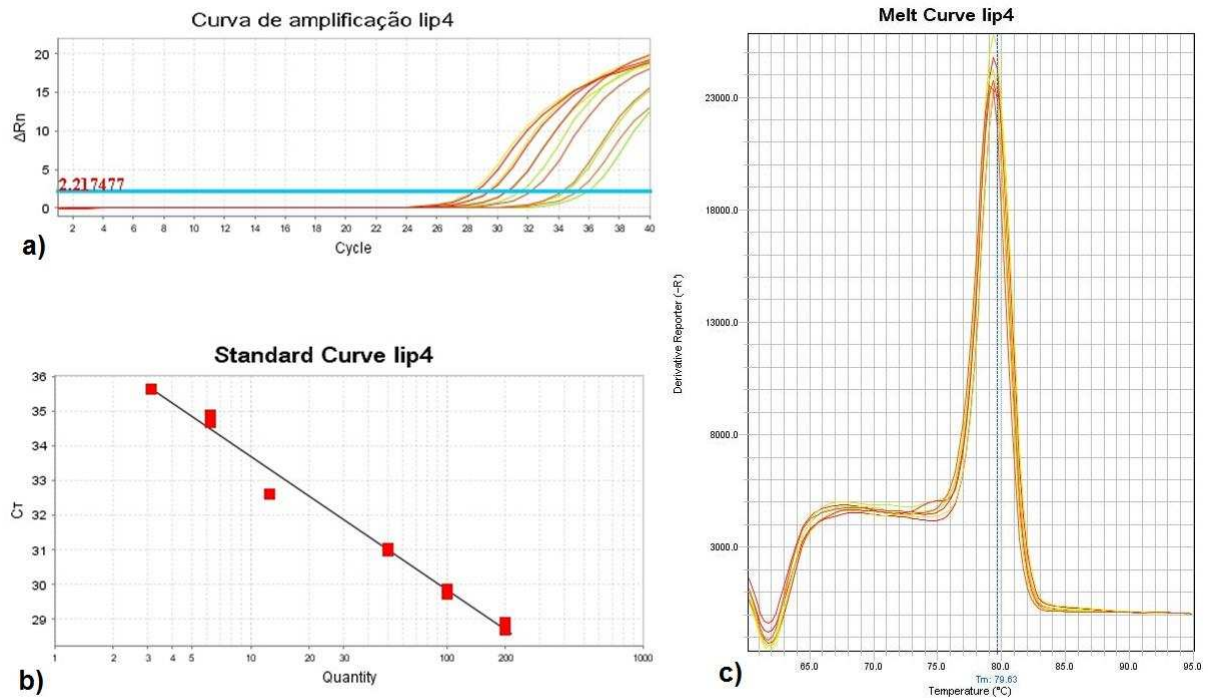


Figura S1: Resultados da análise de real time RT-qPCR do teste de eficiência do primer *lip4* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *lip4*.

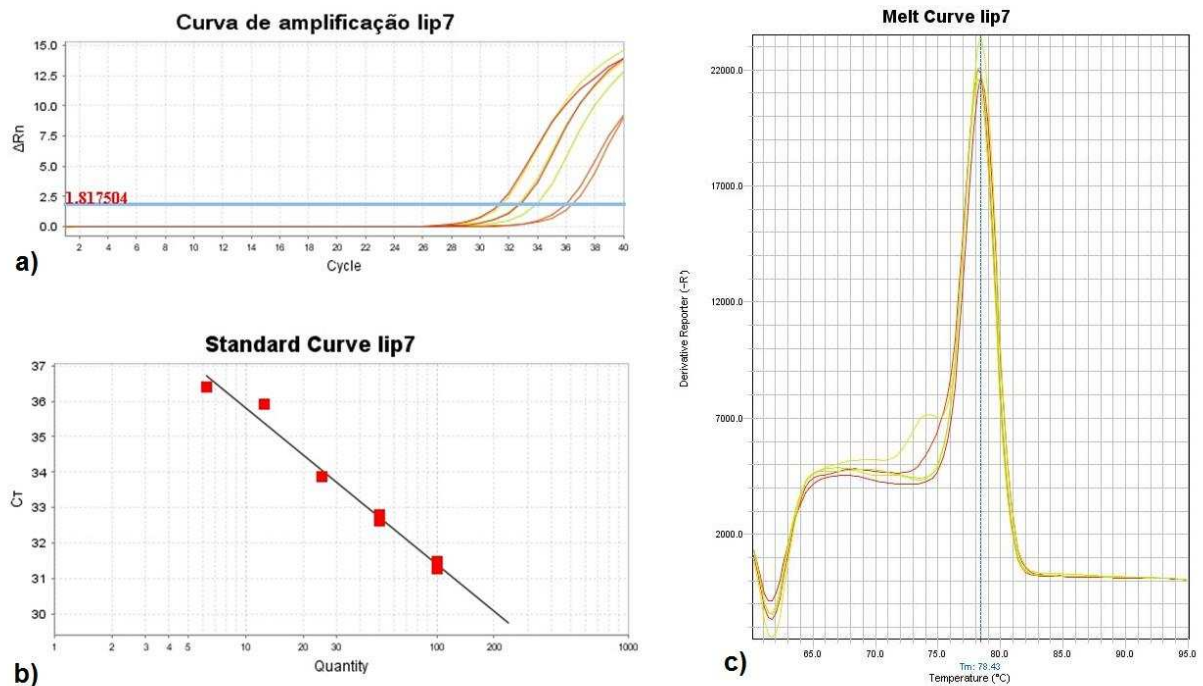


Figura S2: Resultados da análise de real time RT-qPCR do teste de eficiência do primer *lip7* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *lip7*.

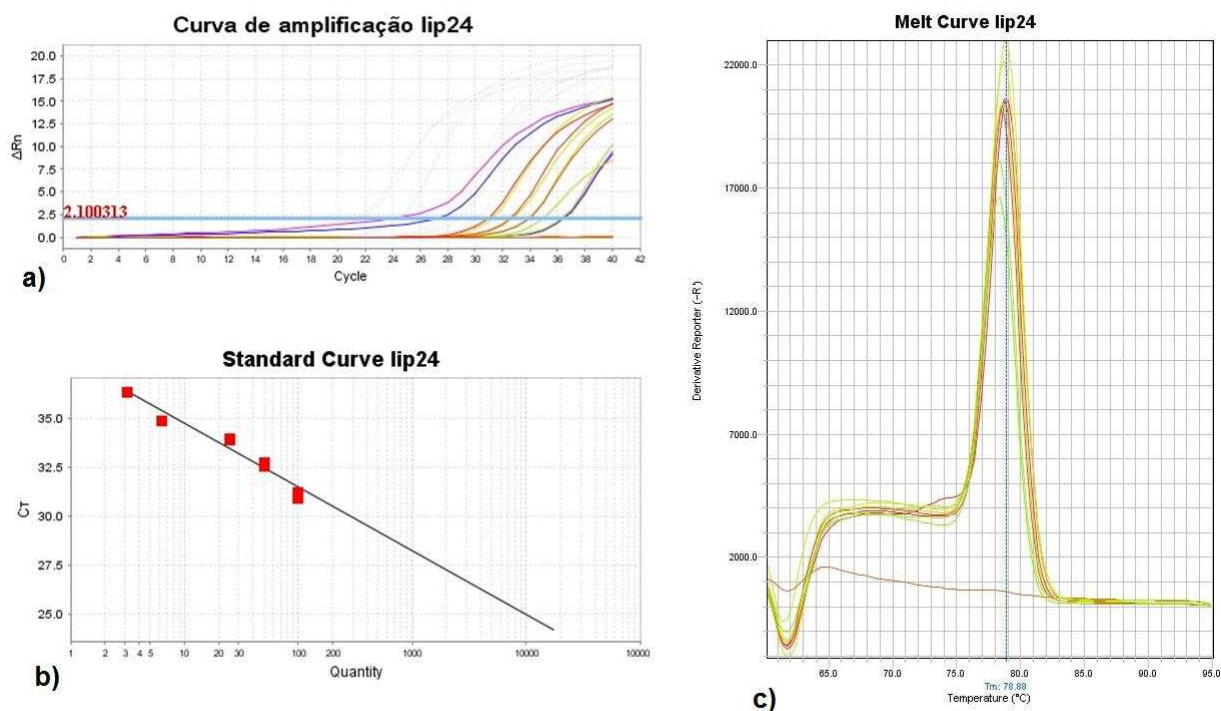


Figura S3: Resultados da análise de real time RT-qPCR do teste de eficiência do primer *lip24* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *lip24*.

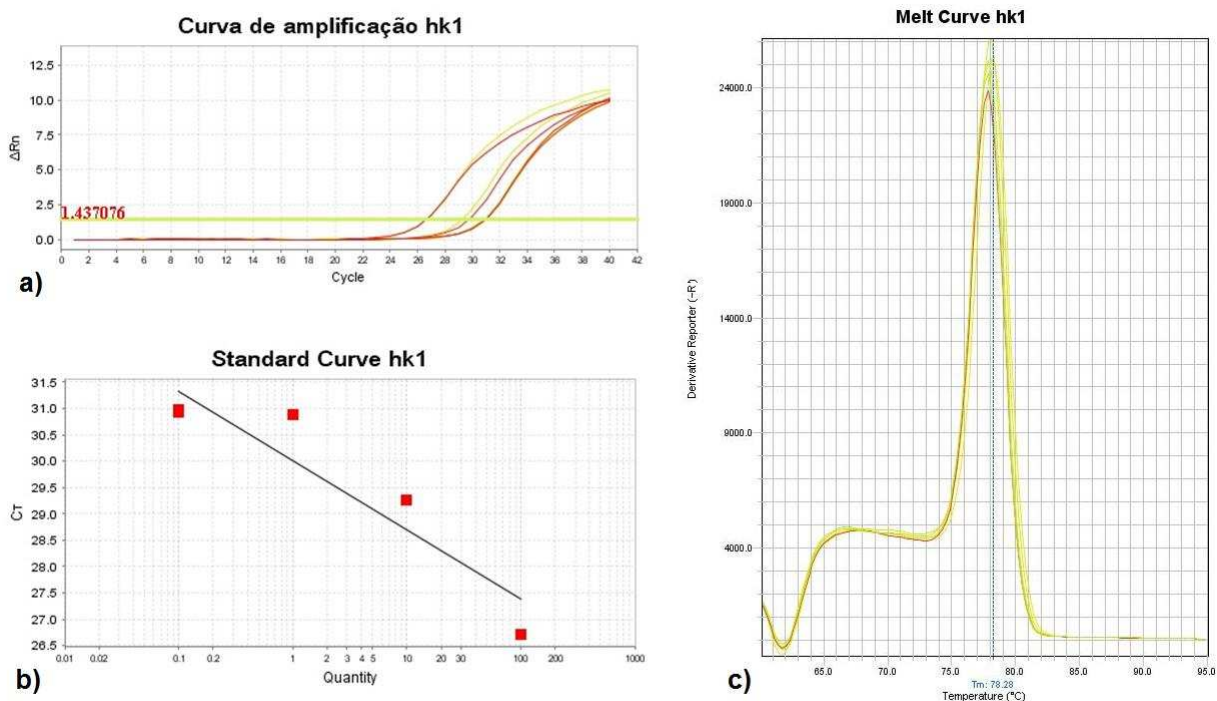


Figura S4: Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk1* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk1*.

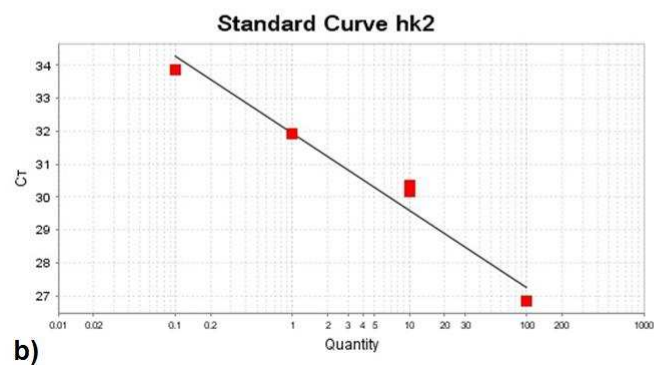
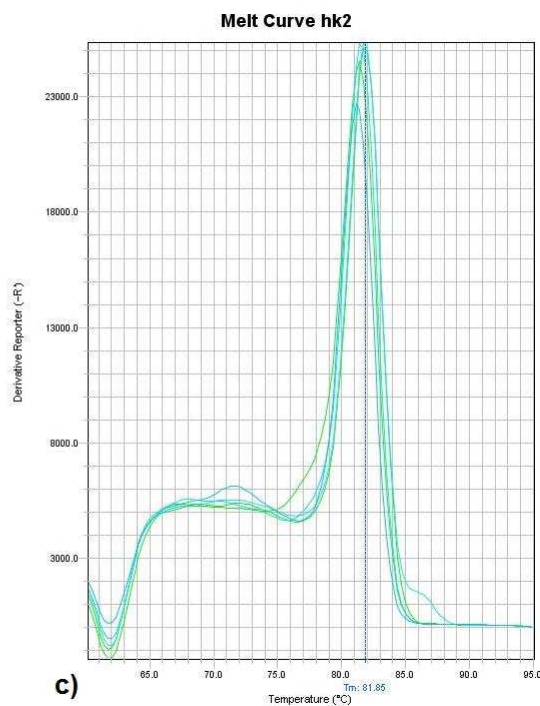
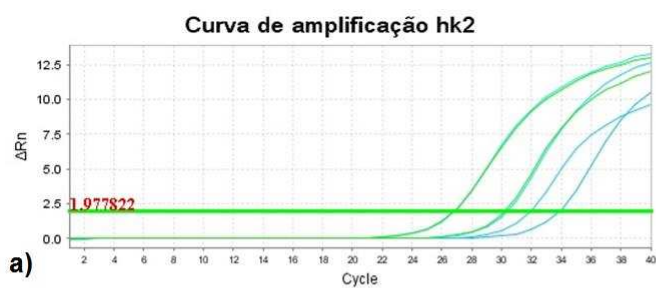


Figura S5: Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk2* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk2*.

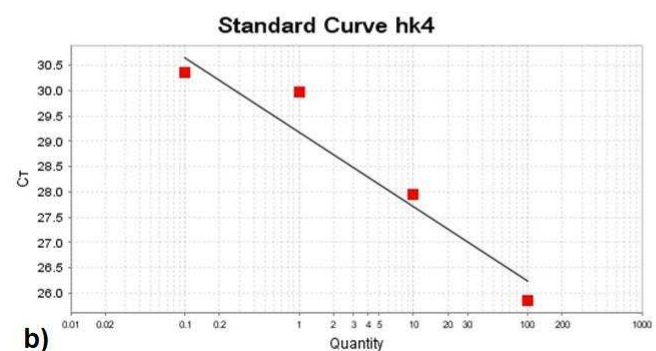
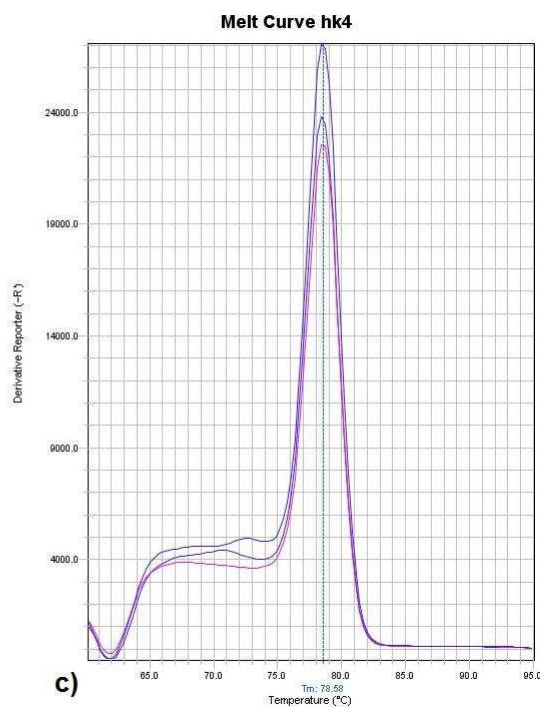
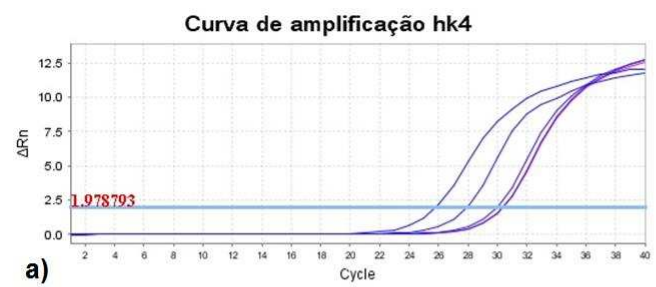


Figura S6: Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk4* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk4*.

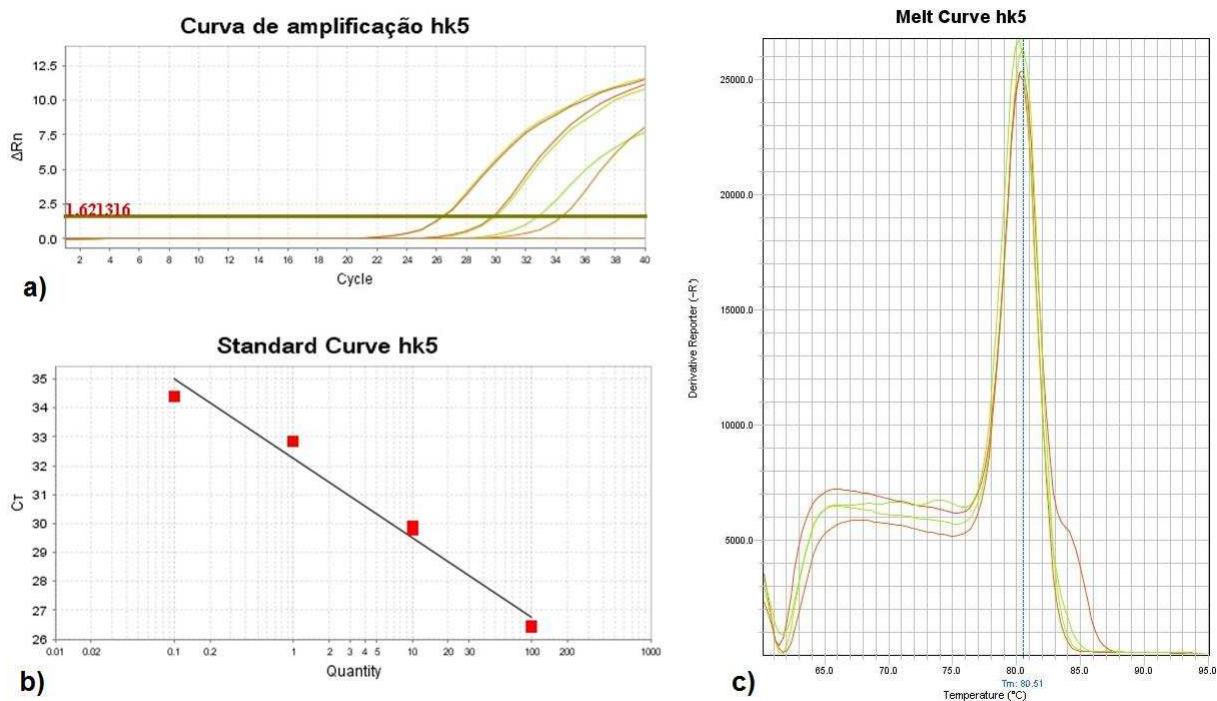


Figura S7: Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk5* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk5*.

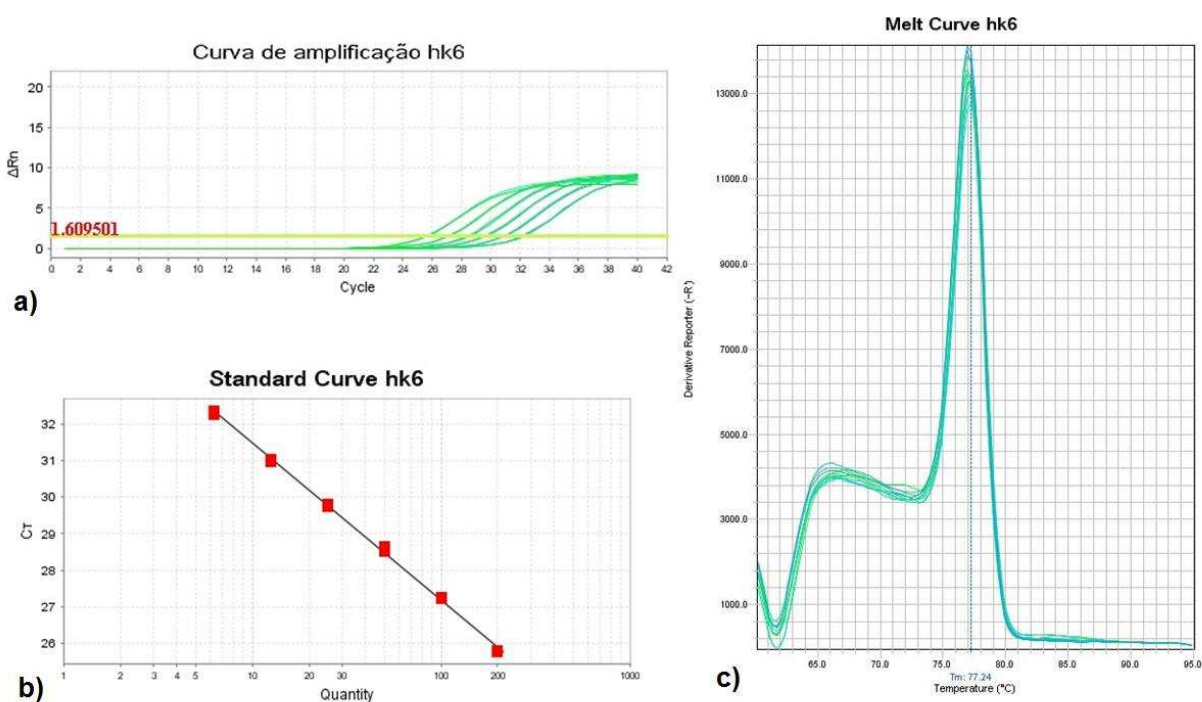


Figura S8: Resultados da análise de real time RT-qPCR do teste de eficiência do primer *hk6* fornecidas pelo aplicativo do aparelho *StepOne* (Life Technologies®). a) Curva de amplificação com o valor de *threshold* indicado; b) Curva de calibração; c) curva de dissociação (curva de *melt*) do gene *hk6*.