



RAFAELLA FABIANA CARNEIRO PEREIRA

ESTUDO DO GENOMA E VIRULÊNCIA DE LINHAGENS DE *Haemophilus influenzae* BIOTIPO aegyptius CAUSADORAS DE FEBRE PURPÚRICA
BRASILEIRA

CAMPINAS
2015

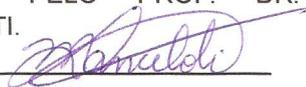
RAFAELLA FABIANA CARNEIRO PEREIRA

ESTUDO DO GENOMA E VIRULÊNCIA DE LINHAGENS DE *Haemophilus influenzae* BIOTIPO
aegyptius CAUSADORAS DE FEBRE PURPÚRICA BRASILEIRA

Tese apresentada ao Instituto de Biologia da
Universidade Estadual de Campinas como parte dos
requisitos exigidos para a obtenção do título de
Doutora em Biologia Funcional e Molecular, na área de
Bioquímica.

ORIENTADOR: MARCELO LANCELLOTTI
CO-ORIENTADORA: LUCIANA MARIA DE HOLLANDA

ESTE EXEMPLAR CORRESPONDE À VERSÃO
FINAL DA TESE DEFENDIDA PELA ALUNA
RAFAELLA FABIANA CARNEIRO PEREIRA E
ORIENTADA PELO PROF. DR. MARCELO
LANCELLOTTI.



CAMPINAS

2015

Agência de fomento: FAPESP
Nº processo: 2011/01319-5

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

P414e Pereira, Rafaella Fabiana Carneiro, 1985-
Estudo do genoma e virulência de linhagens de *Haemophilus influenzae* biotipo aegyptius causadoras de febre purpúrica brasileira / Rafaella Fabiana Carneiro Pereira. – Campinas, SP : [s.n.], 2015.

Orientador: Marcelo Lancellotti.

Coorientador: Luciana Maria de Hollanda.

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. *Haemophilus influenzae*. 2. *Haemophilus*. 3. Febre purpúrica brasileira. 4. Virulência (Microbiologia). 5. Genomas. I. Lancellotti, Marcelo, 1976-. II. Hollanda, Luciana Maria de. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Study of the genome and virulence of *Haemophilus influenzae* biotype aegyptius strains associated with Brazilian purpuric fever

Palavras-chave em inglês:

Haemophilus influenzae

Haemophilus

Brazilian purpuric fever

Virulence (Microbiology)

Genomes

Área de concentração: Bioquímica

Titulação: Doutora em Biologia Funcional e Molecular

Banca examinadora:

Marcelo Lancellotti [Orientador]

Marco Aurélio Ramirez Vinolo

Carlos Emilio Levy

Fabiani Gai Frantz

Wilton Rogério Lustrí

Data de defesa: 06-08-2015

Programa de Pós-Graduação: Biologia Funcional e Molecular

Campinas, 06 de agosto de 2015.

Prof. Dr. Marcelo Lancellotti

Prof. Dr. Marco Aurélio Ramirez Vinolo

Prof. Dr. Carlos Emilio Levy

Profa. Dra. Fabiani Gai Frantz

Prof. Dr. Wilton Rogério Lustrri

"Imagination is more important than knowledge. Knowledge is limited.

Imagination encircles the world."

— Albert Einstein

AGRADECIMENTO

Este texto é dedicado a todos aqueles que à sua maneira, direta ou indiretamente, contribuíram para este trabalho ao longo de seu desenvolvimento e conclusão. Sinto-me honrada por compartilharem comigo risadas, conselhos, sabedoria e momentos inestimáveis. Assim, deixo aqui meus sinceros agradecimentos.

Todos que me conhecem sabem que devo uma vida de agradecimentos à uma senhora de fibra forte e cabelos branquinhos: Dona Eva. Vó, muito obrigada por tudo! À minha mãe, Márcia, e irmã, Marcella, devo também profundos agradecimentos pelo apoio, carinho e compreensão. A vocês três agradeço por ter me ensinado a andar, a sonhar e a viver. Amo vocês.

Agradeço também ao meu pai Ademir, Yara e irmão Gabriel pela torcida e carinho.

Gostaria de agradecer minha tia Rose, “cuirmão” cunhado Carlos Augusto, primas e primos, especialmente Andressa e Alessandra, e tio Cicinho. Sou grata a todos meus familiares, humanos e caninos - Tequila, Buddy, Martini, Cacau, Cascão, Sherlock, Valentina e Athila – seria impossível não citá-los aqui.

A todos meus amigos, obrigada!

À Tati, pela amizade, compreensão e por me aguentar há tanto tempo. Que estes 16 anos de amizade sejam infinitos.

Dani, minha amiga, obrigada pela risadas e companheirismo. Vida longa e próspera!

À Raíssa, Sarah e Nani, pelas histórias, conversas, confissões, risadas e anos de amizade. Rafa, xará de nome e amiga do peito, sou muito grata por sempre estar presente, nos bons e maus momentos. Amigo é aquele que sempre está ao seu lado, independente

da distância, por isso sou muito sortuda em tê-las como amigas. Obrigada.

Aos meus demais amigos de graduação e pós, em especial Lari, Novo, Lucão, André e Daniel, pelas piadas, carinho e torcida.

Às minhas amigas Val, Julia e Carol, pela momentos de descontração, alegrias, camaradagem, conselhos e bolachas de provolone.

Rubens. Profundamente agradecida sou eu. Obrigada pela cumplicidade, companhia e inteligência. *May the force be with you*, sempre.

À Begoña, pela colaboração e entusiasmo que foram fundamentais para as análises de massas deste trabalho. *Gracias* pela amizade tão especial, acima de tudo.

À Professora Doutora Alexandra Christine Helena Frankland Sawaya pela colaboração e dedicação.

Ao Laboratório de Genômica e Expressão, especialmente Marcelo Falsarella Carazolle, Luciana Souto Mofatto, Juliana José e Professor Doutor Gonçalo Amarante Guimarães Pereira, pela amizade, colaboração, dedicação e conhecimentos aplicados neste trabalho.

À Professora Doutora Clarice Weis Arns, pela oportunidade de participar do Programa Estágio Docente durante meu Doutorado. Obrigada pela experiência e conhecimentos transmitidos.

À Professora Doutora Fernanda Ramos Gadelha, pela cordialidade, incentivo e contribuição para o aprimoramento deste trabalho.

Ao Professor Doutor Marcelo Brocchi pelo exemplo e conselhos sábios, não apenas nas etapas finais de meu Doutorado, mas também ao longo de minha formação acadêmica. Aos membros do Laboratório do Professor Brocchi, pela amizade e colaboração.

Ao Professor Doutor Marco Aurélio Ramirez Vinolo pela atenção e

conhecimentos de grande valia para este trabalho. Aos alunos do Professor Doutor Marco Aurélio, também deixo meus sinceros agradecimentos, especialmente à Aline.

À Professora Doutora Fabiani Gai Frantz e ao Professor Doutor Wilton Rogério Lustri, sou grata pelas colocações pertinentes e disponibilidade para tornar este trabalho melhor compreendido.

Ao Professor Doutor Carlos Emilio Levy, pela sabedoria e entusiasmo inestimáveis que enriqueceram de forma significativa este trabalho.

Agradeço também meus colegas/amigos do LABIOTEC pela colaboração, risadas e momentos de descontração que foram fundamentais para o desenvolvimento desta tese: Gisele, Mário, Maria Cecília, Cibele, Felipe, Washington e Tobias.

Ana (garota-plasmídeo) e Danilo, muito obrigada pela amizade, apoio, descontração e convivência, principalmente nos dias e noites de laboratório.

Thaís, a parte experimental *in vivo* deste trabalho não teria sido feita sem você; além do auxílio, agradeço pela alegria e estilo contagiante.

Daisy, gostaria muito de agradecer pela ajuda (acadêmica e automotiva), companhia, sugestões, cappuccinos e amizade.

À Professora Doutora Luciana Maria de Hollanda, pelos ensinamentos, críticas, sugestões, elogios, disponibilidade, atenção e estima.

Professor Doutor Marcelo Lancellotti. Além de orientador, mestre e professor, tenho o orgulho de chamá-lo de amigo. Faltam-me palavras para agradecer pelos ensinamentos durante tantos anos de parceria, entusiasmo, incentivo e compreensão.

Ainda, agradeço também a todos do Departamento de Bioquímica, alunos, funcionários e professores, em especial ao Cleyton, Bruna, Verônica, Márcio, Erika, Juliana Mattoso, Juliana, Bernardo, Marina e Daiane, que tornaram este lugar mais amigável e especial de se trabalhar.

A UNICAMP, finalmente, lugar do qual tive o orgulho de chamar de casa desde 2005. E os funcionários do programa de pós em Biologia Funcional e Molecular, Andréia, Sílvia e Beatriz, prestativas e sempre tão gentis.

Às agências de fomento CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico – processo 141710/2011-0), que financiou os meses iniciais do trabalho, e FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo – processos 2011/01319-5 e 2012/15046-3), pelo financiamento e confiança.

Por fim, se me esqueci de alguém, peço-lhe desculpas e agradeço-o em dobro! Esta tese foi resultado de um trabalho em equipe, pois todos foram importantes para meu crescimento profissional e pessoal. Por isso, serei eternamente grata.

*“My dear, here we must run as fast as we can, just to stay in place. And if
you wish to go anywhere you must run twice as fast as that.”*

— Lewis Carroll, *Alice in Wonderland*

RESUMO

A Febre Purpúrica Brasileira (BPF) é uma doença infantil sistêmica caracterizada por conjuntivite, febre, púrpura e sepse. Seu agente etiológico é o *Haemophilus influenzae* biotipo aegyptius (Hae), uma bactéria anteriormente somente relacionada a casos de conjuntivite purulenta. Os fatores de virulência desse patógeno ainda não são totalmente conhecidos. Assim, o presente estudo teve por objetivo aprofundar o conhecimento sobre esta bactéria. Para tal, nove linhagens de Hae, das quais sete são causadoras de BPF, foram comparadas quanto ao seu genoma, à quantificação da expressão de possíveis genes de virulência e ao perfil de lipídios em infecções *in vitro* e *in vivo*. Linhagens causadoras de BPF apresentaram genes ausentes nas demais linhagens Hae estudadas, tais como *crgA* e *hadA*. Hae não foi competente ao processo de transformação e análises dos genomas de Hae revelaram a ausência de genes relacionados à competência deste processo no gênero *Haemophilus*, o que justifica a incapacidade de manipulação genética nesta bactéria. Os genes *hadA* e *tabA* foram expressos de forma distinta em linhagens de Hae nos hospedeiros *in vitro* e *in vivo*, não sendo possível inferir seu papel na virulência de Hae. Enquanto que o gene *las* parece ser modulado negativamente em contato com o hospedeiro. Apenas linhagens causadoras de BPF foram recuperadas de ratos Sprague-Dawley, o que sugere a presença de determinantes específicos de virulência que auxiliam na sobrevivência no hospedeiro nestas linhagens. Dados de espectrometria de massas dos lipídios totais mostraram modificações em Hae após a passagem *in vivo*. O conjunto de resultados mostra que, apesar de semelhantes entre si quanto ao genoma, as linhagens analisadas de Hae associadas a BPF não apresentam característica clonal. Tais linhagens apresentam fatores de virulência específicos possivelmente correlacionados à ocorrência de BPF, os quais são candidatos promissores para estudos futuros sobre a doença.

ABSTRACT

Brazilian Purpuric Fever (BPF) is a fulminant pediatric disease characterized by conjunctivitis, fever, purpura and sepsis. The illness is caused by *Haemophilus influenzae* biotype aegyptius (Hae), a bacterium which was formerly associated with only self-limited purulent conjunctivitis. The virulence factors of Hae are not yet fully known. Therefore, this study proposes a better understanding of this pathogen. Nine strains of Hae, of which seven are associated with BPF, were compared for their genomes, expression of candidate genes for virulence and characterization of Hae's lipids in *in vitro* and *in vivo* infections assays. Hae associated with BPF presented genes that are absent in other Hae strains studied, such as *crgA* and *hadA*. Hae was not competent to transformation process and analysis of Hae genomes revealed the absence of genes related to competence in *Haemophilus*, which could explain the inability to genetically manipulate this bacterium. *hadA* e *tabA* genes showed distinct profiles of expression in Hae in hosts *in vitro* and *in vivo*, therefore it was not possible to infer their role in Hae's virulence. While *las* gene seems to be modulated negatively after interaction with the host. Only Hae strains associated with BPF were recovered from infected Sprague-Dawley rats, suggesting that these strains acquired specific virulence determinants which assist in survival in the host. Mass spectrometry data of total lipids suggest that Hae suffers modifications after passing through the animal model. The findings showed that the strains analyzed associated with BPF not exhibit clonal characteristic, despite the similarities in their genomes. These strains have specific virulence factors possibly correlated with the occurrence of BPF and are promising candidates for future studies on the disease.

LISTA DE ILUSTRAÇÕES

Figura 1. Metodologia inicial para concatenar *contigs* de Hae 258/86.

Figura 2. *Pipeline* utilizado na montagem dos genomas de Hae Karina, 219/86, 284/86, 321/86, 406/86 e KC1018.

Figura 3. Esquemas dos genes *las*, *hadA* e *tabA* de Hae 254/86 feito no programa pDRAW com as regiões de anelamento dos *primers* específicos. As sequências foram obtidas do Genbank – número de acesso - FQ670178. Os fragmentos amplificados apresentam um sítio de corte para uma determinada enzima de restrição, nos quais será inserido um cassete de resistência à eritromicina.

Figura 4. Esquema da metodologia utilizada para a obtenção dos vetores *plas.eri*, *phadA.eri* e *tabA.eri* para os ensaios de *knockout* gênico.

Figura 5. Esquema da metodologia utilizada para a recuperação de bactérias após infecção em células humanas.

Figura 6. Esquema da metodologia utilizada para a recuperação das bactérias após infecção em animais.

Figura 7. *Pipeline* para a estratégia computacional e experimental para identificar genes de referência para Hae e Hi.

Figura 8. Distribuição dos *reads* resultantes do pirosequenciamento do DNA total de Hae 258/86 de acordo com seu tamanho em bases.

Figura 9. Distribuição dos *contigs* obtidos na montagem 1 de novo de acordo com seu tamanho em bases. *O *contig* 19 está em destaque e será discutido em um tópico à parte.

Figura 10. Cobertura genômica dos *contigs* da montagem 1 pelo gsAssembler Newbler 2.8.

Figura 11. Alocação dos genes anotados (CDS -*Coding Sequences*) através da plataforma RAST em subsistemas.

Figura 12. Comparações pareadas entre os genomas de Hae e Hi. Cada bloco colorido representa o número de bases compartilhadas entre os genomas. Quanto mais azul o bloco, mais similares são as sequências entre si. Quanto mais vermelho o bloco, maiores são as diferenças entre os genomas. 1. Karina. 2. 219/86. 3. 254/86. 4. 258/86. 5. 284/86. 6. 321/86. 7. 406/86. 8. KC1018. 9. F3047. 10. RdKW20. 11. R2846. 12. 10810.

Figura 13. Organização estrutural de *crgA*, *mdaB* e *rep* em Hae-BPF; *mdaB* e *rep* em Hae não-BPF e Hi, *crgA* e *mdaB* em *Neisseria* (fora de escala). As setas indicam o sentido de transcrição dos genes.

Figura 14. Clusterização das sequências de *crgA* de linhagens que apresentam o gene disponíveis no NCBI pelo método UPMGA. O grupo de *Neisseria* e *Haemophilus* spp. é subdividido em dois clusters: um composto por *Neisseria* e outro por Hae e Hi.

Figura 15. Gel de agarose a 1,5 % com os *amplicons* resultantes com os *primers* descritos na Tabela 12. Padrão de peso molecular 1 kb. Amplificação do plasmídio de Hae 2.

Karina. 3. 219/86. 3. Karina. 4. 254/86. 5. 258/86. 6. 321/86. 7. 406/86. 8-10. Branco (controle negativo) da reação.

Figura 16. Variação da expressão dos genes candidatos à referência para Hae e Hi em diferentes linhagens e condições experimentais. 254 S (isolada após infecção in vivo por 24 horas), 254 A (isolada após infecção em células A549), 258 (estoque), 258 S (isolada após infecção in vivo por 24 horas), ATCC 11116 (estoque), β -lac (estoque) e β -lac H (isolada após infecção em células Hec-1B).

Figura 17. Expressão relativa de *hadA*, *las* e *tabA* em Hae 254/86. *gyrA*, *rpoC* e *rpoD* foram usados como genes de referência para normalização dos dados e a amostra controle como calibradora (linhagem bacteriana sem passagem em célula ou animal). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por one-way ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação às demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

Figura 18. Expressão relativa de *hadA*, *las* e *tabA* em Hae 258/86. *gyrA*, *rpoC* e *rpoD* foram usados como genes de referência para normalização dos dados e a amostra controle como calibradora (linhagem bacteriana sem passagem em célula ou animal). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por one-way ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação às demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

Figura 19. Expressão relativa de *las* e *tahA* em Hae KC1018. *gyrA*, *rpoC* e *rpoD* foram usados como genes de referência para normalização dos dados e a amostra controle como calibradora (linhagem bacteriana sem passagem em célula ou animal). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por one-way ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação às demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

Figura 20. Expressão relativa de *las* e *tahA* em Hae ATCC 11116. *gyrA*, *rpoC* e *rpoD* foram usados como genes de referência para normalização dos dados e a amostra controle como calibradora (linhagem bacteriana sem passagem em célula ou animal). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por one-way ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação às demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

Figura 21. Expressão relativa de IL-10 e TNF- α em células A549 infectadas por Hae 254/86, 258/86, Karina e KC1018. *GAPDH* foi usado como gene de referência para normalização dos dados e a amostra controle como calibradora (células não infectadas). Os ensaios de PCR em Tempo Real foram feitos em triplicatas para cada amostra. A análise estatística foi feita por one-way ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação às demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

Figura 22. Expressão relativa de IL-10 e TNF- α em células HaCaT infectadas por Hae 254/86, 258/86, KC1018, ATCC 11116 e Hi β -lac. *GAPDH* foi usado como gene de referência para normalização dos dados e a amostra controle como calibradora (células não infectadas). Os ensaios de PCR em Tempo Real foram feitos em triplicatas para cada

amostra. A análise estatística foi feita por one-way ANOVA com as amostras comparadas entre si com pelo método de Tukey. $p < 0,05$. As barras indicam o desvio padrão das amostras.

Figura 23. Expressão relativa de IL-10 em ratos Sprague-Dawley infectados por Hae 254/86, 258/86, KC1018, ATCC 11116 e Hi β -lac. *GAPDH* foi usado como gene de referência para normalização dos dados e a amostra controle como calibradora (animais não infectados). Os ensaios de PCR em Tempo Real foram feitos em triplicatas para cada amostra. A análise estatística foi feita por one-way ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação as demais amostras. As barras indicam o desvio padrão das amostras.

Figura 24. HCA dos dados de ESI(+)-MS de Hae-BPF 254/86 e 258/86 segundo o método de Ward e distância Euclidiana ao quadrado. Pode-se observar o agrupamento das amostras em duplicata do estoque, da infecção em A549 (amostras A), da infecção em Hec-1b (amostras H), da infecção em HaCaT (amostras R) e da infecção em Sprague-Dawley (amostras S).

Figura 25. HCA dos dados de ESI(+)-MS de Hae KC1018 e ATCC 11116 e Hi β -lac segundo o método de Ward e distância Euclidiana ao quadrado. Pode-se observar o agrupamento das amostras em duplicata do estoque, da infecção em A549 (amostras A), da infecção em Hec-1b (amostras H) e da infecção em HaCaT (amostras R) e, no caso de Hi β -lac, também da infecção em Sprague-Dawley (amostras S).

Figura 26. PCA dos dados de ESI(+)-MS das amostras de 254/86: *Scores* – amostras e *loadings* – íons. Pode-se observar a separação das amostras, principalmente as de origem animal (S) em relação às demais.

Figura 27. ESI(+)-MS *fingerprints* das amostras de 254/86. Os íons que se destacaram nestas amostras estão em evidência.

Figura 28. PCA dos dados de ESI(+)-MS das amostras de 258/86: *Scores* – amostras e *loadings* – íons. Pode-se observar a separação das amostras, principalmente as de origem animal (S) em relação às demais.

Figura 29. ESI(+)-MS *fingerprints* das amostras de 258/86. Os íons que se destacaram nestas amostras estão em evidência

Figura 30. PCA dos dados de ESI(+)-MS das amostras de KC1018: *Scores* – amostras e *loadings* – íons. Não é possível observar a separação das amostras, exceto as de origem de infecção em células A549 (A).

Figura 31. ESI(+)-MS *fingerprints* das amostras de KC1018. Os íons que se destacaram nestas amostras estão em evidência.

Figura 32. PCA dos dados de ESI(+)-MS das amostras de ATCC 11116: *Scores* – amostras e *loadings* – íons. É possível observar a separação das amostras de origem de infecção em células A549 (A) e HaCaT (R).

Figura 33. ESI(+)-MS *fingerprints* das amostras de ATCC 11116. Os íons que se destacaram nestas amostras estão em evidência.

Figura 34. PCA dos dados de ESI(+)-MS das amostras de β -lac: *Scores* – amostras e *loadings* – íons. Não é possível observar a separação das amostras, exceto as de origem de infecção em células Hec-1b (H).

Figura 35. ESI(+)-MS *fingerprints* das amostras de β -lac. Os íons que se destacaram nestas amostras estão em evidência.

Figura 36. HCA dos dados de ESI(-)MS de Hae-BPF 254/86 e 258/86 segundo o método de Ward e distância Euclidiana ao quadrado. Pode-se observar o agrupamento das amostras em duplicata do estoque, da infecção em A549 (amostras A), da infecção em Hec-1b (amostras H), da infecção em HaCaT (amostras R) e da infecção em Sprague-Dawley (amostras S).

Figura 37. HCA dos dados de ESI(-)MS de Hae KC1018 e ATCC 11116 e Hi β -lac segundo o método de Ward e distância Euclidiana ao quadrado. Pode-se observar o agrupamento das amostras em duplicata do estoque, da infecção em A549 (amostras A), da infecção em Hec-1b (amostras H) e da infecção em HaCaT (amostras R) e, no caso de Hi β -lac, também da infecção em Sprague-Dawley (amostras S).

Figura 38. PCA dos dados de ESI(-)MS das amostras de 254/86: *Scores* – amostras e *loadings* – íons. Pode-se observar a separação das amostras.

Figura 39. ESI(-)MS *fingerprints* das amostras de 254/86.

Figura 40. PCA dos dados de ESI(-)MS das amostras de 258/86: *Scores* – amostras e *loadings* – íons.

Figura 41. ESI(-)MS *fingerprints* das amostras de 258/86.

Figura 42. PCA dos dados de ESI(-)MS das amostras de KC1018: *Scores* – amostras e *loadings* – íons.

Figura 43. ESI(-)MS *fingerprints* das amostras de KC1018.

Figura 44. PCA dos dados de ESI(-)MS das amostras de ATCC 11116: *Scores* – amostras e *loadings* – íons. Pode-se observar a separação das amostras de infecção de células Hec-1b (H).

Figura 45. ESI(-)MS *fingerprints* das amostras de ATCC 11116.

Figura 46. PCA dos dados de ESI(-)MS das amostras de β -lac: *Scores* – amostras e *loadings* – íons no modo negativo.

Figura 47. ESI(-)MS *fingerprints* das amostras de β -lac.

LISTA DE TABELAS

Tabela 1. Linhagens de *Haemophilus* utilizadas neste trabalho e suas principais características.

Tabela 2. Parâmetros de ionização para cada lipídio padrão.

Tabela 3. Características das montagens dos *reads* do pirosequenciamento do DNA total de Hae 258/86 pelo *gsAssembler* e *gsMapper*.

Tabela 4. Organização dos *contigs* de Hae 258/86.

Tabela 5. Número total de *reads* e percentual destes com qualidade aceitável gerados pelo sequenciamento por Illumina.

Tabela 6. Dados referente à montagem dos genomas das linhagens de Hae, em *contigs* maiores de 1000 bp.

Tabela 7. Presença ou ausência dos genes *mdaB* e *crgA* em Hae, Hi e N. meningitidis e características deste último.

Tabela 8. Comparação entre os plasmídios de linhagens Hae-BPF.

Tabela 9. Presença (+) ou ausência (-) de genes do *regulon* CRP-S de Hi em linhagens de Hae.

Tabela 10. Relação entre o número de animais dos quais foram recuperadas colônias bacterianas e o número de animais infectados.

Tabela 12. Resumo dos principais resultados obtidos neste trabalho.

Tabela 13. Lista de *primers* utilizados neste trabalho, suas funções e principais características.

Tabela 14. Linhagens bacterianas utilizadas nas buscas e comparações de sequências.

LISTA DE ABREVIATURAS E SIGLAS

A549 – Célula imortalizada de carcinoma epitelial pulmonar humano
ANOVA - *One-way analysis of variance*
alaS – gene codificante de alanil-tRNA sintetase
ATCC – *American Type Culture Collection*
BHI – *Brain heart infusion*
BLAST - *Basic Local Alignment Search Tool*
BLASTn – *Nucleotide BLAST*
BPF – *Brazilian Purpuric Fever / Febre Purpúrica Brasileira*
CaCl₂ - Cloreto de Cálcio
CARS - *Compensatory anti-inflammatory response syndrome*
CDC - *Centers for Disease and Control*
CDS - *Coding Sequence*
CEUA - Comissão de Ética no Uso de Animais da UNICAMP
CG – Conteúdo Citosina Guanina
Chang - *Human conjunctival cells*
CID - Dissociação Induzida por Colisão
CONCEA - Conselho Nacional de Controle de Experimentação Animal
CREN - *Contact regulatory element of Neisseria*
crgA – *Contact-regulated gene A*
CRP-S – *Cyclic AMP Receptor Protein Site-regulated*
Ct - *Threshold cycle*
DMSO – Dimetilsulfóxido
EDTA – Ácido Etilenodiamino Tetra-acético
ELISA – *Ezyme Linked Immuno Sorbent Assay*
ermAM - *Erythromycin-resistance marker*
era - gene codificante da proteína Era de ligação ao GTP
ESI - Ionização por eletrospray.
ESI(-) - ESI no modo negativo
ESI(+) - ESI no modo positivo
ESI-MS - *Electrospray tandem mass spectrometry*
ET - *Eletroforetic type*
GAPDH – Gliceraldeído-3-fosfato desidrogenase
gmk - gene codificante da guanilato quinase
gyrA - gene codificante da subunidade A da DNA girase
H₂SO₄ – Ácido Sulfúrico
hadA – *Haemophilus adhesin A*
Hae – *Haemophilus influenzae* biotipo aegyptius
Hae-BPF – Hae causador da BPF
Hae/ Hi A - Linhagem bacteriana isolada após infecção em células A549
Hae/ Hi H - Linhagem bacteriana isolada após infecção em células Hec-1B
Hae/ Hi R – Linhagem bacteriana isolada após infecção em células HaCaT
Hae/ Hi S - Linhagem bacteriana isolada do sangue de ratos Sprague-Dawley
Hae não-BPF- Linhagens de Hae não associadas a BPF
HCA- *Hierarchical Cluster Analysis* / Análise de agrupamentos hierárquicos
Hec-1B – célula imortalizada de adenocarcinoma de endométrio humano
Hi - *Haemophilus influenzae*
HMEC-1 - *Human Microvascular Cell line*

HUVEC - *Human Umbilical Vein Endothelial Cells*
IAL – Instituto Adolfo Lutz de São Paulo
IL-10 - Interleucina 10
IL-10m - Interleucina 10 murina
INCQS – FIOCRUZ – Instituto Nacional de Controle de Qualidade - Fundação Oswaldo Cruz.
IPTG - *Isopropyl β-D-1-thiogalactopyranoside*
Labiotec – Laboratório de Biotecnologia (IB – UNICAMP)
las – gene codificante do autotransportador associado à virulência Las em Hae
lav- gene codificante do autotransportador associado à virulência Lav em Hi e Nm
LB – meio Luria Bertani
LGE - Laboratório de Genômica e Expressão (IB-UNICAMP)
LPS - Lipopolissacarídeo
LOS – Lipooligossacarídeo
m/z – Razão massa/carga
map - gene codificante da metionina aminopeptidase
MCM - *Mauve Contig Mover*
MDa – Mega Dalton
MLEE – *Multilocus enzyme electrophoresis*
Mo(NH₄)₂ - Molibdato de amônio
MS - *Mass Spectrometry*/ Espectrometria de massa
MS/MS – Espectrometria de massa em tandem
N50 - *Contig* ou *scaffold* no qual se obtém 50 % da montagem total do genoma
NAD - Dinucleotídeo de nicotinamida
NCBI – *National Center for Biotechnology Information*
NTHi – *Nontypable H. influenzae*
primase - gene codificante da DNA primase
PBS – *Phosphate buffered saline* / Tampão fosfato-salino
PCA - *Principal Component Analysis* /Análise do componente principal
PCR – *Polymerase chain reaction* / reação em cadeia da polimerase
qRT-PCR – reação de transcrição reversa seguida de PCR quantitativo em Tempo Real
RAST - *Rapid Annotations using Subsystems Technology*
recA – gene codificante da recombinase A
recF - gene codificante da proteína de replicação e reparo RecF
RefSeq – *Reference Sequence*
rho - gene codificante do fator de terminação de transcrição Rho
RNA – Ácido Ribonucléico
rpoA – gene codificante da subunidade alfa da RNA polimerase
rpoC – gene codificante da subunidade beta da RNA polimerase
rpoD - gene codificante do fator sigma da RNA polimerase
rRNA – RNA ribossômico
SDS – *Sodium dodecyl sulfate*
SDS-PAGE - *Sodium dodecyl sulfate – Polyacrylamide gel electrophoresis*
SFB - Soro Fetal Bovino
SIRS - *Systemic inflammatory response syndrome*
TA – Temperatura de anelamento
TAA – *Trimeric autotransporter adhesins*
tabA – gene codificante da adesina autotransportadora TabA em Hae-BPF
tahA – gene homólogo a *tabA* em Hae não-BPF
TNF-α - *Tumor necrosis factor alpha*/Fator de necrose tumoral alfa

TNF- α m – Fator de necrose tumoral alfa murino

tRNA – RNA transportador

UFC – Unidades Formadoras de Colônia

UPLC-ESI-MS - *Ultra-performance liquid chromatography-electrospray tandem mass spectrometry*

UPMGA - *Unweighted Pair Group Method with Arithmetic Mean*

X-Gal - *5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside*

SUMÁRIO

INTRODUÇÃO.....	22
1. Febre Purpúrica Brasileira.....	22
2. <i>Haemophilus influenzae</i> biotipo aegyptius.....	22
OBJETIVOS.....	35
1. Objetivo principal.....	35
2. Objetivos específicos.....	35
MATERIAIS E MÉTODOS.....	36
1. Meios e condições de crescimento bacteriano.....	36
2. Sequenciamento e montagem do genoma de Hae 258/86.....	37
3. Sequenciamento e montagem dos genomas de Hae Karina, 219/86, 284/86, 321/86, 406/86 e KC1018.....	41
4. Comparação entre genomas.....	44
5. Plasmídios de Hae.....	46
6. Transformação e competência em Hae.....	47
7. Ensaio de infecção.....	56
8. Expressão de genes em Hae.....	61
9. Expressão de citocinas <i>in vitro</i> e <i>in vivo</i>	65
10. Análise dos lipídios de Hae por Espectrometria de massas.....	66
RESULTADOS E DISCUSSÃO.....	72
1. Sequenciamento e Montagem do genoma de Hae 258/86.....	72
2. Sequenciamento e Montagem dos genomas de Hae Karina, 219/86, 284/86, 321/86, 406/86 e KC1018.....	78
3. Comparação entre os genomas.....	79
4. Plasmídio de Hae-BPF.....	87
5. Transformação e competência em Hae.....	90
6. Ensaio de infecção <i>in vivo</i>	94
7. Expressão de genes em Hae.....	95
8. Quantificação de citocinas inflamatórias por qRT-PCR.....	105
9. Análise dos lipídios de Hae por Espectrometria de Massas.....	111

10. Considerações finais	116
CONCLUSÕES	119
REFERÊNCIAS	120
ANEXOS.....	128
1. Lista de <i>primers</i>	128
2. Linhagens bacterianas utilizadas nos alinhamentos de sequências de DNA e proteínas	133
3. Aprovação do CEUA.....	135
4. Dados ESI(+)MS	137
5. Dados ESI (-) MS	147
6. Declaração.....	159
7. <i>Curriculum</i> durante o doutorado	160

INTRODUÇÃO

1. Febre Purpúrica Brasileira

Febre Purpúrica Brasileira (BPF – *Brazilian Purpuric Fever*) é uma doença pediátrica descrita pela primeira vez em 1984 no estado de São Paulo, Brasil. É caracterizada por febre alta, dor abdominal, vômitos, evolução para um quadro hemorrágico nos pulmões e adrenais, colapso vascular, petéquias e necrose dos tecidos periféricos. A doença apresenta taxa de mortalidade média de 40 % (nos primeiros casos descritos, a taxa de mortalidade chegou a 70 %) e seu quadro clínico se desenvolve após 7 a 16 dias de um episódio de conjuntivite purulenta (CDC, 1986, CDC, 1985, Harrison et al., 2008).

Os principais surtos de BPF ocorreram de 1984 a 1990 no Brasil e a ocorrência mais recente de BPF foi relatada em 2007 na região do Amazonas (Santana-Porto et al., 2009). Casos da doença também foram relatados na Austrália (McIntyre et al., 1987) e um caso nos Estados Unidos (Virata et al., 1998).

2. *Haemophilus influenzae* biotipo aegyptius

2.1. Taxonomia de *Haemophilus influenzae* biotipo aegyptius

O agente etiológico da BPF é *Haemophilus influenzae* biotipo aegyptius (Hae), bactéria identificada pela primeira vez há mais de um século por Koch no Egito como causadora de conjuntivite purulenta. Anos depois, foi isolada por Weeks e, finalmente, o bacilo Koch-Weeks foi nomeado como *Haemophilus aegyptius* por Pittman & Davis (1950). Estudos de hibridização de DNA de *H. influenzae* e *H. aegyptius* mostraram que se trata de um grupo de organismos relacionados (Albritton et al., 1984, Casin et al., 1986, Brenner et al., 1988); ainda, Kilian (1976) observou que *H. aegyptius* é fenotipicamente muito próximo a *H. influenzae* biotipo III. Apesar de diferenças, que serão discutidas

adiante, diversos autores, como Casin et al. (1986) e Brenner et al. (1988), argumentam que, dadas as semelhanças fenotípicas, bioquímicas e genéticas, *H. influenzae* e *H. aegyptius* representam a mesma espécie. Assim, *H. aegyptius* é comumente designado como *H. influenzae* biogrupo ou biotipo aegyptius (Hae).

Bactérias do gênero *Haemophilus* são cocobacilos Gram-negativos, sem motilidade, não formadores de esporos e aeróbios ou anaeróbios facultativos. O único hospedeiro natural de *H. influenzae*, incluindo Hae, é a espécie humana e tais organismos necessitam dos fatores de crescimento V (NAD - dinucleotídeo de nicotinamida) e do fator X (hemina) presentes no sangue. Linhagens de *H. influenzae* são classificadas em sorotipos e em biotipos. A sorotipagem de *H. influenzae*, em sorotipos de a - f, baseia-se na presença do tipo capsular. As linhagens que não apresentam cápsula são não tipáveis (NTHi – *nontypable H. influenzae*), como é o caso de Hae. A classificação em biotipos I – VIII baseia-se na presença de enzimas como ornitina descarboxilase, urease, na fermentação do açúcar D-xilose e na produção de indol (Kilian et al., 1979). É interessante ressaltar que alguns autores como Drozd et al. (2012) consideram biotipo e biogrupo como sinônimos e usam a nomenclatura *Haemophilus influenzae* biotipo aegyptius, sendo aegyptius um biotipo a parte dos demais. Bactérias pertencentes ao biotipo III e Hae não são capazes de fermentar a xilose e de produzir indol, ainda, tais bactérias são capazes de hemaglutinação e são os principais causadores de conjuntivite entre as demais linhagens de *H. influenzae*. Entretanto, existem diferenças entre linhagens do biotipo III e Hae, como a susceptibilidade de Hae à troleandomicina, morfologia bacilar distinta, menor taxa de crescimento em ágar chocolate (Mazloum et al., 1982) e o perfil de proteínas da membrana externa (Leaves et al., 1994, Carlone et al., 1985).

Ao considerar as diferentes nomenclaturas para Hae: *H. aegyptius*, *H. influenzae* biogrupo aegyptius e *H. influenzae* biotipo aegyptius, optou-se pelo uso da última opção,

da sigla Hae de forma geral e da sigla Hae-BPF para linhagens causadoras da BPF. Para evitar equívocos, utilizaremos a sigla Hi para as demais linhagens de *H. influenzae* (não Hae).

Como *Haemophilus* é um importante patógeno humano, Hi tipo b é um dos principais responsáveis por infecções em crianças, tais como otites, meningites e pneumonia (Farley et al., 1992, Takala et al., 1990, Kostman et al., 1993). Hae, por sua vez, sempre foi associado apenas a casos de conjuntivite purulenta ou “*pink eye*”, com tendência a epidemias de conjuntivite altamente contagiosas em locais de climas quentes, tais como o Brasil (Harrison et al., 2008). Dessa forma Hae não era relacionado a infecções invasivas até ser identificado como causador da BPF.

2.2. Características de Hae

Apesar de Hae ter sido descrita há mais de um século e a BPF, há aproximadamente 30 anos, pouco se sabe sobre a origem e virulência das linhagens causadoras de BPF. Quais fatores de virulência foram responsáveis pela evolução de uma bactéria causadora de conjuntivite a um patógeno invasivo capaz de invadir e sobreviver no sistema intravascular?

Em um estudo realizado por Brenner et al. (1988), cepas de Hae isoladas de casos de BPF de diferentes sítios corpóreos (raspado de pele, sangue, fluido cerebrospinal, conjuntiva e orofaringe) e de cepas controle de Hi e Hae não causadoras de BPF (isoladas da conjuntiva ou orofaringe) foram analisadas e diferenciadas por ferramentas fenotípicas, bioquímicas, epidemiológicas e genéticas. A tipagem por eletroforese de enzima multilocus (MLEE – *multilocus enzyme electrophoresis*) das amostras revelou que todas as linhagens brasileiras de Hae-BPF e algumas linhagens controle de Hae e Hi pertencem ao mesmo ET (*electroforetic type*), o ET 2, o que indica a natureza clonal das

linhagens causadoras da doença (Brenner et al., 1988). Observou-se, também que as linhagens Hae-BPF apresentam 2 tipos de restrição dos genes ribossomais (Brenner et al., 1988, Irino et al., 1988). Tais características estão relacionadas à epidemiologia e caracterização de Hae-BPF e não diretamente à virulência destas bactérias.

2.3. Diagnóstico

A natureza clonal de Hae permite uma fácil distinção entre outras linhagens de *Haemophilus* por características fenotípicas e genotípicas. Porém, especula-se que casos de BPF foram e ainda sejam omitidos, já que o quadro clínico é muito similar à meningocemia causada por *Neisseria meningitidis*, exceto pela ausência de meningite e o fato de ocorrer apenas em crianças de 2 meses a 10 anos de idade. Recém-nascidos, adolescentes, adultos e idosos parecem ser imunes, o que sugere que a proteção contra BPF seja mediada, em parte pelo menos, por anticorpos adquiridos por passagem transplacentar ou produzidos pelo sistema imune maduro do indivíduo (Harrison et al., 2008).

O diagnóstico laboratorial geralmente é realizado pela identificação de Hae, através da cultura de sangue e material da conjuntiva. Até o momento, não se conhece um diagnóstico molecular para BPF. O diagnóstico molecular de *N. meningitidis*, ou meningococo, por exemplo, é normalmente feito por PCR através da identificação de sequências tidas como específicas para *Neisseria* spp., como o gene *crgA* (*contact-regulated gene A*) (Kesanopoulos et al., 2005, Taha et al., 2005, Parra et al., 2007, Tuyama et al., 2008, Hou et al., 2014). Este último gene codifica um regulador transcripcional da família LysR e já foi detectado em linhagens Hi NTHi através de *primers* usados para diagnóstico clínico de *N. meningitidis* (Hou et al., 2014).

Como a BPF pode evoluir para um quadro exarcebado de sepse com elevada mortalidade, seria interessante o desenvolvimento de testes rápidos e rotineiros capazes

de identificar Hae e distinguir casos de BPF de meningococemia.

2.4. Plasmídio Constitutivo

O grupo de Brenner identificou em linhagens causadoras de BPF um plasmídio de 24 MDa com um padrão de restrição com a enzima *AccI* referido como 3031 (devido ao fato desse padrão ser primeiro observado na linhagem de Hae F3031 ou 254/86) (Brenner et al., 1988). Através do sequenciamento, sabe-se que o plasmídio de F3031 tem 32433 pares de base (pb) e 34 genes em potencial, cujas proteínas hipotéticas apresentam pouca identidade e similaridade com as encontradas no banco de dados. Esse plasmídio apresenta dois domínios distintos: um domínio com genes semelhantes aos encontrados em plasmídios conjugativos e outro com genes com funções desconhecidas e que apresenta nove cópias da sequência *uptake* específica de *Haemophilus*. Observou-se que o plasmídio em questão não é similar o suficiente a qualquer outro plasmídio conjugativo caracterizado para ser enquadrado em algum grupo de incompatibilidade e não foram identificados genes diretamente relacionados à virulência (Kroll et al., 2002). Ainda, foram identificadas raras linhagens de Hae causadoras de BPF que não apresentam essa estrutura (Tondella et al., 1995). Tais fatos indicam que o plasmídio não apresenta um envolvimento crítico na virulência de Hae, apesar de não se conhecer totalmente os mecanismos de virulência desta bactéria. Porém, ao considerar a virulência única das linhagens Hae e a presença de um plasmídio singular na maioria destas linhagens, tornam-se importantes estudos mais aprofundados sobre essa estrutura.

2.5. Cápsula

Como as linhagens de Hae-BPF são invasivas, ou seja, capazes de chegar à corrente sanguínea, acreditava-se que estas apresentavam uma cápsula polissacarídica, um dos mais importantes fatores de virulência em Hi. Brenner et al. (1988) observou que tais bactérias não reagem com anticorpos anti-cápsula de Hi tipo b ou anticorpo

polivalente anti-cápsula de a - f e também não conseguiu detectar alguma estrutura capsular. Anos depois, com o sequenciamento do cromossomo da linhagem F3031 (Strouts et al., 2012), confirmou-se que Hae-BPF não apresenta o *locus cap* presente nas linhagens capsuladas de Hi. Porém, estudos de hibridização de DNA mostraram que Hae-BPF apresenta o elemento de inserção IS1016, o qual está localizado no *loci* de *cap* de algumas linhagens de Hi. IS1016 está ausente em linhagens de Hae não causadoras de BPF (Dobson et al., 1992). Isso sugere que Hae-BPF possa ter possuído ou adquirido em algum momento os genes responsáveis pela síntese capsular (em conjunto com outros genes de virulência talvez). Entretanto, suspeita-se que linhagens de Hae-BPF apresentam uma superfície com carboidratos, evidenciada com a coloração por *ruthenium red*. Porém, esse reagente não é específico para polissacarídeos e tal coloração só é observada após a passagem em modelo animal (Carlone et al., 1989). Assim, é possível que essa superfície apresente composição de outra natureza, seja codificada por genes desconhecidos cuja expressão seja somente induzida em contato com o hospedeiro.

Até o momento não há evidência concreta que Hae-BPF seja um organismo portador de cápsula ou não. Se de fato tal bactéria não for capsulada, isso a destaca entre os patógenos bacterianos invasivos, já que essa estrutura é essencial para a sobrevivência destes patógenos na corrente sanguínea do hospedeiro (Harrison et al., 2008).

2.6. Lipooligossacarídeo e Lipídios

Outro fator de virulência comumente relacionado à doenças invasivas e ao quadro de sepse bacteriana é o lipooligossacarídeo (LOS), também conhecido como endotoxina (Preston et al., 1996). O LOS, ou lipopolissacarídeo (LPS) em algumas bactérias, é conhecido por estimular o sistema imune, e é capaz de induzir uma síndrome inflamatória caracterizada por uma fase de SIRS (*Systemic inflammatory response syndrome*) e uma fase de CARS (*Compensatory anti-inflammatory response syndrome*), com a produção

de citocinas como IL-10 (Interleucina 10) e TNF- α (*Tumor necrosis factor* – fator de necrose tumoral) (Cohen, 2002). Vale ressaltar que a resposta imune causada pela BPF é algo que ainda não foi analisado.

O LOS ou LPS é o maior constituinte da membrana externa de bactérias Gram-negativas e é formado por uma porção lipídica, o lipídio A, covalentemente ligada a um oligossacarídeo ou polissacarídeo. O lipídio A, um glicolípido composto por múltiplos ácidos graxos, é o componente biologicamente ativo e responsável por sua atividade endotóxica (Beutler, 2002). Erwin & Munford (1989) determinaram que os LOS de um isolado de Hae-BPF (F3031), um de Hae não causadora de BPF (F3065) e outro de Hi tipo b são similares em tamanho e composição de ácidos graxos, sendo compostos por dois tipos de ácidos graxos, 3-OH ácido mirístico e ácido mirístico, e se assemelham ao LPS de *Salmonella* spp. e de *Escherichia coli*. Não houve diferenças também entre o LOS de linhagens de Hae causadoras ou não de BPF em ensaios de toxicidade. Também foi observado que o LOS de Hae (BPF e não-BPF) diferiu do de Hi tipo b e *N. meningitidis* ao se separar na fase fenol ao invés da fase aquosa na extração água-fenol. Esse tipo de partição foi descrito em *Pseudomonas aeruginosa* e pode refletir em uma maior hidrofobicidade da endotoxina dessas bactérias.

Em ratos Sprague-Dawley de 5 dias de idade, modelo animal estabelecido para o estudo da patogênese de Hi tipo b (Smith et al., 1973), observou-se mudanças no LOS entre uma linhagem de Hae-BPF antes e depois de ser recuperada do sangue de animais infectados. Analisou-se a mobilidade eletroforética em SDS-PAGE (*Sodium dodecyl sulfate* – *Polyacrylamide gel electrophoresis*) e a capacidade de reagir com anticorpos monoclonais específicos para epítomos de oligossacarídeos do LOS de Hi tipo b. O LOS da amostra sem passagem animal, LOS1, migrou em gel de SDS-PAGE como uma única banda e reagiu com os anticorpos. Por sua vez, linhagens recuperadas do sangue de

animais apresentaram LOS (LOS2) que migrou como duas bandas com massa molecular maior e não reagiu com os anticorpos contra LOS de Hi tipo b. Não obstante, em ensaios com o soro de animais infectados, bactérias com LOS2 tiveram uma taxa de sobrevivência muito maior do que aquelas com LOS1. Tais dados sugerem que o fenótipo LOS2 apresenta uma maior vantagem e sobrevivência no sangue de animais e que, em conjunto com fatores ainda indeterminados, pode estar relacionado ao aumento da virulência (Rubin & St Geme, 1993).

Apesar da detecção da variação da parte glicídica do LOS de Hae-BPF, nenhum estudo foi realizado até o momento sobre a possível variação do lipídio A e, em geral, sobre a caracterização lipídica de linhagens de Hae. Sabe-se que fatores ambientais regulam modificações no lipídio A do LPS de patógenos como *Yersinia pestis* (Rebeil et al., 2004), *P. aeruginosa* (Miller et al., 2005) e *Salmonella typhimurium* e, no caso desta última bactéria, os genes envolvidos nesse processo são essenciais à sua patogênese (Ernst et al., 2001, Guo et al., 1997). Como foram observadas semelhanças do LOS de Hae entre o LPS de *P. aeruginosa* e *Salmonella* spp. (Erwin & Munford, 1989), tornam-se importantes estudos sobre a porção lipídica dessa estrutura.

Além disso, há diversos indícios de que lipídios atuam em diversas etapas na interação hospedeiro-patógeno e que tal interação pode induzir mudanças na composição de lipídios tanto do hospedeiro quanto do patógeno. Os lipídios que compõem as membranas interna e externa de bactéria Gram-negativas não-capsuladas formam uma fina camada entre a bactéria e o ambiente. Assim, alterações no perfil lipídico de patógenos podem atuar como um mecanismo de evasão do sistema imune do hospedeiro (van der Meer-Janssen et al., 2010), no estágio metabólico do microrganismo (Helms, 2006) e, também, na virulência. Por exemplo, o aumento da virulência foi correlacionado a alterações na composição de lipídios em *Mycobacterium tuberculosis* (Khuller et al.,

1982).

A lipidômica de patógenos por espectrometria de massas é uma área promissora e ainda pouco explorada (Helms, 2006). Em relação à composição de lipídios totais de Hae, não foi possível encontrar nenhum estudo na literatura, o que é um fato intrigante já que os lipídios, tanto da membrana quanto do LOS, podem estar envolvidos na capacidade de adaptação de Hae-BPF e na ocorrência de BPF.

2.7. Modelos de Estudo

Em estudos *in vivo* com ratos Sprague-Dawley de 6 dias de idade também foi observado que bactérias Hae-BPF são mais virulentas do que as demais linhagens de Hae e NTHi (Rubin & Carlone, 1989, Rubin & St Geme, 1993), se tornam ainda mais virulentas após a passagem pelo modelo animal e perdem a pilina. Todavia, não se sabe ao certo quais alterações representam de forma mais fidedigna o fenótipo da bactéria no sangue de uma criança infectada (Rubin & St Geme, 1993).

Sabe-se que a pilina de Hae é funcional e estruturalmente semelhante a de Hi tipo b (Weyant et al., 1990). Em ensaios *in vitro* foram recuperadas principalmente Hae desprovidas dessa estrutura e mais virulentas do que as piliadas. Curiosamente, as amostras isoladas diretamente de pacientes com BPF eram todas piliadas. Porém, dados do estudo sugerem que esta estrutura não seja um dos principais fatores vinculados à virulência de Hae causadores de BPF (Rubin & St Geme, 1993).

Algumas diferenças foram observadas entre a infecção por Hae-BPF no modelo animal e em crianças: ratos tiveram uma bacteremia autolimitante e meningite, já em humanos, a bacteremia é frequentemente fatal e não há meningite. É possível que a natureza fulminante da infecção em humanos não ofereça tempo suficiente para bactéria chegar às meninges, já que o maior determinante do desenvolvimento de meningite por

Haemophilus é a magnitude e duração da bacteremia (CDC 1987a, CDC, 1987b). Porém, apesar de não mimetizar completamente a evolução da BPF em humanos, esse modelo *in vivo* é um modelo viável para o estudo da infecção por Hae (Rubin & Carlone, 1989).

Ainda, foi identificada uma proteína de 145 kDa sujeita à variação de fase, P145, em isolados de Hae-BPF após a passagem em animais (Rubin & Rizvi, 1991, Rubin, 1995). Acredita-se que P145 exerça algum papel na sobrevivência de Hae no hospedeiro, já que anticorpos específicos a esta proteína tem atividade bactericida e inibem a bacteremia em ratos (Rubin & Rizvi, 1991).

Além do modelo animal, utilizou-se sistemas *in vitro* para estudar a virulência de Hae. Linhagens causadoras de BPF causam uma maior citotoxicidade em células HMEC-1 (*Human Microvascular Cell line*) que linhagens Hae não relacionadas à doença (Quinn et al., 1994). Os dados obtidos por Quinn et al. (1995) sugerem que Hae-BPF é capaz de se aderir, invadir e multiplicar-se internamente nas células HMEC-1.

2.8. Genoma e Genes Possivelmente Relacionados à Virulência de Hae

A transferência horizontal de genes relacionados à virulência apresenta um importante papel na evolução e surgimento de patógenos bacterianos (Bartlett, 2006, Diep et al., 2006). Assim, uma provável explicação sobre a origem clonal de Hae-BPF é a aquisição de um ou mais genes de virulência pelo fenômeno em questão.

Hi é uma bactéria naturalmente competente ao processo de transformação (Mell et al., 2011, Sinha et al., 2012), ou seja, capaz de incorporar DNA exógeno ao seu próprio genoma (Chen et al., 2005). Sinha et al. (2012) identificaram o *regulon* CRP-S (*Cyclic AMP Receptor Protein Site-regulated*) composto por genes relacionados ao processo de transformação presentes em diferentes linhagens de Hi.

Segada & Lesse (1997) transformaram uma linhagem de Hae-BPF através de

modificações no método de Herriot, para obtenção de células competentes de *Haemophilus* (Herriott et al., 1970). Muitos grupos de pesquisa tentaram repetir o método de Segada & Lesse (1997), porém nenhum obteve sucesso até o momento (Li et al., 2003, Serruto et al., 2009). Apesar de análises *in silico* mostrarem que Hae possui os genes e sítios regulatórios necessários para competência e transformação em Hi (Strouts et al., 2012) e que seu plasmídeo de 24 MDa apresenta sequências *uptake* específicas de *Haemophilus* (Kroll et al., 2002), Hae é conhecida por ser resistente à manipulação genética (Segada & Lesse, 1997, Li et al., 2003, Serruto et al., 2009).

Por outro lado, sequências conservadas de *N. meningitidis* foram encontradas em linhagens de Hae associadas aos surtos de BPF. As sequências, denominadas NMB0419 em *N. meningitidis*, foram denominadas como *bfp001* em Hae (Li et al., 2003). Sabe-se que Hi e *N. meningitidis*, também naturalmente competente, apresentam o mesmo nicho e possuem um sistema de reconhecimento de DNA exógeno similar baseado em sequências *uptake* (Hood et al., 1996, Peak et al., 1999).

O gene *lav* é resultado da troca gênica entre *H. influenzae* e *N. meningitidis*. Acredita-se que tal gene pertença a uma família de autotransportadores AIDA-1/VirG/PerT associados à virulência, o qual foi transferido de *Haemophilus* para *Neisseria*. Através de análises de homologia entre sequências, observou-se que um novo tipo de autotransportador está presente em Hae, o qual foi descrito como *las*. Sugere-se que *las* seja um gene quimérico, cuja origem foi a fusão de dois segmentos distintos de origem exógena ao gênero *Haemophilus* (Davis et al., 2001).

Strouts et al. (2012) sequenciaram o cromossomo de F3031 e através da comparação com os genomas de linhagens de Hi, entre estas Hae não-BPF F3047 (não relacionada ao clone BPF) e Hi RdKW20, encontraram 163 possíveis genes específicos

de Hae. Entre tais, 10 sequências são de novas adesinas autotransportadoras triméricas (TAA – *trimeric autotransporter adhesins*). É interessante ressaltar que o genoma de Hae apresenta muito mais genes de TAA do que qualquer outro genoma de *Haemophilus* já sequenciado.

Uma dessas novas adesinas autotransportadoras é *tabA1* em F3031 (que será referido neste trabalho como *tabA*) e seu homólogo em F3047, *tahA1*. Entretanto, TabA e TahA apresentam diferenças, como por exemplo, seus domínios específicos de TAA são distintos, o que sugere funções diferentes. Este fato destaca o *locus tabA1* na busca por fatores de virulência exclusivos de Hae-BPF. Até o momento, não se sabe o papel desse gene na virulência de Hae-BPF *in vitro* e *in vivo*. Ainda, acredita-se que TabA1 seja a proteína P145 (Strouts et al., 2012), identificada em isolados de Hae de animais (Rubin & Rizvi, 1991, Rubin, 1995).

Outro TAA encontrado em F3031 foi HadA, cuja família de proteínas inclui YadA e NadA identificadas em linhagens virulentas de *N. meningitidis* e *Yersinia enterocolitica*, respectivamente. Através da expressão heteróloga de HadA em *E. coli* não patogênica e de ensaios com células Chang (*human conjunctival cells*) e HUVEC (*human umbilical vein endothelial cells*), observou-se que HadA está envolvido na adesão e invasão celular. O *locus hadA* tem 2159 pb e possui o gene *hadA* com 771 pb. De acordo com Serruto et al. (2009), Hi RdKW20 e Hae não-BPF F1947 não possuem *hadA*, mas apresentam as sequências que flanqueiam esse gene, sendo *hadA* apenas encontrado em F3031 e específico de Hae-BPF. Devido ao fato da composição de GC de *hadA* ser significativamente diferente do restante dos genomas de Hi (Serruto et al., 2009) e da própria F3031 (Strouts et al., 2012), acredita-se que o gene em questão tenha sido adquirido por transferência horizontal (Serruto et al., 2009).

Apesar de avanços no conhecimento genômico de Hae, dados do nosso grupo

sugerem pequenas diferenças entre os genomas das linhagens de Hae, assim é necessária a disponibilidade de genomas de outras amostras de Hae-BPF e não-BPF para uma predição mais precisa de genes exclusivos relacionados à sua virulência. Ainda, o crescente número de genomas bacterianos publicados promove comparações mais precisas entre linhagens relacionadas que variam em seu potencial patogênico, as quais podem revelar genes específicos envolvidos na virulência de certas linhagens (Strouts et al., 2012). Como BPF é uma doença relativamente recente e até negligenciada, mais estudos são necessários a fim de se compreender o desenvolvimento do quadro clínico, a origem e virulência única de Hae-BPF.

OBJETIVOS

1. Objetivo principal

O presente estudo tem como objetivo principal o avanço no conhecimento sobre Hae-BPF e a análise de possíveis características relacionadas à virulência dessa bactéria. Para tal, foi realizado um quadro comparativo entre linhagens de Hae causadoras, Hae-BPF Karina, 219/86, 254/86, 258/86, 284/86, 321/86 e 406/86, linhagens de Hae não causadoras de BPF, Hae KC1018 e ATCC 11116, e linhagens de Hi, RdKW20 e β -lac.

2. Objetivos específicos

- Sequenciar, montar, analisar e comparar o genoma das linhagens Karina, 219/86, 258/86, 284/86, 321/86, 406/86 e KC1018;

- Determinar por PCR em Tempo Real a expressão dos genes candidatos à virulência, *las*, *hadA* e *tabA* nas linhagens 254/86, 258/86, KC1018 e ATCC 11116 isoladas de infecção em células humanas A549, Hec-1B e HaCaT, e em ratos Sprague-Dawley;

- Determinar por PCR em Tempo Real e comparar a expressão das citocinas IL-10 e TNF- α em células humanas A549 e HaCaT, e em ratos Sprague-Dawley após infecção por Hae 254/86, 258/86, KC1018, ATCC 11116 e β -lac;

- Determinar por espectrometria de massas o perfil de lipídios das linhagens 254/86, 258/86, KC1018, ATCC 11116 e β -lac isoladas de infecção em células humanas A549, Hec-1B e HaCaT, e em ratos Sprague-Dawley.

MATERIAIS E MÉTODOS

1. Meios e condições de crescimento bacteriano

As linhagens de Hae foram adquiridas da Coleção de Cultura do Instituto Adolfo Lutz (IAL) de São Paulo. Já as linhagens de Hi foram gentilmente cedidas pelo Instituto Nacional de Controle de Qualidade da Fundação Oswaldo Cruz (INCQS – FIOCRUZ) de sua coleção de cultura. As linhagens estão descritas na Tabela 1.

O crescimento das bactérias foi realizado através do inóculo das mesmas em meio BHI ágar suplementado com NAD (Sigma-Aldrich®) a 10 µg/ mL e hemina (Sigma-Aldrich®) a 10 µg/ mL.

Para os ensaios de transformação, utilizou-se BHI líquido, suplementado com NAD (Sigma-Aldrich®) e hemina (Sigma-Aldrich®) como especificado acima (Brenner et al., 1988). Os meios foram suplementados com antibiótico quando necessário: obteve-se a concentração final de 7 µg/ mL para eritromicina em Ágar Chocolate (BHI suplementado com 25 % de sangue de cavalo desfibrinado) para os ensaios de transformação e infecções em modelo *in vivo*.

Tabela 1. Linhagens de *Haemophilus* utilizadas neste trabalho e suas principais características.

Linhagem	Sigla do CDC	Características	Região de Isolamento
Karina	F1946	Hae-BPF, isolado da pele	Promissão- SP
219/86	F3030	Hae-BPF, isolado do sangue	Serrana -SP
254/86	F3031	Hae-BPF, isolado do sangue	Serrana -SP
258/86	F3033	Hae-BPF, isolado do sangue	Serrana -SP
284/86	F3042	Hae-BPF, isolado da orofaringe	Serrana -SP
321/86	F3039	Hae-BPF, isolado do sangue	Serrana -SP
406/86	F3283	Hae-BPF, isolado do sangue	Presidente Prudente -SP
KC1018	-	Hae isolado de um caso de conjuntivite	U.S.A.
ATCC 11116	F3065 -R1967	Hae referência biotipo <i>aegyptius</i> NCTC 8502	Texas, U.S.A.
Rd KW20	-	Hi sorotipo d, cepa padrão de laboratório, ATCC 51907	-
B-lac	-	Hi sorotipo b	-

2. Sequenciamento e montagem do genoma de Hae 258/86

As montagens e algumas análises de todos os genomas de Hae apresentadas neste trabalho foram feitas com a colaboração dos pesquisadores do Laboratório de Genômica e Expressão (LGE – IB, UNICAMP) Marcelo Falsarella Carazzolle e Luciana Souto Mofatto.

2.1. Pirosequenciamento do genoma de Hae 258/86

A linhagem Hae 258/86 foi inoculada em meio BHI suplementado com NAD e hemina (como especificado anteriormente), incubada por 20 horas a 37 °C em atmosfera de 5 % de CO₂ e teve seu DNA genômico (cromossômico e plasmidial) extraído segundo Ausubel & Wiley (2001). O material extraído foi lido em NanoDrop (NanoDrop® 2000 - Thermo Scientific®), a fim de se verificar a pureza e concentração do mesmo e então encaminhado à empresa Helixxa, a qual realizou o pirosequenciamento do DNA total da

amostra no sistema *Genome Sequencer FLX Titanium* (Roche 454[®]).

A partir do serviço de pirosequenciamento obteve-se dois arquivos de extensão *.sff com 187389 e 70451 *reads* e 67679284 e 34082492 bases cada um respectivamente, totalizando em 257840 *reads* e 101761776 bases.

2.2. Montagem do genoma de Hae 258/86

Para que um genoma seja montado, os *reads*, que são sequências curtas (fragmentos) resultantes do sequenciamento, são agrupados para formar uma sequência maior e contígua denominada *contig*. Há dois tipos de montagem: por referência e *de novo*.

A montagem por referência concatena os *reads* em *contigs* com base em um genoma de referência já montado e completo. Já a *de novo* consegue utilizar apenas os *reads* para formar os *contigs* sem utilizar um genoma de referência. A metodologia de montagem *de novo* é comumente utilizada quando não existe nenhum genoma de referência ou quando este genoma está altamente fragmentado.

Os arquivos *.sff obtidos do pirosequenciamento do DNA total de Hae 258/86 foram submetidos à montagem nos *softwares* *gsAssembler* (montagem *de novo*) e *gsMapper* (montagem ancorada por um genoma referência) do pacote Newbler v. 2.8, fornecido pela Roche 454[®].

Foram realizadas três diferentes montagens a fim de se obter a mais eficiente. Duas destas foram realizadas no *gsAssembler*, no qual o primeiro *assembly* foi feito com os 2 arquivos *.sff e o segundo, com apenas o maior arquivo *.sff. A terceira montagem foi realizada no *gsMapper*, na qual utilizou-se os dois arquivos *.sff do sequenciamento e um arquivo *.fasta com o genoma de Hae 254/86 como referência disponível no NCBI (*National Center for Biotechnology Information* - <http://www.ncbi.nlm.nih.gov/>; acesso GenBank n°. FQ670178) (Strouts et al., 2012).

Em todas as montagens, utilizou-se o *default* do programa: valor de sobreposição mínimo de 40 bases e identidade de sobreposição de 90 % entre bases de cada *read*.

A qualidade de cada montagem foi analisada através dos seguintes parâmetros: número de *contigs*, tamanho do maior *contig*, N_{50} (*contig* ou *scaffold* no qual se obtém 50 % da montagem total, por exemplo, supondo-se que o N_{50} seja o valor do *contig* 3, então 50 % da montagem será o *contig* 1 + 2 + 3) e tamanho da montagem. A montagem com o menor o número de *contigs*, maior o tamanho do maior *contig*, menor o N_{50} e maior o tamanho foi considerada como a mais eficiente.

2.3. Concatenação dos *contigs* de Hae 258/86 por referência

A montagem de Hae 258/86 considerada a mais eficiente foi alinhada contra o genoma completo de Hae 254/86 pelo programa Mauve. O arquivo *.fna obtido do *assembly* contendo o *draft* genômico, com apenas os *contigs* formados com mais de 500 bases, foi alinhado com o *.gbk contendo a sequência e anotação do genoma de Hae 254/86 pela ferramenta *progressiveMauve* (Darling et al., 2010).

Após o alinhamento dos genomas, o programa oferece uma ferramenta, *Mauve Contig Mover* (MCM), a qual exhibe a ordem relativa dos *contigs* do *draft* baseando-se no genoma referência. Esse ordenamento e orientação dos *contigs* são feitos através de regiões homólogas entre o genoma referência e as extremidades de cada *contig*. O MCM gera uma série de novos alinhamentos, cada um representando uma reorganização diferente dos *contigs*. Para cada nova organização, são criados um arquivo *.fasta e uma lista dos *contigs* ordenados e suas localizações (Rissman et al., 2009).

Para se concatenar os possíveis *contigs* adjacentes, foi utilizada inicialmente uma abordagem *in silico* com os dados (*reads*) do próprio sequenciamento seguida de uma *in vitro* por PCR.

Inicialmente, um banco de dados com todos os *reads* resultantes do

sequenciamento de 258/86 foi criado. Em seguida, com base na ordem de *contigs* estabelecida pelo MCM, obteve-se arquivos *.fasta com as 300 bases iniciais ou finais das extremidades entre os possíveis *contigs* adjacentes, como está ilustrado na Figura 1.

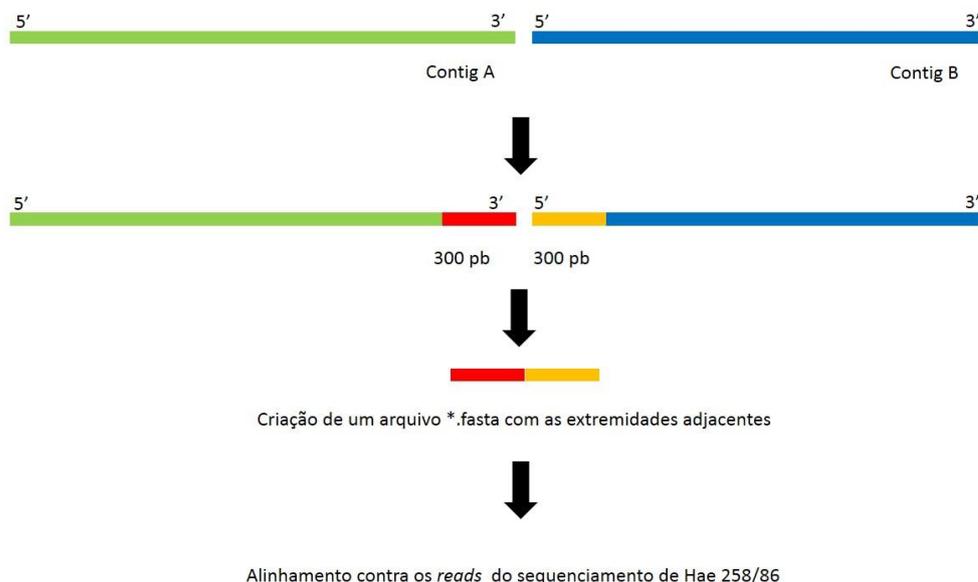


Figura 1. Metodologia inicial para concatenar *contigs* de Hae 258/86.

Estes novos arquivos então foram alinhados contra o banco de dados de *reads*. Caso houvessem *reads* entre as regiões de junção entre *contigs*, esta junção era tida como real e os *contigs*, concatenados. Caso houvessem poucos *reads* ou nenhum indicando a existência de uma região central de união entre os dois *contigs*, utilizou-se uma abordagem *in vitro* para a junção entre tais *contigs*.

O passo seguinte foi testar se as junções entre os *contigs* eram reais ou não através de PCR. Para tal, foram desenhados *primers* entre as extremidades dos *contigs* possivelmente adjacentes. Estes estão em destaque na Tabela 12 em Anexos. Para cada reação de PCR, obteve-se um volume final de 50 μ L de mistura para cada reação com 5 μ L de DNA *template*, 0,5 μ M de cada oligonucleotídeo (Sigma-Aldrich[®]), 2,5 U de *Taq* DNA Polimerase (Invitrogen[®]), 20 mM de Tris-HCL (Invitrogen[®]), 2 mM de MgCl₂, 200 μ M de cada desoxinucleotídeo trifosfatado (Promega[®]) e 2 % de DMSO

(Dimetilsulfóxido - Sigma-Aldrich[®]), conforme as especificações do fabricante da enzima *Taq* DNA Polimerase Recombinante (Invitrogen[®]). A reação de amplificação ocorreu em termociclador (Gene Amp PCR System 9700[®]) com um ciclo inicial de denaturação de 95 °C por 5 minutos, seguido de 35 ciclos com denaturação a 92 °C por 1 minuto, anelamento por 1 minuto (a temperatura de anelamento para cada par de *primers* está na Tabela 12 em Anexos) e extensão a 72 °C por 1 minuto. Um ciclo final de polimerização foi realizado a 72 °C por 10 minutos.

A visualização dos fragmentos amplificados foi realizada em gel de agarose 1,5 %, no qual as amostras foram aplicadas na seguinte proporção: 3 µL de tampão de corrida: 5 µL da amostra, tendo-se como padrão de peso molecular o 1 kb DNA-ladder (Promega[®]). Após o término da corrida de eletroforese, o gel foi corado em banho de Brometo de Etídio a 50µg/mL por 15 minutos e visualizou-se as bandas no gel em transluminador de luz ultravioleta e fotografadas pelo sistema GelDoc It – Imaging System, com filtro vermelho utilizando-se o *software* VisionWorks[®]LS.

2.4. Anotação automática do genoma de Hae /86

Os *contigs* com mais de 500 bases obtidos na montagem 1 foram submetidos à anotação automática através da plataforma RAST (*Rapid Annotations using Subsystems Technology*) v. 4.0, disponível em <http://rast.nmpdr.org/> (Aziz et al., 2008).

3. Sequenciamento e montagem dos genomas de Hae Karina, 219/86, 284/86, 321/86, 406/86 e KC1018

Com o intuito de facilitar e enriquecer as comparações entre os genomas de Hae e Hi, a elucidação de fatores únicos de Hae e de diferenças entre as linhagens deste grupo, foram adicionados a este trabalho o sequenciamento e montagem das pseudomoléculas do DNA cromossômico de Hae Karina, 219/86, 284/86, 321/86, 406/86 e KC1018.

3.1. Sequenciamento dos genomas

O DNA genômico das linhagens de Hae Karina, 219/86, 284/86, 321/86, 406/86 e KC1018 foi extraído seguindo a mesma metodologia descrita para o sequenciamento de Hae 258/86.

O DNA extraído foi lido em NanoDrop (NanoDrop[®] 2000 - Thermo Scientific[®]), a fim de se verificar a pureza e concentração do mesmo e, então, sequenciado pela tecnologia da plataforma Illumina HiSeq2000. O sequenciamento do genoma utilizou um sistema de kits de Illumina para preparação de bibliotecas de DNA com fragmentos *paired-end* (< 1 kb) com *reads* de até 2 x 100 bp, com uma distância de inserto de até 400 bp. Fragmentos do tipo *paired-end* têm ambas extremidades sequenciadas e, como a distância entre estes fragmentos é conhecida, algoritmos de alinhamento podem usar esta informação para mapear de maneira mais precisa *reads* em regiões repetitivas de genomas.

3.2. Montagem dos genomas

A montagem *de novo* foi o método escolhido e o *software* utilizado para isso foi o “*VelvetOptimiser*” versão 2.0 (Gladman, 2009 - *VelvetOptimiser*. US Patent, v. 2, n. 5).

As sequências de *reads* de cada bactéria (R1 e R2) foram unidas separadamente e depois concatenadas através do script em *Perl* chamado “*shuffleSequences_fastq.pl*”, proveniente do pacote “*Velvet*” versão 1.2.03 (Zerbino & Birney, 2008).

Foi feito um corte de número de linhas sobre o arquivo que contém as sequências concatenadas obedecendo a cobertura genômica de 200 x. Cobertura genômica corresponde ao total de pb em *reads* dividido pelo tamanho do genoma. Na prática, corresponde a quantas vezes cada base do genoma foi sequenciada. O cálculo de cobertura foi feito da seguinte maneira:

$$[(\text{soma_linhas}/4) * \text{tamanho_read}] / \text{tamanho_genoma} = 200$$

Onde: soma_linhas = Soma do número de linhas do arquivo concatenado; tamanho_read = Tamanho do read em pb; tamanho_genoma = Tamanho do genoma em bp.

Depois da obtenção desse arquivo com corte, foi rodado o *software VelvetOptimiser 2.0* para fazer a montagem do genoma, passando como parâmetro o *kmer* de corte. O *software* utiliza os chamados grafos de De Bruijn, os quais envolvem a construção de um grafo em que para cada sequência de tamanho $k-1$ é criado um vértice. Um vértice A se conecta a outro vértice B através de uma aresta dirigida se existe uma sobreposição de tamanho $k-2$ do sufixo do vértice A com o prefixo do vértice B. As arestas do grafo representam todos os *kmers* (fragmentos de tamanho k). A sequência é reconstruída através de caminhos eulerianos, caminhos que passam por cada aresta (fragmentos *kmer*) uma única vez (Zerbino & Birney, 2008).

Por fim, é gerado um arquivo *fasta* contendo a montagem final. O resumo do *pipeline* usado neste trabalho e os valores de *kmer* utilizados são mostrados a seguir.

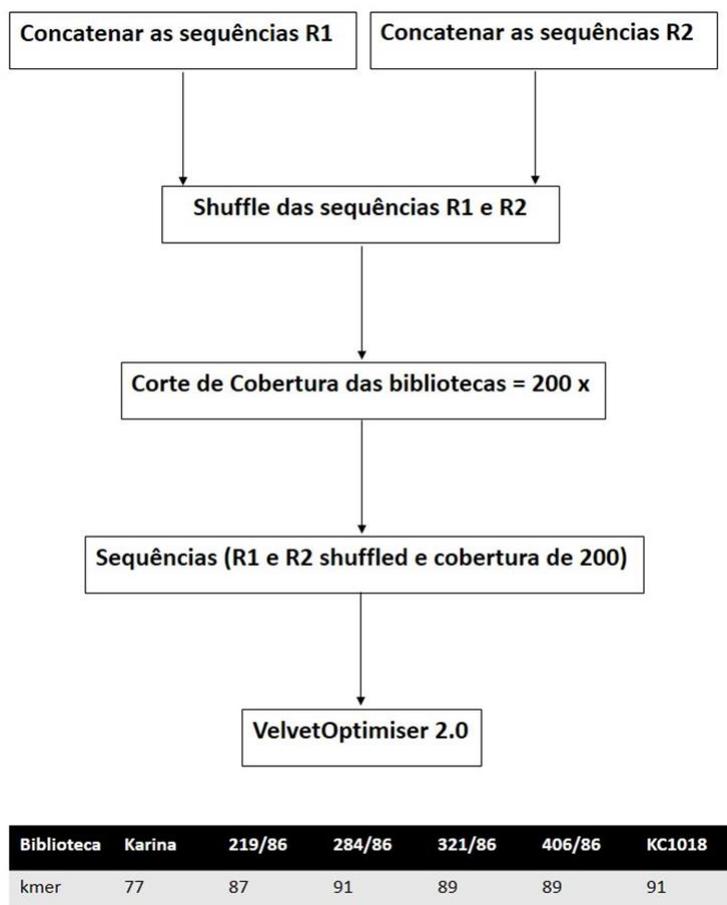


Figura 2. *Pipeline* utilizado na montagem dos genomas de Hae Karina, 219/86, 284/86, 321/86, 406/86 e KC1018.

4. Comparação entre genomas

Além de uma comparação pan-genômica, os genomas sequenciados serviram de base para outras análises mais detalhadas que serão apresentadas no decorrer deste trabalho. Isso foi feito com o objetivo de encontrar possíveis diferenças entre linhagens de Hae, como genes presentes em algumas linhagens e ausentes em outras. Também foram realizadas buscas comparativas por determinados genes de Hi e *N. meningitidis*.

4.1. Comparação pan-genômica entre Hae e Hi

Aos pares, os genomas de linhagens de Hae e Hi descritas na Tabela 13 (Anexos) foram comparados quanto ao alinhamento de bases na plataforma MAUVE (Darling et al., 2010). Para cada alinhamento diferente, obteve-se uma matriz de distância entre os

genomas a qual foi utilizada como base para a criação de um *heat map* mostrando o agrupamento das linhagens analisadas.

4.2. Comparação entre os genomas de Hae 258/86 e outras linhagens Hae-BPF

Os *contigs* de Hae 258/86 foram alinhados contra o genoma de Hae 254/86 na plataforma MAUVE com a ferramenta MCM para exibir a ordem relativa dos *contigs* de 258/86 (Darling et al., 2010, Rissman et al., 2009), como foi abordado no tópico 2.3. Tal alinhamento entre os genomas permitiu a visualização de eventos de escala evolucionária como rearranjos, inversões e até deleções/inserções entre os dois genomas comparados (Darling et al., 2004, Darling et al., 2010).

A partir dos *contigs* que não apresentaram similaridade com o genoma da linhagem 254/86, realizou-se uma busca por similaridade contra outras linhagens da espécie *Haemophilus* através da plataforma BLASTn (*Nucleotide Basic Local Alignment Search Tool*) com o algoritmo *megablast* (Altschul et al., 1990).

4.3. Busca por genes de diagnóstico de *N. meningitidis* em Hae

Sabe-se que os quadros clínicos da BPF e da meningocemia por *N. meningitidis* são muito semelhantes (Harrison et al., 2008) e que isolados clínicos de NTHi apresentam a sequência de *crgA*, um gene usado na identificação de linhagens patogênicas de *N. meningitidis* (Hou et al., 2014). Assim, realizou-se uma busca nos genomas de Hae dos genes de diagnóstico de *N. meningitidis* IS1106, *ctrA*, *porA* (Taha et al., 2005) e *crgA*.

Também foi feita a busca e análise do gene *mdaB* (*modulator of drug activity gene*), adjacente ao gene *crgA*, ativado pelo mesmo e possivelmente envolvido na resposta do meningococo contra o estresse oxidativo (Ieva et al., 2005, Hong et al., 2008).

4.3.1. Identificação de genes de diagnóstico de *N. meningitidis* em Hae por alinhamento de sequências

As sequências dos genes de diagnóstico de *N. meningitidis* citados acima foram

obtidas da linhagem MC58 disponível no NCBI e alinhadas contra os genomas de bactérias listadas na Tabela 13 (Anexos). A busca nos genomas disponíveis no NCBI foi realizada através da plataforma BLASTn com o algoritmo *megablast* (Altschul et al., 1990). Já a busca nos genomas sequenciados neste estudo foi feita também com BLASTn porém em um banco de dados construído com os *contigs* desses genomas.

4.3.2. Comparação das sequências de *crgA* de Hae e *N. meningitidis*

O conteúdo de GC das sequências de *crgA* e de seus respectivos genomas foram comparados. O GC % em *crgA* e dos genomas sequenciados neste estudo foram calculados no *software* Endmemo (disponível em <http://www.endmemo.com/bio/gc.php>). O percentual de GC dos genomas completos foi obtido no NCBI.

As sequências de *crgA* foram alinhadas no *software* MEGA 6.06 pelo algoritmo ClustalW com parâmetros padrão e obteve-se *clusters* pelo método UPMGA (*Unweighted Pair Group Method with Arithmetic Mean*). A análise envolveu as linhagens sequenciadas no presente trabalho e linhagens disponíveis no NCBI que apresentaram o gene em questão, incluindo Hi e principalmente *N. meningitidis* (Tabela 13 em Anexos).

5. Plasmídios de Hae

5.1. Montagem dos plasmídios de Hae

A partir das montagens dos genomas das linhagens Karina, 219/86, 258/86, 284/86, 321/86 e 406/86, foi realizada uma busca por similaridade utilizando a plataforma BLAST com o algoritmo *megablast* (Altschul et al., 1990) contra o plasmídio 254/86, disponível no GenBank, acesso NC_004846.1 (Kroll et al., 2002).

5.2. Fechamento dos plasmídios de Hae 258/86

Para se confirmar a circularidade do *contig* correspondente ao plasmídio de Hae, optou-se por uma abordagem *in vitro*: a construção de *primers* unindo as duas

extremidades do *contig* 19 de Hae 258/86 e a amplificação das mesmas, de forma a unir as duas extremidades por PCR, comprovando, assim, a circularidade da sequência. Os *primers* utilizados estão na Tabela 12 em Anexos (*plasm19 forward* e *reverse*). As reações de PCR e a visualização dos fragmentos seguiram a metodologia já descrita no tópico 2.3 desta seção.

5.3. Comparação *in silico* entre os plasmídios de Hae

Uma análise *in silico* comparativa do perfil de digestão por enzimas de restrição foi feita entre o plasmídio de Hae 254/86 e das demais linhagens de Hae-BPF utilizadas neste trabalho descritas na Tabela 1. Para isso, utilizou-se o *software* pDRAW v.1.1.115 (disponível em <http://www.acaclone.com>).

6. Transformação e competência em Hae

Inicialmente, um dos objetivos principais deste trabalho era inativar os genes *las*, *hadA* e *tabA* nas linhagens de Hae-BPF Karina, 254/86 e 258/86 para analisar a importância destes genes na infecção *in vitro* e *in vivo*. Para tal, foram construídos vetores para o *knockout* desses genes para a mutação sítio-dirigida, os quais foram utilizados posteriormente em ensaios de transformação bacteriana nas linhagens em questão.

Ainda, com base nos dados obtidos, foram realizados ensaios de transformação adicionais a fim de se comparar a capacidade e taxa de transformação entre Hae e Hi. Foi realizada, também, uma busca por genes relacionados à competência em Hi nos genomas sequenciados de Hae.

6.1. Clonagem para *knockout* de genes candidatos à virulência de Hae

A fim de se estudar o papel dos genes *las*, *hadA* e *tabA* na virulência de *Hae*, os mesmos foram amplificados da linhagem 254/86 e clonados em *E. coli* JM 109. A metodologia para a construção do vetor de *knockout* para os genes em questão está descrita a seguir.

6.1.1. Amplificação dos genes *las*, *hadA* e *tabA*

Regiões específicas dos genes delimitadas na Figura 3 foram amplificadas para posterior inativação gênica por recombinação homóloga. As reações de PCR seguiram a metodologia detalhada no tópico 2.3 de Materiais e Métodos e os *primers* utilizados estão descritos na Tabela 12 em Anexos.

6.1.2. Ligação dos *amplicons* ao Sistema pGEMT-Easy

Os fragmentos amplificados de Hae 254/86 com os iniciadores específicos foram re-precipitados em etanol absoluto e submetidos à ligação com o kit pGEMT Easy System II (Invitrogen®). Para tal, adicionou-se 1 µL de pGEMT Easy (Invitrogen®), 1 µL de T4 DNA Ligase (Invitrogen®), 5 µL do Tampão de Ligação 2X (Invitrogen®) e 3 µL do fragmento purificado. A mistura foi incubada a 16 °C por um período de 16 - 18 horas e obteve-se os vetores *plas*, *phadA* e *ptabA*, os quais foram usados em transformação de células competentes de *E. coli* linhagem JM109.

6.1.3. Transformação e clonagem em células competentes de *Escherichia coli*

6.1.3.1. Preparo de células competentes de JM109

A linhagem de *E. coli* JM109 foi inoculada em meio LB ágar em estrias de isolamento e incubada a 37 °C por um período de 16 - 18 horas. Colônias bacterianas foram transferidas em 100 mL meio LB líquido ($DO_{600nm} = 1,00$) e incubadas a 37 °C *overnight*. O crescimento bacteriano foi transferido para um tubo de centrifuga previamente conservado a 4 °C e centrifugado a 4000 g por 10 minutos a 4 °C. O sobrenadante foi descartado e o sedimento ressuscitado em 10 mL de uma solução estéril de $CaCl_2$, 100 mM a 4 °C. Tal suspensão foi novamente centrifugada à mesma velocidade de microcentrífuga supracitada e ao sedimento recuperado adicionou-se 8 mL de uma solução estéril de cloreto de cálcio a 100 mM com 30 % de glicerol. A suspensão foi alíquotada em tubos *ependorf* (alíquotas de 200 µL) e congelada em nitrogênio

líquido e, posteriormente, conservada a -70 °C.

6.1.3.2. Ensaio de transformação em JM109

Uma alíquota de 200 µL de células competentes foram descongeladas rapidamente e adicionadas a 15 µL de plas, phadA e ptabA. A mistura foi mantida em banho de gelo por no mínimo 30 minutos. Transcorrido esse período, a mistura foi submetida a choque térmico: foi aquecida a 42 °C durante o tempo limite de 90 segundos e, logo após, o tubo de reação foi imediatamente mantido em banho de gelo por 2 minutos. Foram adicionados à mistura 800 µL de meio LB líquido pré-aquecido a 37 °C e a mistura foi incubada a mesma temperatura por 3 horas. O volume total de transformação foi inoculado em meio LB ágar adicionado de ampicilina a 100 µg/ mL, já que o sistema pGEMT-Easy (Invitrogen®) apresenta um cassete de resistência a este antibiótico, e adicionado 30 µL X-Gal (a 25 mg/mL) e 10 µL IPTG (*Isopropyl β-D-1-thiogalactopyranoside* – a 1 M), para permitir a seleção das colônias transformantes. As placas foram incubadas por 16 - 48 horas a 37 °C e as colônias selecionadas, repicadas em meio com ampicilina a 100 µg/ mL e estocadas em LB líquido com 30 % de glicerol à temperatura de – 80 °C.

A seleção das colônias transformantes, aquelas com a construção plasmidial, foi realizada pela resistência à ampicilina e coloração da colônia. Pois, além de apresentar um cassete de resistência à ampicilina, o vetor pGEMT-Easy (Invitrogen®) possui o gene *lacZ*, o qual codifica a enzima β –D-galctosidase, responsável pelo fenótipo de cor azul em colônias bacterianas na presença de IPTG e X-Gal. Porém, este vetor é comercializado na forma linearizada com a abertura na região do *lacZ*. Logo, colônias portadoras da construção plasmidial, ou seja, o vetor com o inserto, apresentam o gene *lacZ* inativo e apresentam o fenótipo de coloração branca, o que permite a seleção das colônias de interesse.

6.1.3.3. Extração plasmidial em pequena escala (mini-prep)

Utilizou-se a metodologia descrita por Birnboim & Doly (1979). As amostras de *E. coli* JM109 transformadas foram inoculadas em 5 mL de meio LB líquido suplementado com o(s) antibiótico(s) necessário(s). As culturas bacterianas foram incubadas a 37 °C por 12 - 18 horas. Uma alíquota de 1,5 mL do crescimento bacteriano foi transferida para um tubo do tipo *ependorf* e centrifugada a 12000 g por 2 minutos. O sobrenadante foi desprezado e o sedimento ressuspendido em 100 µL de Solução I gelada (50mM de glicose, 25mM de Tris-HCl pH 8,0 e 10 mM de EDTA pH 8,0) e, em seguida, mantido em banho de gelo por 30 minutos. A seguir, 200 µL de Solução II (NaOH a 0,2N e 1 % de SDS) foram adicionados e o conteúdo do tubo misturado por inversão suave e vagarosa. Após o tubo ter sido mantido em banho de gelo por não mais que 5 minutos, adicionou-se 150 µL da Solução III (60 mL de Acetato de Potássio a 5M, 11,5 mL de Ácido Acético Glacial e 28,5 mL de água MiliQ). Então, o tubo foi invertido suavemente e mantido em banho de gelo por mais 15 minutos. O tubo foi centrifugado a 14000 g por 5 minutos e 400 µL do sobrenadante transferido para outro tubo de centrífuga com a posterior adição de 1 mL de etanol 100 % a -20 °C. Após este procedimento, o tubo foi mantido a -20 °C por pelo menos 1 hora. O DNA plasmidial foi precipitado por centrifugação a 12000 g por 5 minutos, ressuspendido em 50 µL de água MiliQ esterilizada e mantido a -20 °C.

Para a verificação da eficácia da extração, as amostras plasmidiais foram aplicadas em gel de agarose a 1,5 % na seguinte proporção: 3 µL de tampão de corrida: 5 µL da amostra, tendo-se como padrão de peso molecular o 1 kb DNA-ladder (Promega®) Após o término da corrida, o gel foi corado em solução de Brometo de Etídio a 50µg/mL por 15 minutos e visualizou-se as bandas no gel em transluminador de luz ultravioleta e fotografadas pelo sistema GelDoc It – Imaging System, com filtro vermelho utilizando-

se o *software* VisionWorks[®]LS.

6.1.4. Inserção de cassete de resistência para inativação gênica

As construções plasmidiais *plas* (pGEMT easy + *las*Δ), *phadA* (pGEMT easy + *hadA*Δ) e *ptabA* (pGEMT easy + *tabA*Δ), obtidos através da clonagem em *E.coli* linhagem JM109, com os fragmentos dos genes a serem inativados, foram digeridos com endonucleases de restrição em sítios específicos nos quais, posteriormente, inseriu-se o cassete de restrição. Em *plas*, utilizou-se a enzima HindIII, pois sabia-se que seu único sítio de restrição presente no plasmídio era no interior do *amplicon* em questão. Já no caso de *phadA* e *ptabA* utilizou-se as enzimas ClaI e XhoI, respectivamente, pelo mesmo motivo, como é possível analisar na Figura 3.

A reação de digestão enzimática foi realizada com 10 μL do plasmídio, 2 μL da enzima de restrição, 2 μL do tampão de restrição e 6 μL de Água MiliQ estéril. A reação permaneceu *overnight* a temperatura de 37 °C e, no dia seguinte, a enzima foi inativada por aquecimento a uma temperatura de 60 °C por 20 minutos.

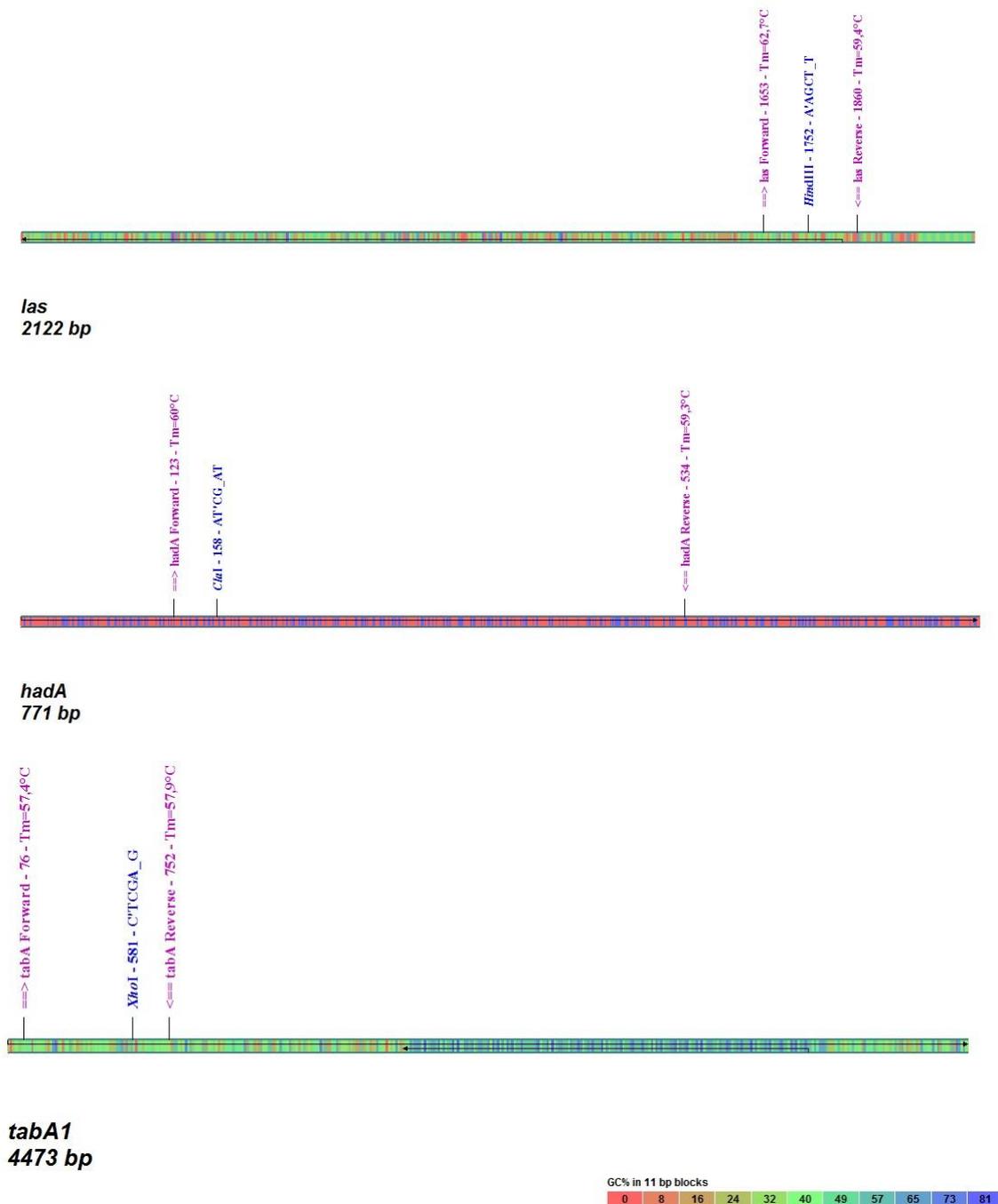


Figura 3. Esquemas dos genes *las*, *hadA* e *tabA* de Hae 254/86 feito no programa pDRAW com as regiões de anelamento dos *primers* específicos. As sequências foram obtidas do Genbank – número de acesso - FQ670178. Os fragmentos amplificados apresentam um sítio de corte para uma determinada enzima de restrições a qual será inserido um cassete de resistência à eritromicina.

O cassete *ermAM* (*erythromycin-resistance marker*) que confere resistência à eritromicina foi amplificado (segundo o protocolo descrito em 2.3) pela utilização de iniciadores específicos (Tabela 12 em Anexos) capazes de amplificar todo o gene de

resistência. Tais fragmentos foram, então, re-precipitados com etanol 100 % a $-20\text{ }^{\circ}\text{C}$ e utilizados para a reação de ligação, descrita em 6.1.2, a *plas*, *phadA* e *ptabA* previamente digeridos.

Os plasmídios obtidos, *plas.eri*, *phadA.eri* e *ptabA.eri* foram clonados novamente em *E. coli* JM 109 e as colônias mutantes foram selecionadas em meio LB ágar adicionado de ampicilina e eritromicina a 100 e 300 $\mu\text{g/ mL}$, respectivamente.

Com a obtenção dos clones, os plasmídios foram extraídos e utilizados nos ensaios de transformação descrito no próximo tópico.

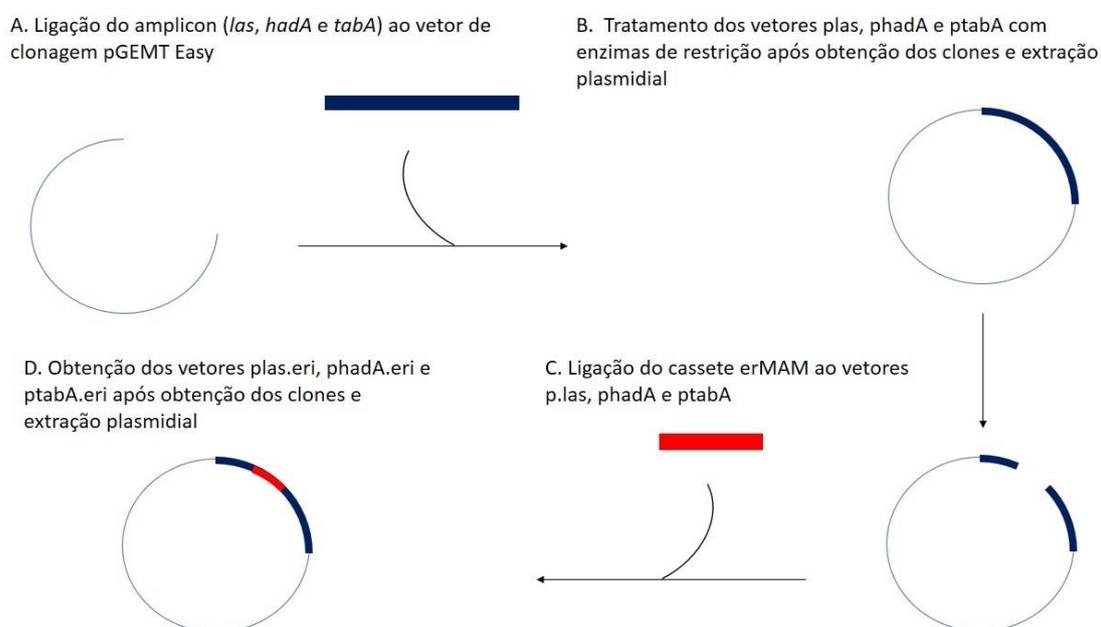


Figura 4. Esquema da metodologia utilizada para a obtenção dos vetores *plas.eri*, *phadA.eri* e *tabA.eri* para os ensaios de *knockout* gênico em *Hae*.

6.2. Transformação em *Hae* para *knockout* gênico

Apesar de existir uma metodologia publicada para transformação de *Hae* (Segada & Lesse, 1997), nenhum outro grupo de pesquisa conseguiu reproduzi-la (Li et al., 2003, Serruto et al., 2009). Como a manipulação genética em *Hae* é um assunto controverso e apenas uma vez registrada como bem sucedida na literatura (Segada & Lesse, 1997), optou-se por uma nova abordagem.

Pode-se induzir a competência *in vitro* de bactérias não competentes pelo tratamento com agentes químicos, como CaCl_2 e choque térmico, como foi o caso de *E.coli* JM109 no tópico 6.1.3 de Materiais e Métodos. Os tratamentos químico e térmico desestabilizam a membrana celular, permitindo a captação de DNA.

Sabe-se também que o uso de nanopartículas de sílica aumenta a taxa de transformação em *N. meningitidis* (Hollanda et al., 2011) e em Hi (Varela et al., 2014).

Assim, foram realizados quatro tipos de ensaios de transformação de Hae: 1. sem tratamento químico e térmico e sem o uso de nanotubos de sílica SBA 15 5x2; 2. com tratamento químico e térmico e sem o uso dos nanotubos; 3. sem tratamento químico e térmico e com uso dos nanotubos ;4. com tratamento químico e térmico e com nanotubos. As nanopartículas de sílica SBA 15 5x2 usadas nestes ensaios foram obtidas do grupo de pesquisa do Centro de Desenvolvimento da Tecnologia Nuclear da Comissão Nacional de Energia Nuclear em Belo Horizonte, Minas Gerais, Brasil.

As linhagens Hae-BPF Karina, 254/86 e 258/86 foram crescidas em meio BHI suplementado com NAD e hemina e incubadas a 37 °C em 5 % de CO_2 por 18 - 20 horas. Após o crescimento, as amostras foram ou ressuspensas em 1 mL BHI líquido (suplementado com NAD e hemina) ou então encaminhadas ao tratamento para indução de competência. Esse tratamento consistiu em lavagens das bactérias em banho de sais e choque térmico, similar à indução de competência para linhagens de *E. coli* JM109. As amostras foram ressuspensas em 1 mL de solução estéril de $\text{MgCl}_2\text{-CaCl}_2$, centrifugadas a 5000 g por 10 minutos. O *pellet* foi novamente lavado com $\text{MgCl}_2\text{-CaCl}_2$ e centrifugado e depois, ressuspensado em 1 mL de BHI líquido (suplementado com NAD e hemina).

Em uma placa de 24 *wells*, adicionou-se por *well* 40 μL dos vetores contruídos para *knockout* gênico (*plas.eri*, *phadA.eri* ou *ptabA.eri*), 200 μL da suspensão bacteriana

e, em alguns *wells*, 30 ng do nanotubo SBA 15 5x2. As amostras foram então incubadas a 37 °C em 5 % de CO₂ por 30 minutos e, após esse período, as amostras que receberam tratamento térmico foram colocadas em gelo por 3 minutos para ocorrer choque térmico. Então, adicionou-se 800 µL de BHI pré-aquecido a 37 °C em cada *well* e as placas foram novamente incubada em condições já citadas por 4 horas. Após esse tempo, o conteúdo de cada *well* foi inoculado em placas de meio apropriado: transformantes em ágar chocolate com eritromicina 7 µg/ mL e os controles das linhagens tratadas e não tratadas (para verificar a viabilidade e não contaminação das amostras), em BHI com NAD e hemina. As placas foram incubadas 37 °C em 5 % de CO₂ por 48 horas.

6.3. Transformação comparativa em Hae e Hi

A capacidade e taxa de transformação entre Hae e linhagens competentes de Hi também foi comparada. Para tal, ao invés de usar as construções para *knockout* dos genes de interesse, utilizou-se um vetor proveniente de outro trabalho do grupo: o pLG2, que contém o gene *las* de 254/86 fusionado a um cassete de resistência a eritromicina *ermAM* (Cury et al., 2014). Trata-se de um vetor que não tem como objetivo deletar e inativar o gene de interesse e, sim, apenas transferi-lo, já que o cassete de resistência está adjacente ao gene *las* e não no meio deste como em *plas.eri*.

Para estes novos ensaios de transformação, foram escolhidas 2 linhagens Hae-BPF, 254/86 e 258/86, 2 linhagens Hae não-BPF, KC1018 e ATCC 11116, e 2 linhagens Hi, RdKW20 e β-lac, como controle de transformantes pois sabe-se que estas linhagens são competentes.

As linhagens bacterianas foram crescidas em meio BHI suplementado com NAD e hemina e incubadas a 37 °C em 5 % de CO₂ por 18 - 20 horas. Após o crescimento, as amostras foram ressuspensas em 1 mL BHI líquido (suplementado com NAD e hemina). Em uma placa de 24 *wells*, adicionou-se por *well* 40 µL do vetor, 200 µL da

suspensão bacteriana e, em alguns *wells*, 30 ng do nanotubo SBA 15 5x2. As amostras foram então incubadas a 37 °C em 5 % de CO₂ por 30 minutos e, então, adicionou-se 800 µL de BHI pré-aquecido a 37 °C em cada *well* e as placas foram novamente incubadas em condições já citadas por 4 horas. Após esse tempo, o conteúdo de cada *well* foi inoculado em placas de BHI com NAD e hemina com eritromicina a 7 µg/ mL e os controles das linhagens, para verificar a viabilidade e não contaminação das amostras, em placas sem antibiótico. As placas foram incubadas 37 °C em 5 % de CO₂ por 48 horas

6.4. Busca por genes relacionados à competência de *Haemophilus* em Hae

As sequências dos genes do *regulon* CRP-S de Hi (*pilA*, *pilB*, *pilC*, *comA*, *comB*, *comE*, *comF*, *comN*, *comO*, *comP*, *comEI*, *pilF2*, *rec2*, *ligA*, *HI0659*, *HI0660*, *HI1631*, *comM*, *drpA*, *radC* e *ssb*) foram obtidas no site do NCBI de uma linhagem Hi padrão de laboratório, RdKW20, e alinhadas contra os genomas de Karina, 219/86, 254/86, 284/86, 321/86, KC1018 e ATCC 11116. A busca pelos genes e alinhamento foram conduzidos através da plataforma BLAST da mesma forma como foi descrito no tópico 4.3.1 desta seção.

7. Ensaios de infecção

As infecções *in vitro* e *in vivo* objetivaram a recuperação de bactérias após o contato com o hospedeiro, a fim de se analisar possíveis características relacionadas à virulência de Hae.

7.1. Ensaio de infecção *in vitro*

7.1.1. Cultivo celular

Células Hec-1b (carcinoma de endometrial humano), A549 (carcinoma epitelial pulmonar humano) e HaCaT (queratinócito humano) foram cultivadas em meio RPMI 1640 (Cultilab[®]) suplementado com 10 % de Soro Fetal Bovino (SFB - Cultilab[®]) e mantidas a 37 °C em atmosfera de 5 % de CO₂. BPF é uma doença invasiva com

hemorragia pulmonar (Harrison et al., 2008), por isso, optou-se por utilizar A549, uma linhagem de células pulmonares. Como a linhagem de Hae Karina foi isolada da pele de um caso de BPF (Brenner et al., 1988), a linhagem HaCaT foi escolhida.

Após a confluência da monocamada celular, as células foram tripsinizadas e subcultivadas em placas de 24 *wells*. Cada *well* recebeu uma alíquota de 1 mL da suspensão celular. As placas em questão foram incubadas sob as mesmas condições supracitadas até a confluência da monocamada celular.

7.1.2. Infecção em células humanas

O teste de adesão em células foi baseado em Pereira et al. (2011). Para a realização do mesmo, 50 μ L das suspensões bacterianas de Hae-BPF 254/86 e 258/86, Hae não-BPF KC1018 e ATCC 11116, e Hi β -lac contendo 1×10^8 UFC/ mL foram adicionados a cada *well* contendo 1 mL de meio RPMI 1640 suplementado com 10 % de SFB. A placa teste de células e bactérias foi incubada a 37 °C por 6 horas. Transcorrido o tempo de infecção, cada *well* foi lavado três vezes com PBS pH 7,4 estéril a fim de se retirar bactérias que não aderiram ou invadiram as células. Então, cada *well* poderia ser destinado ou à extração de RNA com Trizol® ou à recuperação das bactérias.

Para a recuperação das bactérias, adicionou-se aos *wells* já lavados 1 mL de meio RPMI 1640 e, com uma ponteira de micropipeta estéril, raspou-se o fundo de cada *well* a fim de se lisar as células. Então, o lisado celular foi inoculado em BHI ágar suplementado com NAD e hemina (10 μ g/ mL) e as placas, então, incubadas a 37 °C em atmosfera de 5 % de CO₂ por 24 horas. As placas nas quais houve crescimento bacteriano foram selecionadas e as colônias repicadas para outra placa de ágar chocolate. Após essa segunda passagem, fez-se uma coloração de Gram de cada placa (resultante de uma única colônia inicialmente) e as bactérias foram estocadas em BHI suplementado com 30 % de glicerol a – 80 °C. Isso foi feito para garantir a pureza de cada colônia, aumentar o número

de bactérias e garantir que todas as análises da expressão de genes candidatos à virulência e do perfil de lipídios de Hae após infecção fossem realizadas com o mesmo estoque proveniente da colônia original isolada (clone). Todos os ensaios seguintes foram realizados com a terceira passagem do estoque das bactérias isoladas. A recuperação das bactérias está ilustrada na figura a seguir.

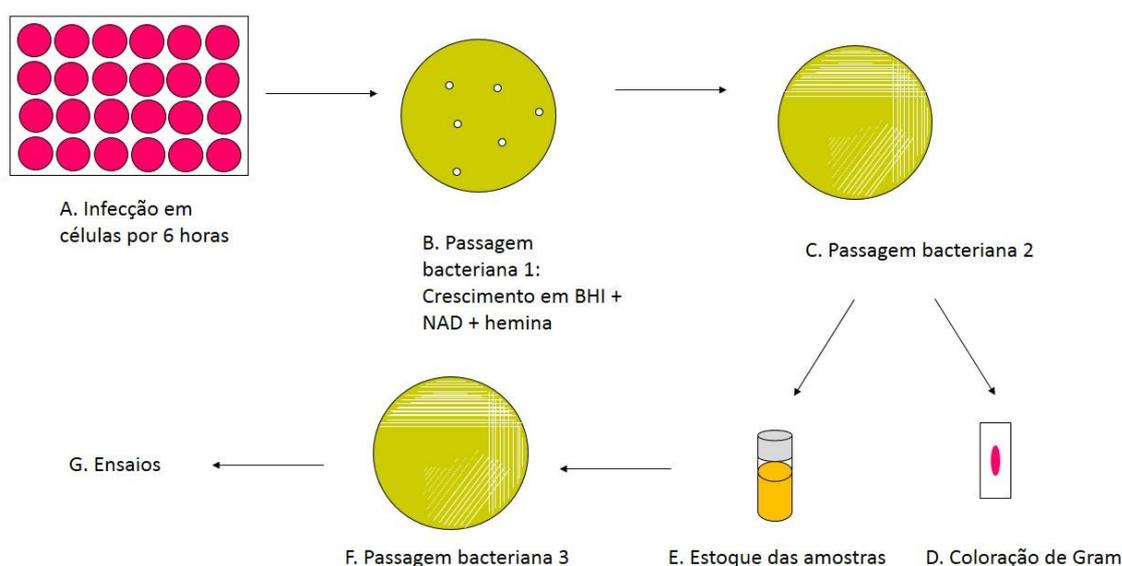


Figura 5. Esquema da metodologia utilizada para a recuperação de bactérias após infecção em células humanas.

7.2. Ensaio de infecção *in vivo*

O modelo animal escolhido foi ratos Sprague-Dawley (*Rattus norvegicus*) com 8 - 10 dias de idade. Optou-se por essa linhagem de animal e idade com base em estudos anteriores com Hae (Rubin & Carlone, 1989, Peters & Rubin, 1992, Rubin, 1995, Rubin & St Geme, 1993) e com Hi (Smith et al., 1973, Tsao et al., 2012, Hempel et al., 2013). Ainda, sabe-se que este modelo animal tem uma suscetibilidade relacionada à idade a Hi do sorotipo b (Smith et al., 1973). Como os animais eram muito jovens e, portanto, ainda não atingiram a idade de desmame, estes permaneceram o tempo todo com a rata progenitora a fim de se evitar a desnutrição dos animais. Não houve distinção entre ratos

machos e fêmeas.

Os experimentos em animais foram realizados com as linhagens bacterianas Hae-BPF 254/86, Hae-BPF 258/86, Hae não-BPF KC1018 e Hi sorotipo b β -lac. Tais linhagens bacterianas foram escolhidas pois tratam-se de 2 linhagens de Hae causadoras de BPF, uma linhagem de Hae não causadora de BPF e uma linhagem de Hi. Para cada linhagem bacteriana e tempo de infecção (24 e 48 horas), utilizou-se um n de 5 animais por grupo, totalizando 40 animais. Como a obtenção dos animais era dependente da aquisição de ratas grávidas e o número de animais era dependente da quantidade de recém nascidos, os ensaios *in vivo* foram feitos em etapas. Ainda, o número de ratos ultrapassou o n desejado (45 no total) e 3 animais foram infectados com Hae ATCC 11116 (outra Hae não-BPF) e 2 foram utilizados como controle negativo.

A infecção foi feita via intraperitoneal através da injeção de uma suspensão bacteriana em salina com 1×10^5 UFC de bactéria (Rubin & St Geme, 1993). Optou-se por essa via de infecção pois em estudos anteriores não obteve-se êxito em tentativas de se recuperar linhagens de Hae de animais infectados por via intranasal ou inoculação na conjuntiva (Rubin & Carlone, 1989).

Transcorrido o tempo de infecção de 24 e 48 horas, os animais foram anestesiados via intraperitoneal com cetamina/ xilasina à 50 mg/ kg e 5 mg/ kg, respectivamente. Retirou-se o sangue dos animais por punção cardíaca e a alíquota obtida foi transferida para um tubo de coleta de sangue com EDTA. O volume total de sangue de cada animal, o qual variou entre 0,5 a 1 mL por animal, foi dividido em duas partes: a maior parte foi destinada à recuperação das bactérias após a passagem em animal e outra, congelada em Trizol® para extração de RNA.

A eutanásia dos animais foi realizada por aprofundamento anestésico. Os protocolos de experimentação animal estão em conformidade com as regras sanitárias e

éticas descritas pelo CONCEA (Conselho Nacional de Controle de Experimentação Animal) e seguiram a aprovação do CEUA (Comissão de Ética no Uso de Animais) da UNICAMP (documento em anexo). Os animais foram descartados juntamente com os animais do Departamento de Genética, Evolução e Bioagentes (Setor de Imunologia – Instituto de Biologia – UNICAMP).

Para a recuperação das bactérias do sangue dos animais, alíquotas de sangue foram inoculadas em meio ágar chocolate em estrias de isolamento e as placas, então, incubadas a 37 °C em atmosfera de 5 % de CO₂ por 48 horas. As placas nas quais houve crescimento bacteriano foram selecionadas e as colônias repicadas para outra placa de ágar chocolate. Após essa segunda passagem, fez-se o mesmo procedimento para isolamento e estoque das bactérias isoladas das infecções *in vitro* (descrita em 7.1.2.). Ainda, é importante ressaltar que obteve-se muito pouco volume de sangue por animal, cuja maior parte, como já discutido, foi destinada à recuperação e isolamento das bactérias (metodologia ilustrada na figura abaixo).

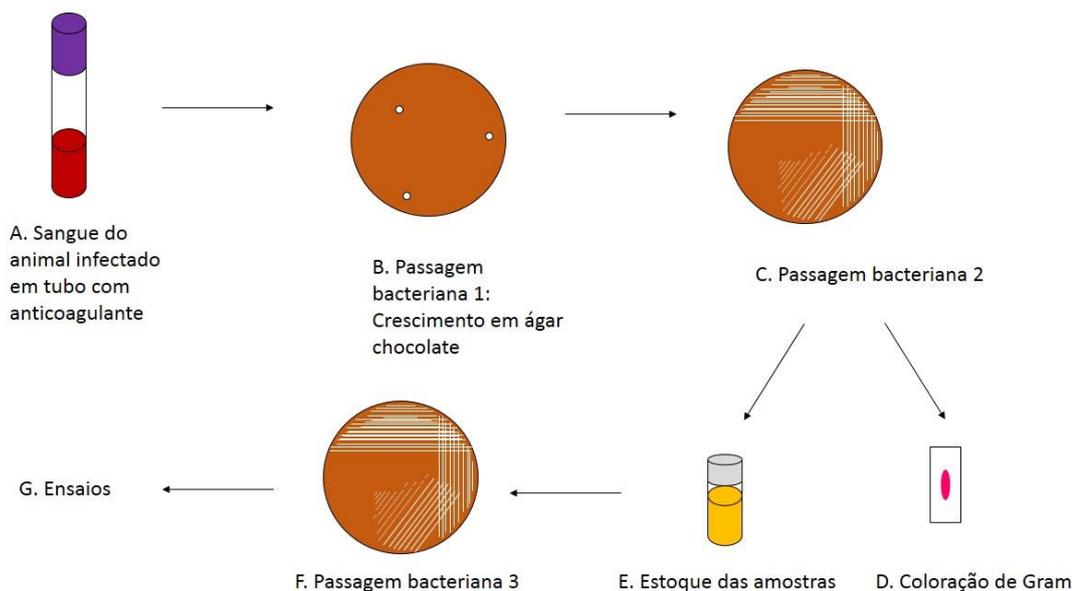


Figura 6. Esquema da metodologia utilizada para a recuperação das bactérias após infecção em animais.

8. Expressão de genes em Hae

Inicialmente, a infecção em células humanas e em ratos Sprague-Dawley visava comparar linhagens de Hae-BPF selvagens e *knockout* para os genes *las*, *hadA* e *tabA*. Porém, como não obteve-se êxito na transformação de Hae, optou-se determinar a expressão desses genes antes e depois da infecção *in vitro* e *in vivo*.

8.1. Extração de RNA das amostras

Para a extração de RNA somente das bactérias, estoques bacterianos originais e os provenientes de passagens *in vitro* e *in vivo* foram crescidas em BHI ágar suplementado com NAD e hemina em condições discutidas no tópico 1 dos Materiais e Métodos. Após o crescimento das amostras bacterianas, uma alíquota equivalente a 10 mg do crescimento bacteriano foi adicionada a 300 µL de Trizol Reagent® e homogeneizada.

As amostras em Trizol Reagent® foram estocadas a -80 °C e o RNA destas foi extraído baseado em Chomczynski (1993). Cada amostra recebeu um volume de 200 µL de clorofórmio, agitada e incubada 3 minutos a temperatura ambiente. As amostras, então, foram centrifugadas a 12000 g por 15 minutos a 4 °C. Após a centrifugação, houve a formação de 3 fases em cada *ependorf*: uma fase superior aquosa com RNA, uma interface e outra orgânica, ambas com DNA e proteínas. A fase que continha RNA foi transferida para um novo *ependorf* RNase free. Adicionou-se uma alíquota de 500 µL de isopropanol 100 % a cada tubo e estes foram incubados a temperatura ambiente por 10 minutos. Após esse período, os tubos foram centrifugados a 12000 g por 10 minutos a 4 °C e, posteriormente, os sedimentos foram lavados com 1000 µL de etanol 75 %. As amostras foram novamente centrifugadas a 7500 g por 5 minutos a 4 °C. O sobrenadante foi descartado e os *pellets* foram secos à temperatura ambiente por 10 minutos e, então, ressuspendidos em água destilada DEPC UltraPure™ (Invitrogen®) e conservado à -80 °C. O material extraído foi lido em NanoDrop (NanoDrop® 2000 - Thermo Scientific®)

para a verificação da pureza e concentração do mesmo.

8.2. Seleção dos genes de referência

A fim de se quantificar a expressão relativa de *hadA*, *las* e *tabA* em linhagens de Hae durante a infecção *in vitro* e *in vivo*, foram escolhidos 12 genes como candidatos a gene de referência para Hae, já que estes tem como objetivo a normalização entre amostras (Bustin et al., 2009). A figura abaixo representa de forma esquemática a estratégia utilizada para a seleção dos genes de referência.

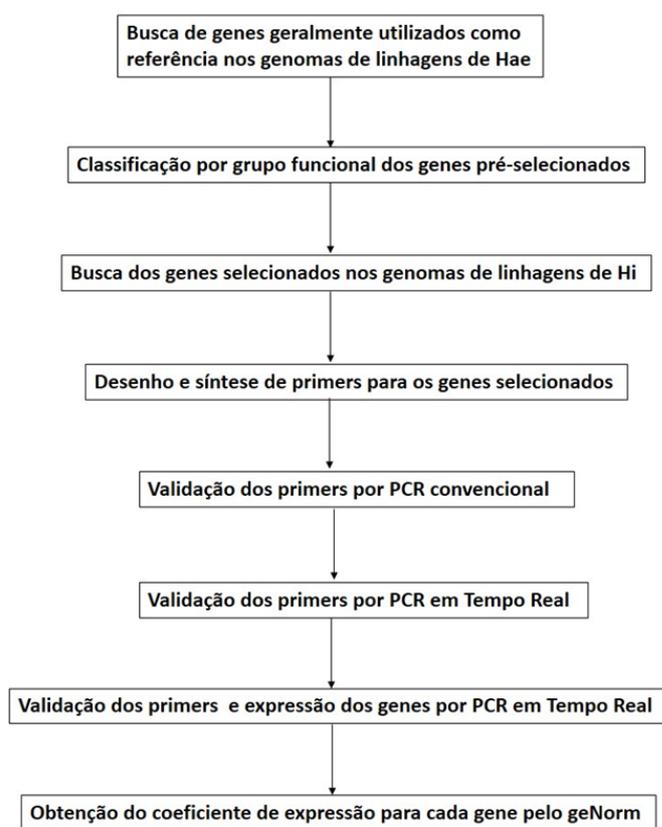


Figura 7. Pipeline para a estratégia computacional e experimental para identificar genes de referência para Hae e Hi.

Os genes em questão são: *alaS* (alanil-tRNA sintetase), *era* (proteína Era de ligação ao GTP), *gmk* (guanilato quinase), *gyrA* (subunidade A da DNA girase), *map* (metionina aminopeptidase), *primase* (DNA primase), *recA* (recombinase A), *recF* (proteína de replicação e reparo de DNA RecF), *rho* (fator de terminação de transcrição

Rho), *rpoA* (subunidade alfa da RNA polimerase), *rpoC* (subunidade beta da RNA polimerase) e *rpoD* (fator sigma da RNA polimerase). Tais genes correspondem a diferentes classes funcionais (como metabolismo de DNA, RNA e proteínas). Os *primers* estão listados na Tabela 12 em Anexos e foram desenhados de forma a terem 20 - 25 bases, um conteúdo médio de CG de 50 %, uma temperatura de melting de 60 °C e com *amplicon* de 100 a 70 - 150 pb.

Para cada reação de PCR convencional, utilizou-se o protocolo descrito em 2.3 de Materias e Métodos. Já as reações de PCR quantitativo em Tempo Real, seguiu-se a metodologia descrita a seguir.

O próximo passo foi avaliar a estabilidade da expressão dos genes sobre três condições experimentais distintas: após infecção *in vivo* e *in vitro* e estoque. Para isso, foram utilizadas as seguintes linhagens: 254/86 S (amostra isolada após infecção por 24 horas *in vivo*), 254/86 A (isolada após infecção em células A549), 258/86 (amostra estoque), 258/86 S (isolada após infecção por 24 horas *in vivo*), ATCC 11116 (amostra estoque), β -lac (amostra estoque) e β -lac H (amostra isolada após infecção em células Hec-1B).

Para a síntese do cDNA e reação de PCR foi utilizado o kit SuperScript® III Platinum® One-Step Quantitative RT-PCR System (Invitrogen®), o qual combina as enzimas Transcriptase Reversa e Taq Polimerase em um único mix enzimático. Como marcação fluorescente utilizou-se SYBR Green Fast (Applied Biosystems®). Cada reação era composta por 50 ng de RNA e 600 nM de cada *primer*. As concentrações de enzimas e fluoróforos seguiram as recomendações dos fabricantes.

As reações foram efetuadas em placas de 96 *wells* no StepOne Plus (Applied Biosystems®), nas condições de 50 °C/ 15 min (para a ação da transcriptase reversa), seguido de 45 ciclos de 95 °C/ 15 seg, 60 °C/ 1 min.

A estabilidade da expressão gênica foi avaliada pelo algoritmo geNorm (Vandesompele et al., 2002) presente no *software* DataAssist v.3.01 (Applied Biosystems®). O algoritmo em questão determina a estabilidade da expressão de acordo com o Ct (*Cycle threshold* - ciclo no qual a fluorescência emitida pelos produtos da reação de amplificação passa a ser estatisticamente significativa) e calcula um coeficiente de estabilidade para cada gene de acordo com todas as amostras.

Assim, obteve-se um coeficiente de estabilidade de expressão para cada gene, a partir destes valores, escolheu-se os candidatos a genes de referência para os próximos ensaios.

8.3. Quantificação da expressão de genes candidatos à virulência de Hae por PCR quantitativo em Tempo Real (qRT-PCR)

Com o objetivo de se avaliar o papel dos genes de interesse em Hae, adotou-se um *pipeline* semelhante ao apresentado na Figura 7, incluindo a validação por PCR convencional e em Tempo Real, exceto pela validação pelo algoritmo geNorm e obtenção de um coeficiente de estabilidade de expressão.

Os pares de oligonucleotídeos para os genes em questão foram desenhados com os mesmos critérios considerados para os genes de referência (20 - 25 bases, conteúdo médio de CG de 50 %, temperatura de *melting* de 60 °C e *amplicons* de 100 a 70 - 150 pb), como é possível ver na Tabela 12 em Anexos.

A análise dos resultados foi realizada pelo método de comparação de expressão do gene de referência, de tal modo que o Ct de cada gene alvo de uma amostra é comparado ao Ct obtido do gene de referência dessa mesma amostra, gerando o ΔCt . O $\Delta\Delta\text{Ct}$ de uma amostra é obtido através da subtração do ΔCt do gene analisado dessa amostra com o ΔCt da amostra controle. A razão da expressão desses valores é expressa nos valores de $2^{-\Delta\Delta\text{Ct}}$, que fornece a expressão de cada citocina expressa em cada tipo

celular frente à infecção bacteriana (Schmittgen & Livak, 2008).

Para a análise dos dados, foram escolhidos três genes como controle endógeno e o valor de Ct máximo permitido foi 35. Para cada linhagem bacteriana, a amostra controle (referência) foi a amostra da própria linhagem sem passagem em células ou em animal. Foram feitas triplicatas biológicas para todos os ensaios e controle negativo (reação sem *template*) para detectar falsos positivos como dimerização de *primers* e/ou contaminação da reação.

A análise estatística foi feita com o *software* Graphpad Instat 3.0 com ANOVA através da comparação das amostras entre si com pelo método de Tukey. Foram considerados significativamente diferentes os valores com $p < 0,05$. Foi comparada a expressão de um mesmo gene por linhagens diferentes.

9. Expressão de citocinas *in vitro* e *in vivo*

9.1. Extração de RNA das amostras

Como sabe-se muito pouco sobre a resposta inflamatória do hospedeiro na BPF, avaliou-se também a expressão *in vitro* e *in vivo* de IL-10 e TNF- α por qRT-PCR. Células cultivadas em placas de 24 *wells* foram submetidas ao processo de infecção por Hae, segundo metodologia descrita no tópico 7.1.2. Após o período de infecção, o meio de cada *well* foi retirado. A cada *well*, foi adicionada a alíquota de 300 μ L de Trizol Reagent[®] (Invitrogen[®]) e com o auxílio de uma ponteira de micropipeta estéril, o conteúdo total de extração de cada *well* foi retirado e transferido para um eppendorf de 1,5 mL RNase free.

Tais amostras de células e as de sangue dos animais, também estocadas em Trizol Reagent[®], foram estocadas a -80 °C e o RNA destas foi extraído da mesma forma apresentada no tópico anterior. O material extraído foi lido em NanoDrop (NanoDrop[®] 2000 - Thermo Scientific[®]) para a leitura da pureza e concentração do mesmo.

Vale ressaltar que no caso das amostras de animais, não foi possível realizar a

separação do plasma sanguíneo devido ao pequeno volume de sangue obtido para análise. Dessa forma, as análises por qRT-PCR, descritas a seguir, foram feitas com o sangue total dos animais.

9.2. Quantificação da expressão citocinas inflamatórias por qRT-PCR

Para o ensaio em questão, foram utilizados GAPDH como gene de referência e citocinas inflamatórias como genes alvo; os oligonucleotídeos estão descritos na Tabela 12 (Anexos). Os *primers* foram avaliados quanto à sua especificidade para *Homo sapiens* e *R. norvegicus*, no caso dos ensaios com células e ratos, respectivamente. Isso foi feito pela plataforma Primer-BLAST (Ye et al., 2012) em uma base de RefSeq mRNA (*Reference Sequence* – RNA mensageiro), a fim de garantir a amplificação do mRNA funcional, considerando os possíveis processamentos pós-transcricionais desta molécula.

Para a síntese do cDNA e condições das reações de qPCR em Tempo Real, utilizou-se a mesma metodologia descrita anteriormente.

A quantificação da expressão dos genes foi realizada pelo método de comparação de expressão do gene de referência GAPDH (Gliceraldeído-3-fosfato desidrogenase) e células e animais não infectados foram utilizados como amostras controle. Para todos os genes foram feitas triplicatas biológicas e controle negativo (reação sem *template*) para detectar falsos positivos como dimerização de *primers* e/ou contaminação da reação.

A análise estatística dos dados de real time PCR foi realizada utilizando ANOVA com as amostras comparadas entre si com pelo método de Tukey no *software* Graphpad Instat 3.0. Foram considerados significativamente diferentes os valores com $p < 0,05$.

10. Análise dos lipídios de Hae por Espectrometria de massas

Acredita-se que os lipídios constituintes de um patógeno exerçam um papel de destaque na interação com o hospedeiro e que alterações desta composição possam estar relacionadas à virulência. Assim, optou-se por estudar o papel de lipídios na virulência

de Hae. Através de Espectrometria de massas (MS - *Mass Spectrometry*), o perfil de lipídios de amostras estoques e amostras isoladas da infecção *in vitro* e *in vivo* de Hae 254/86, Hae 258/86, Hae KC1018, Hae ATCC 11116 e Hi β -lac, foram analisados e comparados, a fim de se verificar se há alterações nos lipídios de Hae após o contato com o hospedeiro.

As análises por MS são resultado da colaboração da Professora Doutora Alexandra C.H.F. Sawaya e da doutoranda Begoña Gimenez-Cassina (Depto. Fisiologia Vegetal, IB, UNICAMP).

10.1. Amostras padrões

As análises por massas foram realizadas com as linhagens Hae 254/86, 258/86, KC1018, ATCC 11116 e Hi β -lac. Utilizou-se amostras destas linhagens isoladas do sangue de ratos Sprague-Dawley infectados (passagem em animal), isoladas após a infecção em células Hec-1B, A549 e HaCaT (passagens em células) e amostras dos estoques originais (sem passagem de infecção).

Antes da análise lipídica de Hae, realizou-se uma busca na literatura dos ácidos graxos e fosfolipídios que seriam possivelmente encontrados nas amostras de Hae. E, com base nessa pesquisa literária, utilizou-se como padrões os lipídios cardiolipina, monosialotetrahexosilgangliosido, fosfatidilcolina, fosfatidiletanolamina, fosfatidilinositol, fosfatidilserina e esfingomielina.

10.2. Extração de lipídios totais de Hae

A extração de lipídios totais de Hae seguiu o protocolo de Bligh & Dyer (1959). As bactérias foram crescidas em BHI suplementado com NAD e hemina em condições já descritas por até 20 horas. Obteve-se suspensões bacterianas em salina 0,5 % estéril na escala de McFarland ($1,5 \times 10^8$ UFC/ mL).

A uma alíquota de 1 mL de suspensão bacteriana em salina 0,5 % estéril,

adicionou-se 1 mL de uma solução de 2:1 de clorofórmio:metanol. As amostras foram incubadas a temperatura ambiente (25 °C) e sob leve agitação durante 1 hora. Depois, centrifugou-se a 3500 g por 15 minutos, o que resultou na formação de 2 fases: fase aquosa (fase superior) e de clorofórmio (fase inferior). A fase aquosa foi descartada e a fase de clorofórmio, a qual continha os lipídios, foi totalmente seca em fluxo de nitrogênio e mantida a -80 °C.

É interessante ressaltar que no caso da análise dos lipídios das bactérias após infecção, as bactérias foram crescidas novamente em meio BHI ágar, suplementado com NAD e hemina a 10 µg/ mL, antes da extração lipídica. Isso foi feito para descartar os lipídios das células infectadas que dificultariam a análise dos lipídios de Hae.

Para a corrida das amostras, os lipídios extraídos foram ressuspensos em metanol. Deste volume, retirou-se uma alíquota de 50 µL e esta foi adicionada a 50 µL de água deionizada previamente colocada em um vial próprio para ser alocado no espectrofotômetro de massas.

10.3. Dosagem de fosfato

A fim de se padronizar quantidade de lipídio das amostras utilizadas em cada corrida do massas, o teor de fosfolipídio de cada amostra foi quantificado. Como os fosfolipídios apresentam um grupo fosfato em sua estrutura, pode-se determinar a concentração de tais lipídios na amostra através da dosagem de fosfato. Para tal, utilizou-se como base o protocolo de New (1990).

Após a secagem das amostras em fluxo de nitrogênio, estas foram diluídas em 1 mL de clorofórmio. Desse 1 mL, retirou-se 100 µL para a dosagem de fosfato e o restante da amostra foi seco novamente em nitrogênio e estocado a - 80 °C para as análises com o massas. Os tubos para a dosagem de fosfato foram incubados à 40 °C em estufa para total evaporação.

Após a secagem das amostra na estufa, adicionou-se 50 µL de ácido perclórico (HClO₄) e as amostras foram incubadas a 180 °C por 1 hora. Então, adicionou-se 900 µL do reativo de trabalho (Molibdato de amônio – Mo(NH₄)₂ a 2,78x10⁻² M; H₂SO₄ a 0,4 M, Ácido Ascórbico a 8,2x10⁻² M) e as amostras foram incubadas a 45 °C por 20 minutos. Após a reação, a absorvância ótica do fosfato foi medida em 820 nm.

A concentração de fosfato em cada amostra foi determinada com o auxílio de uma curva padrão, obtida com concentrações conhecidas de fosfato.

10.4. Análise por *Ultra-performance liquid chromatography-electrospray tandem mass spectrometry* (UPLC-ESI-MS)

Para analisar os lipídios presentes nas amostras foi desenvolvido inicialmente um método para determinar as melhores condições de ionização por *electrospray* (ESI) e análise cromatográfica usando padrões de lipídios conhecidos.

Foram otimizados todos os parâmetros de ionização para cada lipídio (modos de ionização mais adequados, voltagem do capilar, voltagem do cone), assim como as energias de colisão ótimas para realizar a fragmentação por ESI-MS/MS. As condições padronizadas da ionização ESI no modo negativo (-) e positivo (+) estão descritas na Tabela 2. Em ambos os casos, a temperatura da fonte e a temperatura de dessolvatação foram de 130°C e 350°C, respectivamente. Na identificação dos íons foi realizada uma Dissociação Induzida por Colisão (CID) e os espectros MS/MS foram comparados com os padrões, usando uma energia de colisão entre 30 e 90 segundo o composto analisado.

A análise cromatográfica das amostras (1 mg/ mL) foi feita em duplicata em um cromatógrafo UPLC Acquity acoplado com um espectrômetro de massas TQD Acquity (Micromass-Waters Manchester, England), com fonte de ESI, usando uma coluna C18 BEH Waters Acquity (2.1mm x 150 mm x 1.7 µm tamanho de partícula). O gradiente do método cromatográfico foi: solvente A usado foi água Mili-Q e o solvente orgânico B

metanol e um fluxo de 0,25mL/ min. O gradiente de eluição otimizado começou com 55 % de metanol até 100 % no minuto 8, 100 % desde 8.01 até 9 minutos e finalmente voltando à situação inicial.

10.5. Análise por *Electrospray tandem mass spectrometry* (ESI-MS)

Foi realizada uma análise das amostras por *fingerprint* usando um espectrômetro de massas TQD Acquity (Micromass-Waters Manchester, England), com fonte de ESI, e realizando inserção direta das amostras de Hae. As condições da ionização ESI no modo negativo e no modo positivo foram as mesmas descritas no tópico anterior.

Íons em comum encontrados em mais de uma amostra foram comparados quanto ao seu perfil de fragmentação por ESI-MS/MS para confirmar que eram de fato a mesma molécula.

10.6. Análise dos dados e estatística

Foi realizada uma análise multivariada dos dados, por meio da análise de agrupamentos hierárquicos (HCA - *Hierarchical Cluster Analysis*) e análise de componentes principais (PCA – *Principal Component Analysis*) usando o *software* The Unscrambler 10.2. Os dados foram previamente normalizados pela média e foi aplicado o pré-tratamento de autoescalamento.

Tabela 2. Parâmetros de ionização para cada lipídio padrão.

	íon (m/z)	Tempo (min)	Modo	Capilar (kV)	Cone (V)	Energia de Colisão	Faixa das massas analisadas (m/z)
Cardiolipina	1449.5	2.04	ESI -	3.85	50	90	100 - 2000
	1471.5	2.07					
Esfingolmielina	726.5	4.9	ESI +	3.83	43	30	600 - 1600
Fosfatidilcolina	781.5	5.17	ESI +	3.83	43	35	600 - 1600
	783.5	4.5 - 6.5					
Fosfatidiletanolamina	761.5	5.99	ESI +	3.83	43	30	600 - 1600
	763.5	5.98					
Fosfatidilserina	811	6.09/ 7.63	ESI +	3.83	43	35	600 - 1600
	833	6.15				35	
Monosialotetrahexosilgangliosido	1545.5	1.88	ESI -	3.85	50	90	100 - 2000

RESULTADOS E DISCUSSÃO

1. Sequenciamento e Montagem do genoma de Hae 258/86

1.1. Pirosequenciamento de DNA de Hae 258/86

O DNA total de Hae 258/86 extraído apresentou valor de razão de absorvância (A) A_{260}/A_{280} entre 1,8 e 2,0 e concentração de aproximadamente 75 g/ μ L. O material obtido foi então encaminhado à empresa Helixxa para sequenciamento.

Considerando-se que o genoma de Hae 258/86 seja igual em tamanho ao genoma de Hae 254/86 com 1,99 Mb (Strouts et al., 2012) e que 258/86 apresente 10 cópias do plasmídio constitutivo de aproximadamente 32 kb, o mesmo número estimado para a linhagem 254/86 (Kroll et al., 2002), foi possível estimar que a cobertura genômica resultante do pirosequenciamento foi de 44 vezes, com um total de 101761776 bases.

Uma análise mais detalhada dos *reads* presentes em ambos os arquivos *.sff revelou que o maior *read* apresenta tamanho de 844 bases e o tamanho médio dos 257840 *reads* é de 394 bases.

Ao analisar o gráfico da Figura 8, com a distribuição das leituras do sequenciamento conforme seu tamanho, pode-se notar que a maioria dos *reads* (218017 leituras – 84,56 %) encontra-se na faixa de 301 a 600 bases de tamanho.

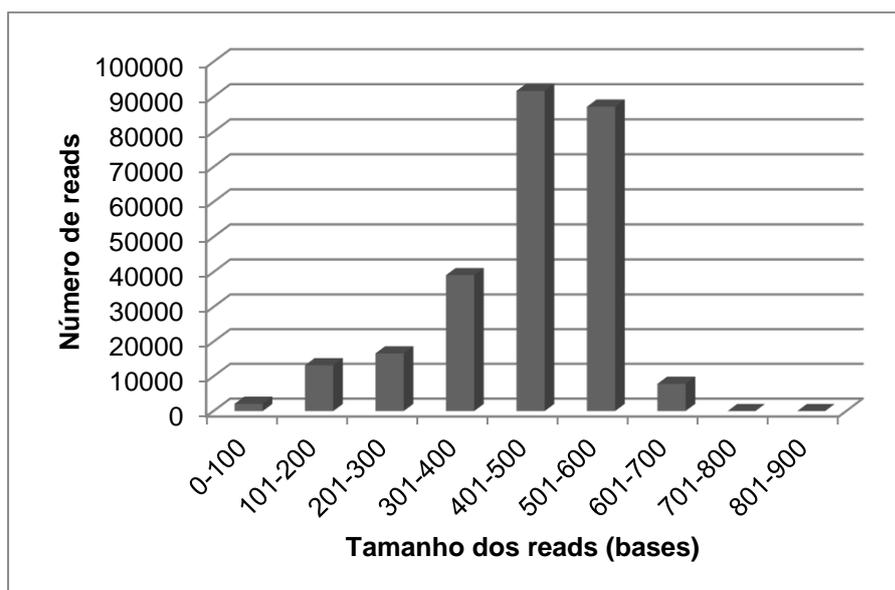


Figura 8. Distribuição dos *reads* resultantes do pirosequenciamento do DNA total de Hae 258/86 de acordo com seu tamanho em bases.

1.2. Montagem do genoma de 258/86

Como abordado no tópico Materiais e Métodos, foram realizadas três montagens: duas no *gsAssembler* e uma no *gsMapper*. A montagem 1 e 3 foram realizadas com os dois arquivos *.sff no *gsAssembler* e no *gsMapper* respectivamente. A montagem 2 foi feita no *gsAssembler* com apenas o arquivo *.sff com o maior número de *reads*.

As montagens foram avaliadas pelo número de *contigs*, tamanho do maior *contig*, N50 e tamanho da montagem, sendo considerado o *assembly* mais eficiente o com o menor número de *contigs*, maior tamanho médio dos *contigs* e maior montagem. Tais parâmetros de cada montagem estão descritos na Tabela 3 abaixo.

Tabela 3. Características das montagens dos *reads* do pirosequenciamento do DNA total de Hae 258/86 pelo *gsAssembler* e *gsMapper*.

	Montagem 1	Montagem 2	Montagem 3
Número de Bases - tamanho da montagem	1871196	1869082	1621061
Número total de <i>contigs</i>	55	56	241
Número de <i>contigs</i> \geq 500 bases	39	40	185
Maior <i>contig</i>	242761 pb	242758 pb	53376 pb
Tamanho médio dos <i>contigs</i>	47979 pb	46727 pb	8762 pb
N50	146556 pb	146554 pb	17977 pb

De acordo com os dados da Tabela 3, observa-se que as montagens *de novo* (1 e 2) apresentaram poucas diferenças entre si com a diferença de 1 *contig* entre as mesmas, provavelmente, os *reads* presentes somente na montagem 1 foram os responsáveis pelo o único *contig* a menos em tal montagem. Já a montagem por referência (3) apresentou cerca de 4 vezes a mais o número de *contigs* e a montagem final foi significativamente menor.

Por fim, a montagem genômica 1 foi considerada a melhor. Os *contigs* dessa montagem com mais de 500 bases, 39 no total, estão representados pelo seu tamanho na Figura 9 e, a partir destes, prosseguiu-se a finalização do genoma de Hae 258/86.

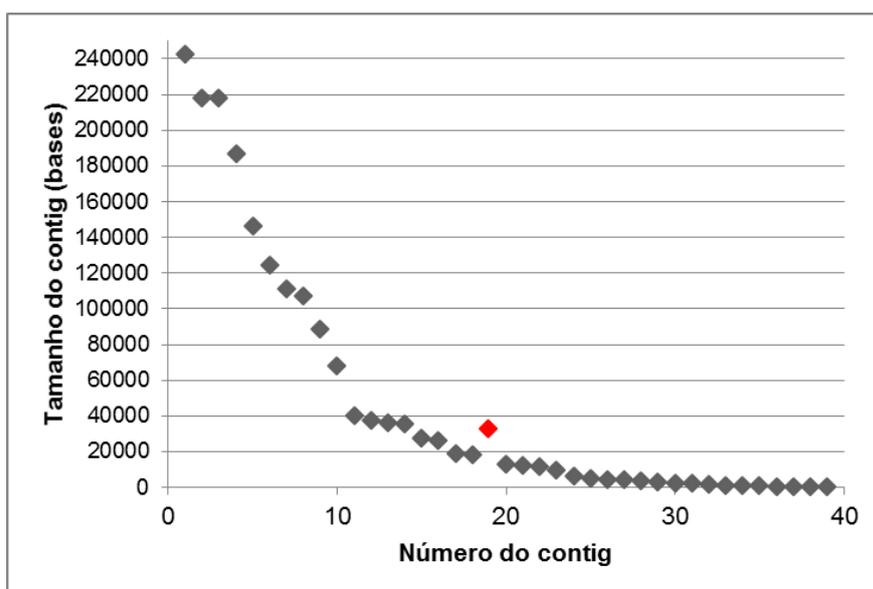


Figura 9. Distribuição dos *contigs* obtidos na montagem 1 *de novo* de acordo com seu tamanho em bases. *O *contig* 19 está em destaque e será discutido em um tópico à parte.

Uma análise da soma acumulada de tais *contigs* da montagem 1, dos maiores para os menores, foi feita com o objetivo de visualizar a cobertura genômica obtida no processo de sequenciamento (Figura 10). O *contig* 19 foi descartado dessa somatória e será discutido no tópico referente ao plasmídio constitutivo de 258/86.

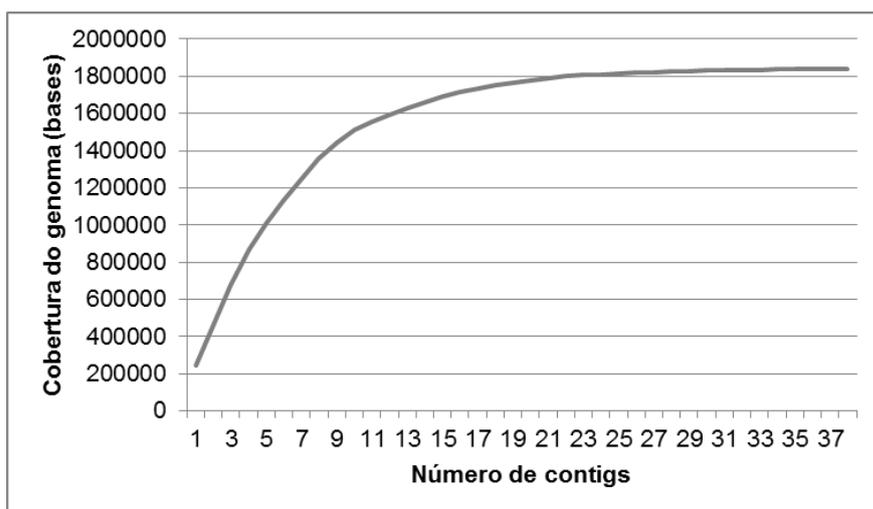


Figura 10. Cobertura genômica dos *contigs* da montagem 1 pelo *gsAssembler* Newbler 2.8.

Estimando-se novamente que o cromossomo de Hae 258/86 seja similar em tamanho ao genoma da linhagem 254/86, ou seja, presente 1985832 bases (Strouts et al., 2012), o draft genômico de Hae 258/86, correspondente a 1838571 bases (tamanho da montagem 1 menos o *contig* 19), representa aproximadamente 92,5 % do genoma completo. Isso significa que restam cerca de 150 kb para a obtenção do genoma completo. Para tal, pode-se sequenciar novamente o genoma, preferencialmente por outra técnica, utilizar os *reads* não alinhados na montagem para o fechamento de *gaps* do draft genômico ou recorrer à técnica de PCR para *gaps* não muito grandes.

1.3. Concatenação *contigs* de Hae 258/86 por referência e por PCR

Obteve-se a ordem relativa dos *contigs* de 258/86 tendo como referência o genoma de 254/86 através da ferramenta MCM (Darling et al., 2010, Rissman et al., 2009). Para concatenar os possíveis *contigs* adjacentes, os arquivos *.fasta com as 300 bases iniciais ou finais das extremidades entre os possíveis *contigs* adjacentes foram alinhados contra os *reads* do sequenciamento.

Quando haviam *reads* que atestam a união entre os *contigs*, a junção entre estes foi considerada verdadeira e estes foram concatenados. É importante ressaltar que toda vez que dois *contigs* eram concatenados, os *contigs* eram novamente alinhados contra o

genoma de 254/86 pelo MCM. Isso foi feito até não ser mais possível a junção entre *contigs*.

No caso de poucos ou nenhum *read* entre as extremidades dos *contigs*, foram construídos *primers* para se avaliar a junção entre os mesmo por PCR convencional.

Por fim, o *draft* do genoma de 258/86, com 1838571 pb, ficou esquematizado da seguinte forma:

Tabela 4. Organização dos *contigs* de Hae 258/86.

Junção entre <i>contigs</i> de 258/86	Número de bases
6 -- 23 --22 -- 27 -- 26	154127
24 -- 17 -- 18 -- 04 -- 20 -- 16	270521
25	5134
8	106928
13--11	76741
28 -- 14 -- 21 -- 31 -- 33 -- 35 -- 36 -- 34 -- 02 -- 01 -- 05	664906
15	27482
10 -- 12	105655
9	88913
3 -- 7	328975
29	3082
30	2655
32	1653
37	611
38	601
39	587
Total da montagem	1838571

1.4. Anotação automática do genoma de 258/86

A anotação através da plataforma RAST (Aziz et al., 2008) dos 39 *contigs* gerados pela montagem 1 resultou em um total de 1.941 características anotadas, as quais incluem CDS (*Coding Sequence* – equivalente a genes codificantes de proteínas) e RNAs (como tRNAs e rRNAs). Foram observadas 1890 CDS, destas 64 % estão alocadas em subsistemas e 16,61 % são CDS hipotéticas. Segundo Strouts et al. (2012), o genoma da

linhagem 254/86 apresenta 1892 CDS, o que reforça a hipótese de que os genomas dessas duas linhagens de Hae são similares quanto ao seu tamanho. A distribuição das CDS em subsistemas permitiu uma análise qualitativa dos genes anotados de acordo com sua função ou grupo e pode ser observada na Figura 11.

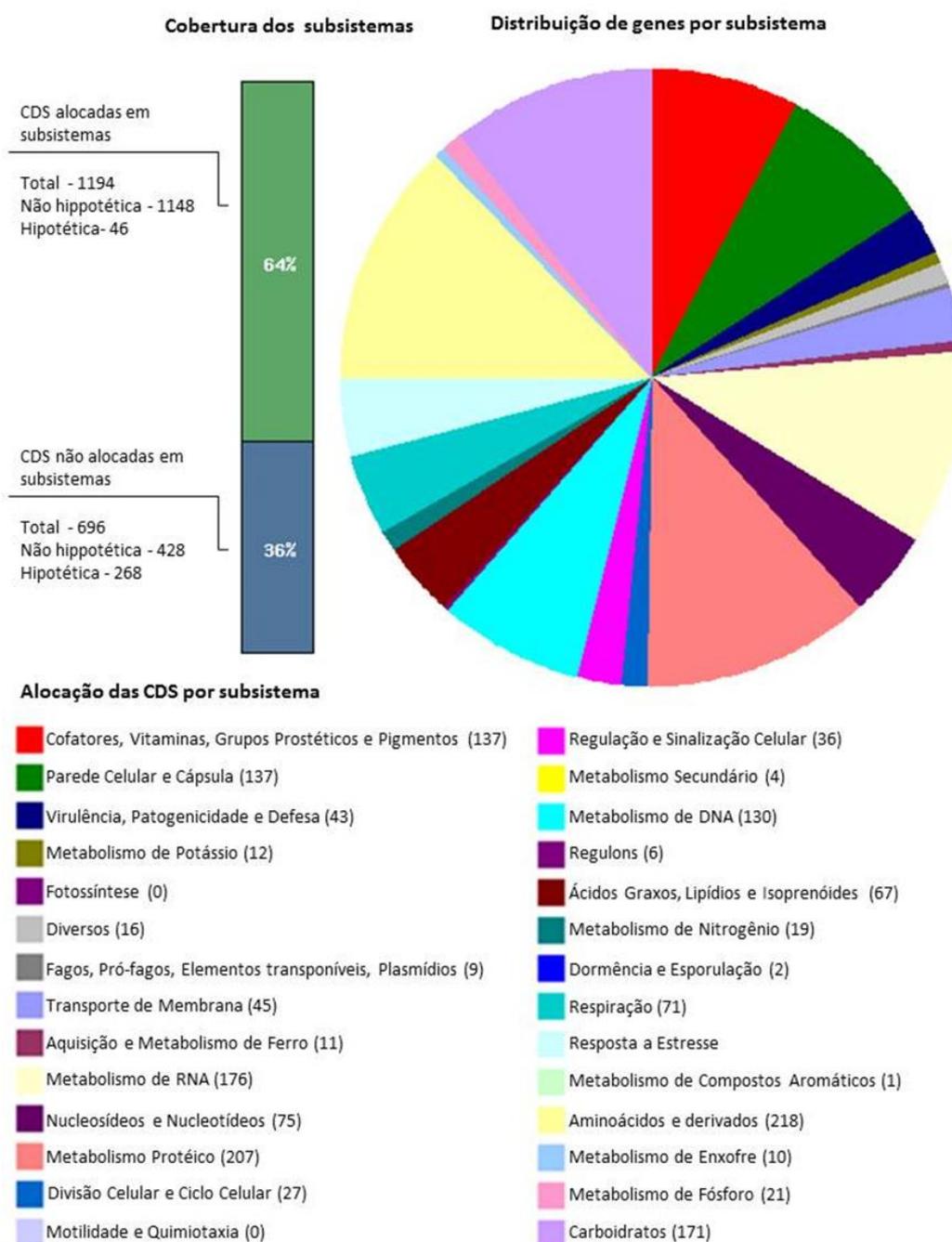


Figura 11. Alocação dos genes anotados (CDS - *Coding Sequences*) através da plataforma RAST em subsistemas.

2. Sequenciamento e Montagem dos genomas de Hae Karina, 219/86, 284/86, 321/86, 406/86 e KC1018

O DNA extraído das linhagens Karina, 219/86, 284/86, 321/86, 406/86 e KC1018 foi destinado ao sequenciamento pela plataforma Illumina. A tabela abaixo contém os dados relativos a cada amostra sequenciada.

Tabela 5. Número total de *reads* e percentual destes com qualidade aceitável gerados pelo sequenciamento por Illumina.

Amostra	Número de <i>reads</i>	% bases \geq Q30
Karina	38647860	94,4
219/86	38479940	94,53
284/86	35636936	94,71
321/86	39906340	95,14
406/86	36692872	95,19
KC1018	53773000	92,67

* % Bases \geq Q30 corresponde ao percentual de bases com a probabilidade de 1 em 1000 de estarem incorretas

Obteve-se uma média de 40522825 *reads* dos quais 94,44 % apresentaram um *score* \geq Q30. Em comparação ao sequenciamento pela plataforma 454 do genoma da linhagem 258/86 com um total de 257840 *reads*, o sequenciamento por Illumina gerou mais dados.

Os *reads* provenientes do sequenciamento foram então montados segundo o *pipeline* esquematizado na Figura 2 e os dados seguem na Tabela 6.

Tabela 6. Dados referente à montagem dos genomas das linhagens de Hae, em *contigs* maiores de 1000 bp.

	Karina	219/86	284/86	321/86	406/86	KC1018
Kmer	77	87	91	89	89	91
Número de <i>contigs</i>	53	34	35	44	39	53
Tamanho total	1959268	1965011	1976293	1967622	1967898	1912257
Tamanho médio do <i>contig</i>	36967.32	57794.44	56465.51	44718.68	50458.92	36080.32
N50	91123	126358	126928	99264	100077	95241
Tamanho do maior <i>contig</i>	251830	298923	299495	299419	299750	382954
Tamanho do menor <i>contig</i>	1056	1056	1056	1056	1056	1009

Ao comparar os dados da tabela acima com os obtidos para o sequenciamento de 258/86, pode-se observar que o número médio de *contigs* (43 e 55 para os sequenciamentos por Illumina e 454, respectivamente) e o tamanho médio dos *contigs* (47080 e 47979 pb para Illumina e 454, respectivamente) são similares. Já o o tamanho dos maiores *contigs* dos sequenciamentos por Illumina (média de 305395 pb) é relativamente maior do que o maior *contig* da montagem de 258/86. Ainda, o tamanho total das montagens de Karina, 219/86, 284/86, 321/86 e KC1018 com média de 1958058 pb é maior que a montagem de 258/86, com 1808273 pb, e corresponde a 98,6 % do tamanho do cromossomo de Hae 254/86.

Dessa forma, pode-se concluir que o sequenciamento e, logo, a montagem dos genomas obtidos por Illumina foram mais eficientes.

3. Comparação entre os genomas

3.1. Comparação pan-genômica entre Hae e Hi

Foi realizada uma comparação dos genomas de linhagens de Hae e Hi listadas na

Tabela 13 (Anexos) aos pares, ou seja, dois genomas foram comparados por vez. Os dados foram plotados em um *heat map* os quais podem ser observados na Figura 12 a seguir.

Os resultados apresentados sugerem pequenas diferenças entre os genomas de Hae-BPF, ainda que as linhagens sejam similares entre si, como já foi relatado na literatura.

A linhagem não causadora de BPF KC1018 destacou-se entre as demais, tanto Hae quanto Hi. Tal fato pode ser justificado pela origem geográfica da amostra (Estados Unidos, Tabela 1).

Entre as linhagens associadas a BPF, 258/86 e 284/86 destacaram-se entre as demais. Segundo os dados, a organização genômica de 258/86 é mais próxima a linhagens Hi RdKW20, R2846 e 10810 (Figura 12). A linhagem 284/86, por sua vez, possui organização genômica mais semelhante a 321/86 e 406/86 (Figura 12) e curiosamente é a única linhagem de Hae-BPF neste estudo isolada da orofaringe (Tabela 1). Apesar de tais diferenças, tais linhagens apresentam genes específicos do grupo Hae-BPF, tais como *las*, *hadA* e *tabA*. Contudo, tais diferenças serão foco de estudo de próximos trabalhos do grupo.

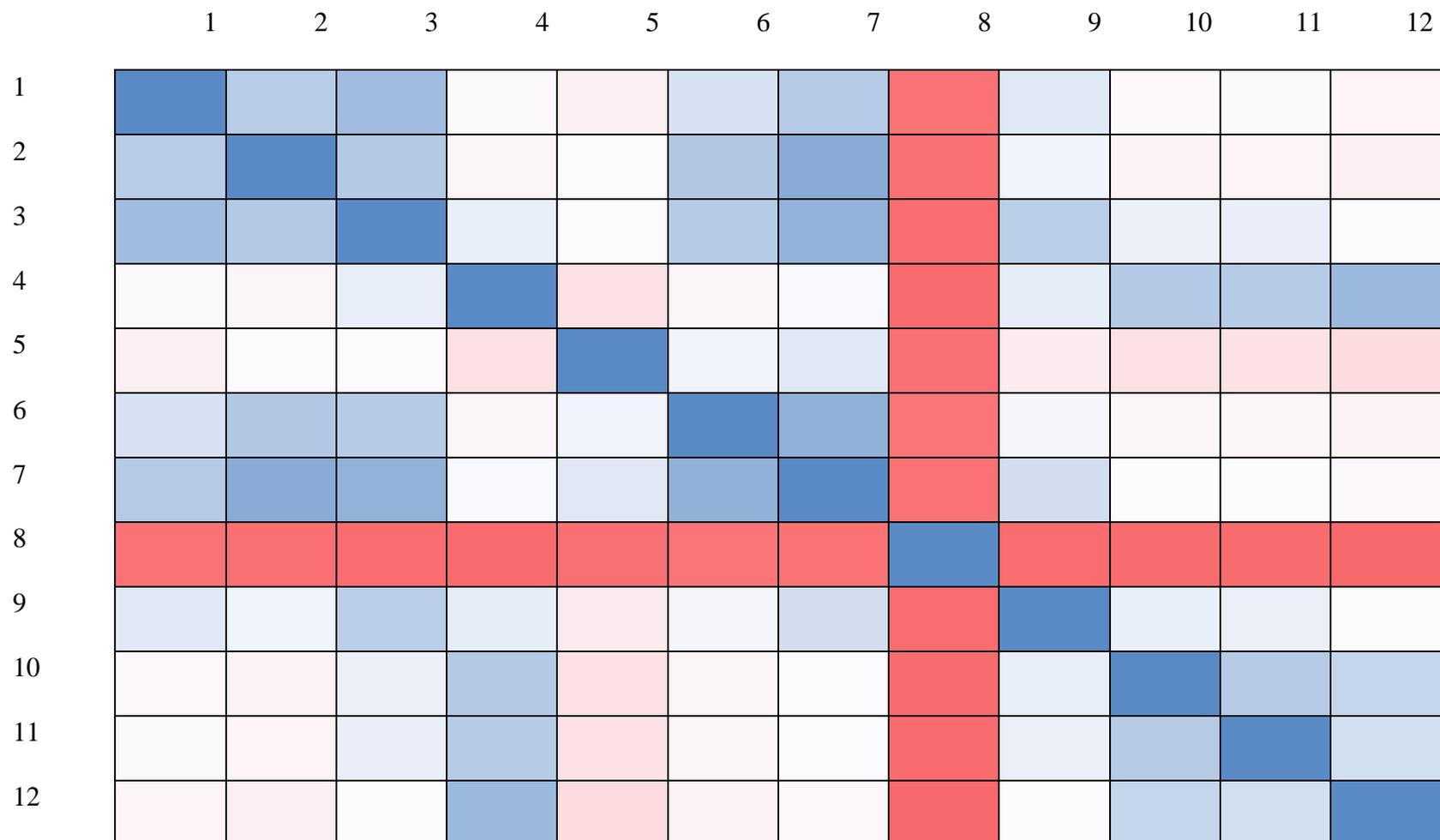


Figura 12. Comparações pareadas entre os genomas de Hae e Hi. Cada bloco colorido representa o número de bases compartilhadas entre os genomas. Quanto mais azul o bloco, mais similares são as sequências entre si. Quanto mais vermelho o bloco, maiores são as diferenças entre os genomas. 1. Karina. 2. 219/86. 3. 254/86. 4. 258/86. 5. 284/86. 6. 321/86. 7. 406/86. 8. KC1018. 9. F3047. 10. RdKW20. 11. R2846. 12. 10810.

3.2. Comparação entre os genomas de Hae 258/86 e outras linhagens Hae-BPF

Os *contigs* resultantes da montagem 1 foram alinhados contra o genoma completo de Hae 254/86 na plataforma MAUVE apesar de rearranjos e inversões entre os genomas. Porém tais observações foram descartadas, já que o genoma de 258/86 é o mais incompleto.

Os *contigs* 20, 29, 30, 32, 35, 36, 37, 38 e 39 não apresentaram alinhamento ao genoma da linhagem 254/86 na plataforma MAUVE. Assim, foi feita uma busca por similaridade através da plataforma BLAST, a qual resultou no alinhamento de todos os *contigs*, exceto o 20, com os genomas de Hae-BPF 254/86, Karina, 219/886, 284/86, 321/86, 406/86.

O *contig* 20 com 11508 pb não apresentou similaridade ao genoma referência, mas alinhou-se com o DNA de outras linhagens da família *Pasteurellaceae*, como Hi R2846, *Pasteurella multocida* subsp. *multocida* linhagem Pm70 e *Haemophilus somnus* 2336. O alinhamento com Hi R2846 correspondeu a uma provável proteína de transporte de arsenito e a uma hipotética proteína conservada com 96 % de identidade e 0 % de gaps entre as duas sequências das proteínas e a sequência do *contig* 20. Tal fato pode justificar a maior semelhança da organização genômica entre Hi R2846 e Hae 258/86, como observado no *heat map* da Figura 12.

Não obstante, foram observadas regiões presentes em Hae 254/86 e ausentes em 258/86, o que podem significar diferenças entre as duas linhagens em questão. Uma região do genoma de 254/86, também presente nas demais linhagens de Hae-BPF estudadas (Karina, 219/86, 284/86, 321/86 e 406/86) não foi encontrada no *draft* genômico de 258/86. Tal região apresenta uma série de genes, como o *HIBPF13250*, considerado como parte do genoma acessório específico de Hae-BPF e encontra-se entre

regiões de bacteriófago no genoma de 254/86 e hipoteticamente codifica uma proteína conservada com 75 % de identidade à de *Haemophilus parasuis* (Strouts et al., 2012). Tal fato pode indicar uma possível inserção ou deleção desse gene no genoma de 258/86 (dependendo de qual linhagem for a ancestral). Assim, sugere-se que existem diferenças entre Hae 258/86 e as demais Hae-BPF, como sugere também a Figura 12. É interessante ressaltar, contudo, que o *draft* genômico de 258/86 é o mais incompleto e, por isso, para análises comparativas mais detalhadas este foi descartado.

3.3. Busca por genes de diagnóstico de *N. meningitidis* em Hae

Acredita-se que BPF já foi e ainda pode ser erroneamente diagnosticada como meningococemia. Assim, realizou-se uma busca por genes utilizados no diagnóstico de *N. meningitidis* nos genomas de Hae. Através de alinhamento de sequência, foram encontrados o gene *crgA*, comumente utilizado no diagnóstico de *N. meningitidis* e também recentemente descrito em isolados clínicos de Hi (Hou et al., 2014), e o gene adjacente e regulado pelo mesmo, *mdaB*.

Todas as linhagens de Hae associadas a BPF apresentaram o *crgA*, o qual está ausente nas linhagens Hae não-BPF estudadas (Tabela 13 em Anexos). Ao considerar este fato e a similaridade dos quadros clínicos de BPF e da meningococemia (Harrison et al., 2008), é possível de fato que casos de BPF já foram diagnosticados de maneira equivocada como infecção por *N. meningitidis*. Assim, torna-se de extrema importância o desenvolvimento de métodos para a distinção entre tais patógenos.

Ainda, o *crgA* está presente em sete linhagens de *Haemophilus* disponíveis no NCBI (até agosto de 2015), as quais cinco são NTHi (dentre estas a Hae 254/86) e uma pertencente ao sorotipo f (KR494), ao contrário de Hou et al. (2014), que observaram este gene em apenas linhagens NTHi (Tabela 7).

Tabela 7. Presença ou ausência dos genes *mdaB* e *crgA* em Hae, Hi e *N. meningitidis* e características deste último.

Linagem	Bactéria	<i>crgA</i> *	<i>mdaB</i> *	Comprimento <i>crgA</i> (bp)	% GC <i>crgA</i>	% GC genoma**
Karina	Hae NTHi	+	+	888	43,8	34,9
219/86	Hae NTHi	+	+	888	43,8	35,12
254/86	Hae NTHi	+	+	888	43,8	38,2
284/86	Hae NTHi	+	+	888	43,8	31,48
321/86	Hae NTHi	+	+	888	43,8	32,8
406/86	Hae NTHi	+	+	888	43,8	36,76
KC1018	Hae NTHi	-	+	-	-	-
F3047	Hae NTHi	-	+	-	-	-
Hi375	Hi NTHi	+	+	888	43,8	38,1
R2846	Hi NTHi	+	+	888	43,7	38
MC58	Nm	+	+	900	51,4	51,8
Z2491	Nm	+	+	900	51,8	51,5

Nm - *N. meningitidis*

*Todas as linhagens de Hae, exceto as que não são associadas a BPF, apresentam *crgA*.

** percentual do *draft* genômico para Hae Karina, 219/86, 284/86, 321/86 e 406/86.

*** As linhagens Hi disponíveis no NCBI que não apresentam *crgA* mas que apresentam o gene *mdaB* não estão na tabela.

Através de comparação das sequências de *crgA* entre as linhagens da tabela acima, pode-se observar que o gene em questão difere tanto em comprimento (em pb) como no conteúdo de GC entre *N. meningitidis* e *Haemophilus*. O percentual de GC no *crgA* de Hae e Hi é maior do que o de seus respectivos genomas, enquanto que em Nm *crgA* e os genomas apresentam um conteúdo de GC semelhantes.

Ainda, o *cgrA* em todas as linhagens de *Haemophilus* analisadas (Hae e Hi –

Tabela 7), quando presente, é flanqueado por *mdaB* e *rep* (que codifica uma helicase ATP-dependente) (Figura 13). A região entre *crgA* e *mdaB* em Hae-BPF apresenta 112 pb, enquanto que em *Neisseria* sabe-se que há 239 pb entre tais genes, onde está situada uma região regulatória CREN (*Contact regulatory element of Neisseria*) (Deghmane et al., 2000), ausente em Hae e Hi. É importante ressaltar que *mdaB* é um gene adjacente ao *crgA* e regulado pelo mesmo no meningococo, porém *rep* não é adjacente ao *crgA* em *Neisseria*.

Não obstante, linhagens de *Haemophilus* que somente apresentam o *mdaB* (tais como Hae KC1018 e F3047 – Tabela 7) possuem na região *downstream* desta sequência, o gene *rep* (Figura 13). A semelhança estrutural da organização desses dois genes em linhagens de Hae e Hi, sugere uma possível perda do *crgA* na evolução de algumas linhagens de *Haemophilus*.

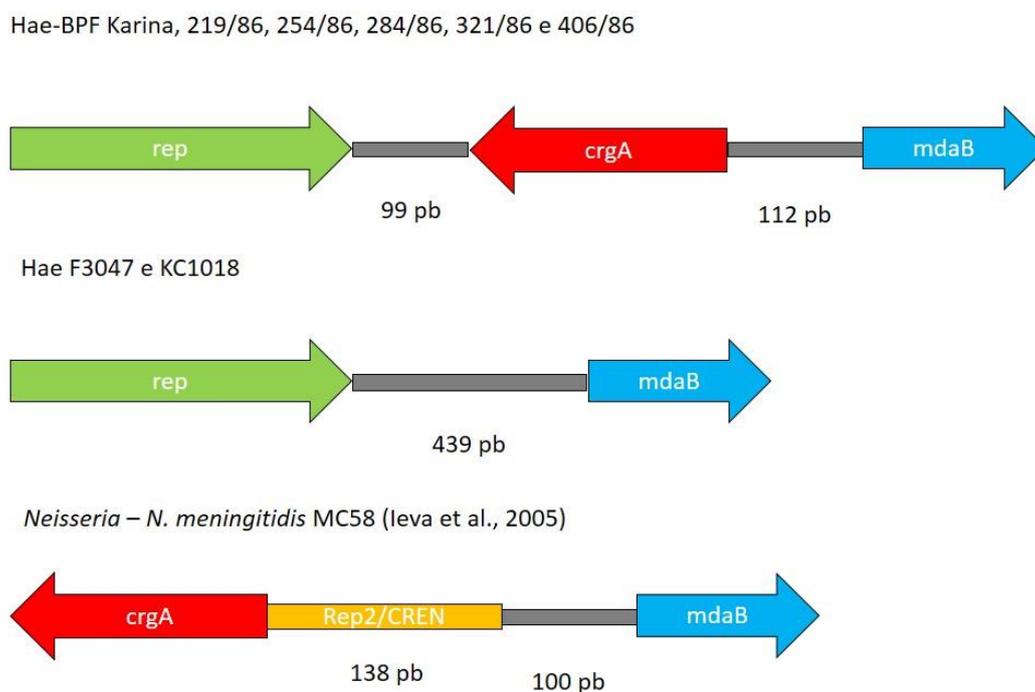


Figura 13. Organização estrutural de *crgA*, *mdaB* e *rep* em Hae-BPF; *mdaB* e *rep* em Hae não-BPF e Hi; *crgA* e *mdaB* em *Neisseria* (fora de escala). As setas indicam o sentido de transcrição dos genes.

Uma análise filogenética preliminar com linhagens que apresentam a sequência do *crgA* disponíveis no NCBI e os genomas sequenciados neste trabalho resultou em um grande *cluster* subdividido em dois grupos: um composto apenas por Hi e Hae e outro, por *Neisseria* (Figura 14). A distribuição no dendograma abaixo reforça a diferença entre o *crgA* do meningococo e de linhagens de *Haemophilus*. Também sugere que o *crgA* foi presente em um ancestral comum entre *Neisseria* e *Haemophilus* spp. ou foi adquirido via transferência lateral de genes por uma linhagem ancestral de *Haemophilus*. A diferença no conteúdo de GC observada reforçaria a segunda hipótese, já que este seria um marcador de transferência lateral gênica (Hayek, 2013). Ainda, ao considerar a incapacidade de transformação e a ausência de genes necessários à competência natural em Hae (como será abordado adiante), se houve de fato a transferência lateral de *crgA*, esta deve ter ocorrido em uma linhagem ancestral em comum a Hae e Hi.

A transferência horizontal de genes entre *Haemophilus* e *Neisseria* já foi identificada (Kroll et al., 1998, Davis et al., 2001), entretanto, acredita-se que o fluxo gênico de *Neisseria* para *Haemophilus* seja mais raro já que a transformação natural é mais frequente em linhagens de *Neisseria*. Dessa forma, mais estudos são necessários sobre a presença de *crgA* em *Haemophilus*.

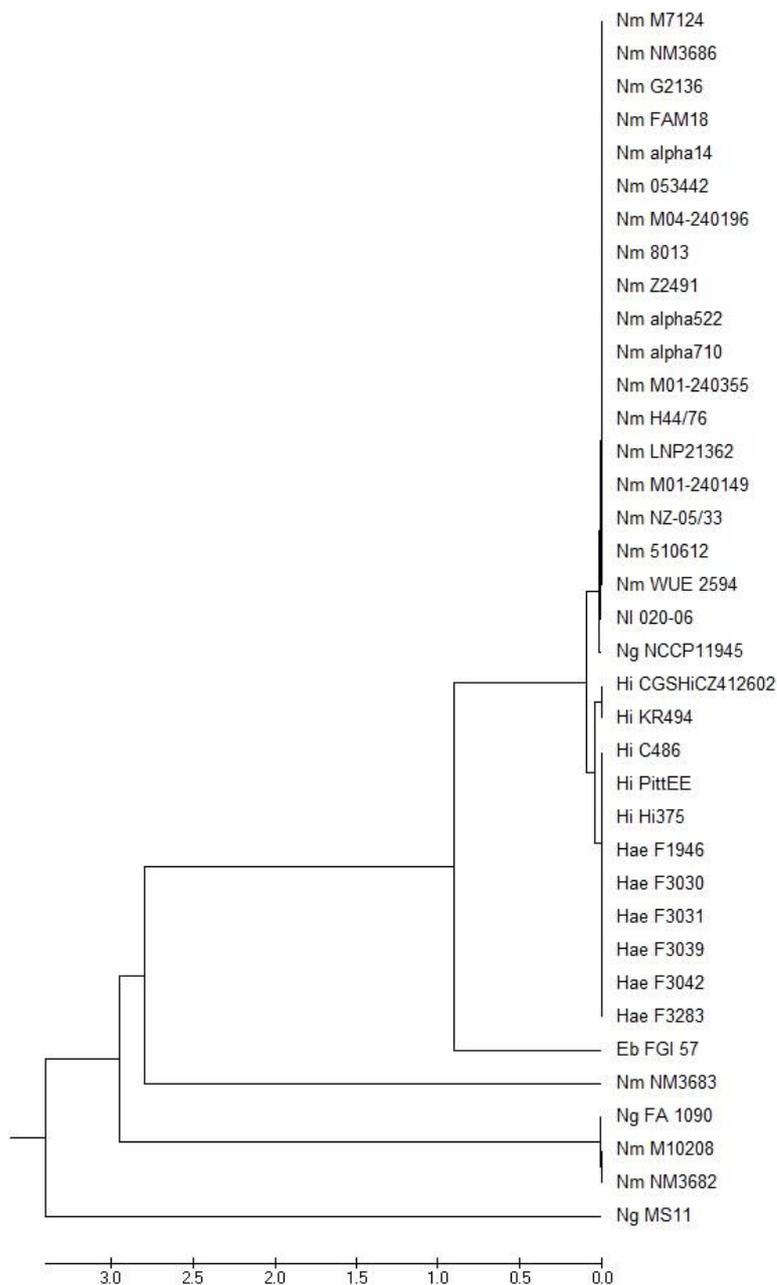


Figura 14. Clusterização das sequências de *crgA* de linhagens que apresentam o gene disponíveis no NCBI pelo método UPMGA. O grupo de *Neisseria* e *Haemophilus* spp. é subdividido em dois clusters: um composto por *Neisseria* e outro por Hae e Hi.

4. Plasmídio de Hae-BPF

4.1. Montagem e Fechamento do plasmídio de Hae

Os genomas das linhagens de Hae Karina, 219/86, 258/86, 284/86, 321/86, 406/86 e KC1018 foram sequenciados e montados. Todas as linhagens, exceto KC1018, apresentaram sequências que se alinharam contra o plasmídio de Hae 254/86. Os *contigs* correspondentes a essa estrutura de cada genoma estão descritos na Tabela 8.

Ainda, a fim de se verificar a circularidade do *contig* 19 e, logo, a confirmação de que trata-se do plasmídio constituinte de Hae 258/86, foi utilizada uma abordagem *in vitro*, a qual consistiu em desenhar *primers* para cada uma das extremidades do *contig* 19, unindo-as. Abaixo encontra-se a figura com os *amplicons* resultantes da reação descrita no tópico 5 de Materiais e Métodos.

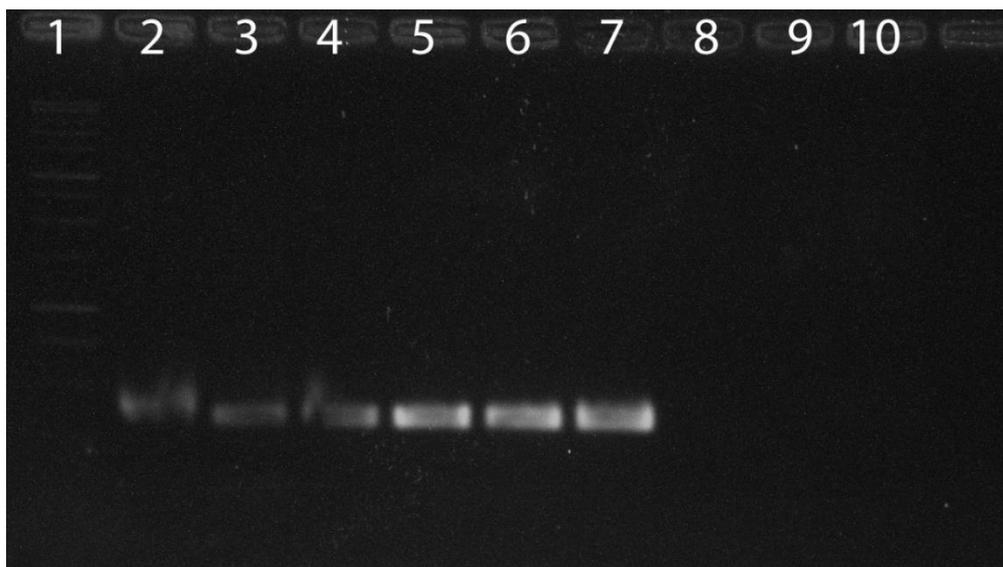


Figura 15. Gel de agarose a 1,5 % com os *amplicons* resultantes com os *primers* da descritos na Tabela 12. 1. Padrão de peso molecular 1 kb. Amplificação do plasmídio de Hae 2. Karina. 3. 219/86. 3. Karina. 4. 254/86. 5. 258/86. 6. 321/86. 7. 406/86. 8-10. Branco (controle negativo) da reação.

Como é possível observar na figura acima, houve amplificação na amostra de 258/86, o que indica a circularidade do *contig* 19 e, portanto, trata-se de fato do plasmídio constitutivo desta linhagem. A amplificação positiva nas demais amostras também indica a presença dessa estrutura nas demais linhagens de Hae-BPF.

4.2. Comparação *in silico* entre os plasmídios de Hae-BPF

Os plasmídios constitutivos das linhagens de Hae-BPF são classificados por apresentarem padrão de restrição do tipo 3031 com 7 sítios de restrição para *AccI*, 2 para *EcoRI*, 2 para *HindIII* e um único sítio para *EcoRV* (Brenner et al., 1988). Kroll et al. (2002) verificaram esse padrão para *EcoRI* e *HindIII* no plasmídio da linhagem F3028, em ensaios pré-sequenciamento dessa estrutura. Dessa forma, a fim de se verificar a

presença desse padrão nas demais linhagens de Hae-BPF deste estudo, realizou-se uma digestão *in silico* com tais enzimas através do *software* pDRAW. Os dados obtidos estão na tabela a seguir.

Tabela 8. Comparação entre os plasmídios de linhagens Hae-BPF.

Linhagem	<i>Contig(s)</i> alinhados ao plasmídio de 254/86	Tamanho (pb)	Conteúdo de CG (%)	EcoRI*	EcoRV*	HindIII*	AccI*
Karina	13 e 57	31849	36,55	2	1	2	9
219/86	1 e 5	31889	36,56	2	1	2	9
254/86	-	32433	34,52	2	1	2	10
258/86	19	32625	37,04	2	1	0	13
284/86	1, 10 e 18	31903	36,56	2	1	2	9
321/86	2 e 4	31897	36,55	2	1	2	9
406/86	5 e 23	31886	36,56	2	1	2	9

*número de sítios de restrição

Todas as sequências analisadas apresentaram um único sítio EcoRV e 2 para EcoRI. Todas, exceto o *contig* 19, apresentaram 2 sítios para HindIII. É possível que o plasmídio de Hae 258/86 apresente de fato esta diferença em relação aos demais ou essa diferença seja, provavelmente, resultado de um erro no sequenciamento, já que a mudança de uma única base afetaria o sítio de restrição da enzima.

Ainda, o número de sítios para AccI diferiram nas amostras e também do esperado segundo o padrão 3031 (Brenner et al., 1988). Tal diferença pode ser explicada pelo fato de que o sequenciamento e digestão *in silico* dessas sequências permite uma visão mais clara e maior detalhamento dos fragmentos em comparação à visualização da mesma digestão *in vitro* em gel de agarose.

5. Transformação e competência em Hae

5.1. Construção de vetores para *knockout* de genes candidatos à virulência de Hae

Os DNAs de Hae Karina, 219/86, 254/86, 258/86, 284/86, 321/86, 406/86, KC1018 e ATCC 11116 (Tabela 1) foram utilizados para a amplificação dos genes *las*, *hadA* e *tabA* com a finalidade de se avaliar a possível presença dos mesmos nessas linhagens e para posterior inativação desses genes por recombinação homóloga.

As nove linhagens de Hae apresentaram amplificação dos genes *las* e *tabA*. A presença de *las* em Hae, independente de estarem associadas ou não a BPF, atesta os dados observados por Davis et al. (2001), que identificaram o gene em questão em linhagens Hae-BPF e em Hae ATCC 11116, não causadora da doença.

É importante ressaltar que a amplificação com *primers* específicos gene *tabA* de Hae-BPF em linhagens não-BPF ocorreu na região conservada entre os genes *tabA* e *tahA* (Strouts et al., 2012). Dessa forma, estes *amplicons* em KC1018 e ATCC 11116 se referem ao gene *tahA*. Apesar de tais genes possuírem uma região de sequência em comum, acredita-se que apresentem funções diferentes, o que torna *tabA* um possível determinante da virulência de Hae.

O gene *hadA*, por sua vez, foi identificado apenas em linhagens causadoras de BPF, como já era previsto na literatura (Serruto et al., 2009, Strouts et al., 2012). Desse modo, acredita-se que *hadA* seja um dos principais responsáveis pela virulência de Hae-BPF, já que é exclusivo destas linhagens.

Os resultados do PCR de *las*, *hadA* e *tabA*, positivos ou negativos, foram corroborados por análises posteriores nos genomas das linhagens de Hae em questão. Não obstante, a expressão dos mesmos foi analisada em amostras de Hae isoladas de infecção *in vitro* e *in vivo*.

5.2. Transformação em Hae para *knockout* gênico

Como a competência de Hae é incerta, foi utilizada uma nova metodologia para tentar obter colônias de Hae transformadas. A metodologia, descrita no tópico 6.2 de Materiais e Métodos, baseou-se na competência adquirida em linhagens de *E. coli* através de tratamento com sais seguido de choque térmico e no uso de nanotubos de sílica para aumentar a taxa de transformação, como já observado em *N. meningitidis* (Hollanda et al., 2011) e Hi (Varela et al., 2014).

Diferentes abordagens foram utilizadas na tentativa de se obter transformantes de Hae com *las* ou *hadA* ou *tabA* inativado: 1. sem tratamento químico e térmico e sem o uso de nanotubos de sílica; 2. com tratamento químico e térmico e sem o uso de nanotubos; 3. sem tratamento químico e térmico e com uso de nanotubos; 4. com tratamento químico e térmico e com nanotubos. Controles negativos de cada ensaio de transformação foram feitos para se certificar da ausência de organismos contaminantes e da viabilidade das amostras, principalmente as submetidas a tratamento químico e térmico.

Obteve-se crescimento de todos os controles negativos, o que garantiu que os tratamentos não inviabilizaram as amostras de Hae. Porém, não se obteve o crescimento de nenhuma colônia de Hae transformante em qualquer um dos quatro tratamentos aplicados e para qualquer um dos genes.

Uma explicação para isso seria que se tratam de genes essenciais, ou seja, os mutantes *knockout* dos genes *las*, *hadA* e *tabA* seriam inviáveis. Outra justificativa, previamente abordada na literatura, seria a ausência da capacidade de transformação nessa bactéria, sendo esta incapaz de ser manipulada geneticamente (Segada & Lesse, 1997, Serruto et al., 2009, Li et al., 2003). Apesar de Hae aparentemente apresentar todos os genes envolvidos na competência de Hi segundo Strouts et al. (2012), a transformação

de Hae foi apenas relatada como bem sucedida uma única vez na literatura (Segada & Lesse, 1997). Posteriormente não houve relatos de sucesso na obtenção de transformantes de Hae (Li et al., 2003, Serruto et al., 2009). Infelizmente, mesmo com o uso de técnicas inovadoras, como as descritas neste trabalho, não foi possível obter linhagens de Hae *knockout* dos genes *las*, *hadA* e *tabA*.

5.3. Transformação comparativa em Hae e Hi

A manipulação genética em Hae é questionável e apenas uma vez registrada como bem sucedida na literatura. Assim, optou-se por uma nova abordagem: comparar a taxa de transformação entre Hae e Hi, uma bactéria sabidamente competente a esse processo. A comparação foi feita através da transferência de um gene ao invés de deletá-lo, já existe a possibilidade dos genes *las*, *hadA* e *tabA* serem essenciais e a deleção dos mesmos resultaria em células inviáveis.

O vetor escolhido foi o pLG2 cuja eficácia de transferência para linhagens de Hi e *N. meningitidis* já foi observada (Cury et al., 2014). Todavia, novamente Hae não foi competente ao processo de transformação. Em contrapartida, foram observadas 19 e 4 UFC de Hi RdKW20 e Hi β -lac, respectivamente.

Sabe-se que Hi é naturalmente competente ao processo de transformação e apresenta conjunto de genes relacionados ao processo de transformação (Sinha et al., 2012). E, como será discutido a seguir, através da análise dos genomas sequenciados aqui neste trabalho, observou-se que linhagens de Hae não apresentam todos os genes relacionados à competência de Hi identificados por Sinha et al. (2012). Tais dados são de grande valia já que podem explicar o porquê de Hae não ser competente ao processo de transformação, observação já realizada por outros estudos os quais também não apresentaram êxito em obter transformantes de Hae (Li et al., 2003, Serruto et al., 2009).

5.4. Busca por genes relacionados à competência de *Haemophilus* em Hae

Apesar de Hi ser uma bactéria naturalmente competente à transformação bacteriana, não foi possível transformar linhagens de Hae, associadas ou não a BPF. Assim, optou-se por buscar nos genomas de Hae os genes que fazem parte do *regulon* CRP-S de Hi (Sinha et al., 2012), envolvido no processo de competência de *Haemophilus* spp.

Como é possível observar na Tabela 9, três genes do *regulon* em questão estão ausentes em linhagens de Hae-BPF. O gene *HII631* está ausente em todas as linhagens encontradas, independentemente de ser associada ou não a BPF. Já os genes *HI0659* e *HI0660*, presentes em F3047, KC1018 e ATCC 11116, não fazem parte do genoma das demais linhagens de Hae. Tais dados estão em desacordo com Strouts et al. (2012), cujas análises *in silico* mostraram que Hae possui os genes e sítios regulatórios necessários para competência e transformação em Hi.

Não há função descrita e homólogos existentes para os três genes em questão, porém, sabe-se que estão relacionados à competência em Hi. Acredita-se que *HII631* e *HI0660* codificam produtos citoplasmáticos e que participam do processo de captação de DNA ou transformação de fato. Ambos os genes também estão ausentes em diversas espécies de *Pasteurellaceae*, incluindo outras linhagens de Hi (Sinha et al., 2012).

Já *HI0659* parece ter papel fundamental na competência de Hi, já que sua deleção em Hi eliminou a captação de DNA e transformação nestas linhagens (Sinha et al., 2012). Tal fato explicaria, em parte, a incapacidade de transformação de Hae-BPF.

A ausência desses três genes (ou de apenas *HII631*) poderia ser o motivo pelo qual Hae não é competente à transformação. Porém, como algumas linhagens de Hae (F3047, ATCC 11116 e KC1018) apresentam o gene *HI0659*, é possível que outros fatores (ou a ausência dos mesmos) impossibilitem a competência em Hae.

Tabela 9. Presença (+) ou ausência (-) de genes do *regulon* CRP-S de Hi em linhagens de Hae.

	Karina	219 /86	254 /86	284 /86	321 /86	406 /86	F3047	KC1018	ATCC 11116
<i>pilA</i>	+	+	+	+	+	+	+	+	+
<i>pilB</i>	+	+	+	+	+	+	+	+	+
<i>pilC</i>	+	+	+	+	+	+	+	+	+
<i>comA</i>	+	+	+	+	+	+	+	+	+
<i>comB</i>	+	+	+	+	+	+	+	+	+
<i>comC</i>	+	+	+	+	+	+	+	+	+
<i>comD</i>	+	+	+	+	+	+	+	+	+
<i>comE</i>	+	+	+	+	+	+	+	+	+
<i>comF</i>	+	+	+	+	+	+	+	+	+
<i>comN</i>	+	+	+	+	+	+	+	+	+
<i>comO</i>	+	+	+	+	+	+	+	+	+
<i>comP</i>	+	+	+	+	+	+	+	+	+
<i>comQ</i>	+	+	+	+	+	+	+	+	+
<i>comE1</i>	+	+	+	+	+	+	+	+	+
<i>HI0365</i>	+	+	+	+	+	+	+	+	+
<i>pilF2</i>	+	+	+	+	+	+	+	+	+
<i>rec2</i>	+	+	+	+	+	+	+	+	+
<i>ligA</i>	+	+	+	+	+	+	+	+	+
<i>HI0659</i>	-	-	-	-	-	-	+	+	+
<i>HI0660</i>	-	-	-	-	-	-	+	+	+
<i>HI1631</i>	-	-	-	-	-	-	-	-	-
<i>comM</i>	+	+	+	+	+	+	+	+	+
<i>dprA</i>	+	+	+	+	+	+	+	+	+
<i>radC</i>	+	+	+	+	+	+	+	+	+
<i>ssb</i>	+	+	+	+	+	+	+	+	+

6. Ensaio de infecção *in vivo*

Ratos da linhagem Sprague-Dawley de 8 - 10 dias de idade foram infectados com linhagens de Hae e Hi β -lac, a fim de isolar tais bactérias após a passagem em animal. A

tabela abaixo mostra o número de animais dos quais obteve-se êxito na obtenção de linhagens com passagem animal em relação ao número total de animais infectados por linhagem bacteriana.

Tabela 10. Relação entre o número de animais dos quais foram recuperadas colônias bacterianas e o número de animais infectados.

Tempo de infecção	254/86	258/86	KC01018	ATCC 11116	Hi β -lac
24 horas	2/5	2/5	0/5	0/3	1/5
48 horas	0/5	2/5	0/5	-	0/5

Como é possível observar, não foram recuperadas colônias de linhagens de Hae não-BPF. Tal dado corrobora com o estudo de Rubin & Carlone (1989), no qual, de apenas 1 de 50 animais, recuperou-se uma linhagem de Hae não causadora de BPF do sangue de Sprague-Dowley infectados por 24 horas.

Ainda, os dados apresentados sugerem que linhagens de Hae-BPF são capazes de causar uma bacteremia mais contundente em comparação a linhagens de Hae não-BPF e Hi, o que foi também já observado previamente por Rubin & Carlone (1989). Isso sugere que linhagens de Hae-BPF são mais virulentas e apresentam fatores específicos que auxiliam na sobrevivência no soro tanto de animais infantis quanto de crianças, já que é provável que o desenvolvimento de BPF em crianças seja principalmente consequência da virulência única destas linhagens do que anomalias imunes do paciente (Rubin & Carlone, 1989).

7. Expressão de genes em Hae

7.1. Seleção dos genes de referência e padronização dos ensaios de qRT-PCR

Um dos parâmetros mais importantes a serem definidos em uma reação de quantificação relativa por PCR em Tempo Real da expressão de genes de interesse é a seleção de genes de referência, já que estes normalizam a variabilidade resultante das

condições experimentais, preparação de amostras e até análise de resultados (Bustin et al., 2009). O gene de referência ideal deve ter expressão estável e similar em todas as amostras, independente da origem e/ou tratamento.

Assim, foram selecionados 12 genes como possíveis candidatos para gene de referência para Hae, com produtos de diferentes classes funcionais. Os *primers* foram testados por PCR convencional nas linhagens de Hae e Hi presentes na Tabela 1 para verificar a presença e amplificação dos genes em questão e a especificidade dos oligonucleotídeos.

Obteve-se êxito nas reações com todos os *primers* testados em todas as linhagens, exceto para o gene *rho*, cujos *primers* mostraram-se inespecíficos resultando em duas bandas visualizadas no gel de agarose e, portanto, dois *amplicons* distintos.

A estabilidade da expressão dos genes em questão sobre três condições experimentais distintas: após infecção *in vivo* e *in vitro* e estoque foram avaliadas, exceto para o gene *rho*, o qual foi descartado. Assim, obteve-se o coeficiente de estabilidade de expressão para cada gene nas amostras analisadas (Figura 16).

Quanto menor o valor do coeficiente, menor a variação da expressão do gene entre as amostras (Vandesompele et al., 2002). Com base nisso, foram selecionado os quatro genes com menor coeficiente: *gyrA* (0,2934), *primase* (0,4917), *rpoC* (0,3362) e *rpoD* (0,3086). O gráfico abaixo mostra os coeficientes dos quatro genes em questão.

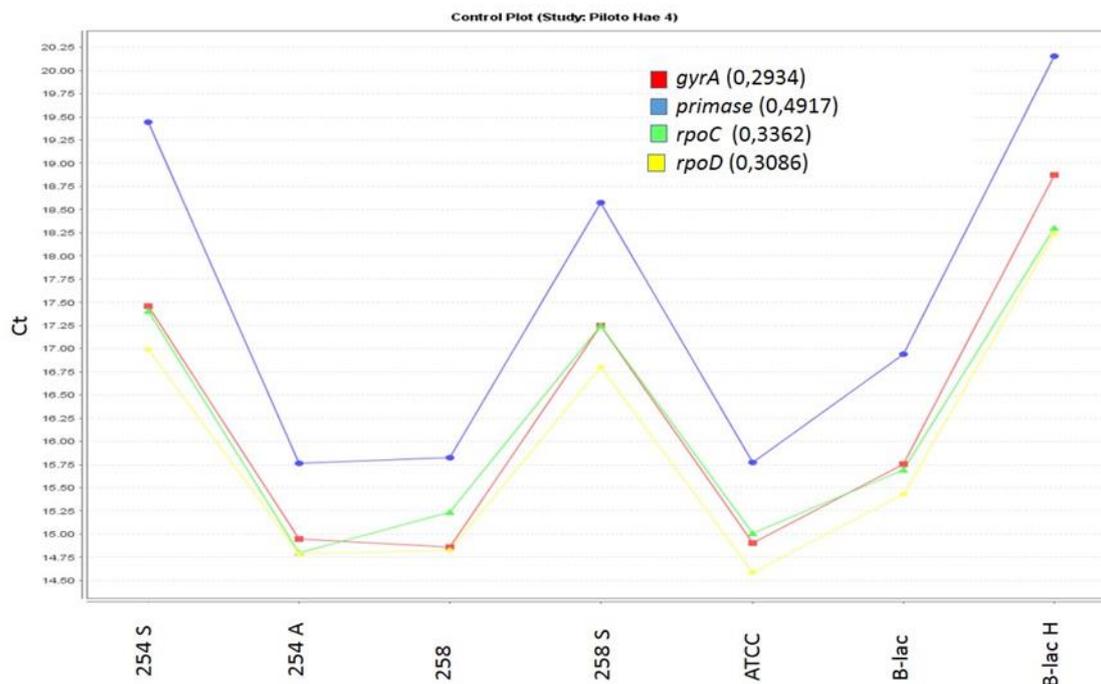


Figura 16. Variação da expressão dos genes candidatos à referência para Hae e Hi em diferentes linhagens e condições experimentais. 254 S (isolada após infecção *in vivo* por 24 horas), 254 A (isolada após infecção em células A549), 258 (estoque), 258 S (isolada após infecção *in vivo* por 24 horas), ATCC 11116 (estoque), β -lac (estoque) e β -lac H (isolada após infecção em células Hec-1B).

Baseando-se nos valores dos coeficientes mostrados no gráfico acima, foram selecionados os genes *gyrA*, *rpoD* e *rpoC* como genes de referência para os ensaios seguintes de quantificação da expressão gênica de Hae e Hi.

7.2. Quantificação da expressão de genes candidatos à virulência de Hae por qRT-PCR

Como o *knockout* dos genes *hadA*, *las* e *tabA* em linhagens de Hae não foi possível, optou-se por avaliar a expressão dos mesmos durante a infecção *in vitro* e *in vivo*, a fim de se analisar seu papel na virulência de Hae. Para tal, utilizou-se o método $\Delta\Delta Ct$ de expressão relativa qRT-PCR com os genes *gyrA*, *rpoD* e *rpoC* como genes de referência (como abordado no tópico anterior).

A maioria das linhagens de infecção *in vitro* e *in vivo* de Hae 254/86, 258/86, KC1018 e ATCC 1116 apresentaram expressão diferenciada dos genes *las* e *tabA* em

relação à amostra controle sem passagem animal ou em célula ($p < 0,001$), com exceção de casos como *lasHae* em 258/86 S, *tabA* em 258/86 A e em ATCC 11116 A (Figuras 17 - 20).

A expressão de *las* em 254/86 A, 254/86 S, 258/86 A, 258/86 R e 258/86 S foi significativamente menor em relação ao controle (*las* e *lasHae* nas Figuras 17 e 18), o que é um fato interessante já que se trata de amostras com passagem em células epiteliais (A549 e HaCaT) e em ratos Sprague-Dawley. Tal dado poderia sugerir que uma regulação negativa da expressão este transportador associado à virulência em Hae-BPF em contato com células epiteliais. Porém, mais dados são necessários para se afirmar isso.

Nas amostras não-BPF, foi observado uma menor expressão de *las* por amostras isoladas da infecção em A549 (KC1018 A e ATCC 11116 A, $p < 0,001$ em relação aos controles). Já nas demais amostras, isoladas da infecção em células Hec-1b e HaCaT, a expressão de *las* foi maior quando comparada aos respectivos controles ($p < 0,001$, Figuras 19 e 20).

Já a expressão de *tabA* por 254/86 A, H e S foi significativamente maior em relação à amostra controle (estoque) ($p < 0,001$). Acredita-se que TabA seja a proteína P145 (Strouts et al., 2012), envolvida na sobrevivência em hospedeiros (Rubin & Rizvi, 1991, Rubin, 1995). Tal fato seria uma provável explicação para o aumento da expressão de *tabA*, principalmente nas linhagens isoladas de animais (Figura 17). Porém, a análise da expressão de *tabA* em 258/86 revelou uma menor expressão desse gene em 258/86 R ($p < 0,001$) e uma expressão similar à amostra controle por 258/86 A e S.

Curiosamente, nas amostras não-BPF KC1018 A, KC1018 H e ATCC H, *tabA* teve uma expressão maior em relação às amostras controle ($p < 0,001$; Figuras 19 e 20). Sabe-se que nestas amostras *tabA* corresponde, na verdade, ao gene *tahA*, o qual acredita-se que apresente uma função diferente de *tabA* (Strouts et al., 2012). Porém, com base

nos dados aqui apresentados, é possível que tal gene também esteja envolvido na sobrevivência ao microambiente do hospedeiro.

O gene *hadA*, por sua vez, presente somente em linhagens associadas a BPF, apresentou comportamento inverso nas duas linhagens Hae-BPF analisadas (Figuras 17 e 18). Em 254/86 A, H e S, sua expressão foi maior em relação à linhagem controle (estoque), com destaque para 254/86 H ($p < 0,001$) (Figura 17). Estes dados corroboram com o observado por Serruto et al. (2009) que relatou o papel de HadA na adesão e invasão celular através da expressão heteróloga em *E. coli*. Entretanto, em 258/86 A, R e S, este gene teve menor expressão em relação ao controle, com destaque para 258/86 R ($p < 0,001$) (Figura 18). É importante ressaltar que não foi possível obter a expressão relativa de *hadAHae* para 258/86 A e S e de nenhum gene para 258/86 H.

Apesar de significativamente diferentes, a expressão dos três genes em relação aos controles foi maior em algumas casos e, em outros, menor, até mesmo entre diferentes passagens (A, H, R e S) de uma mesma linhagem bacteriana. Porém, de maneira geral, os dados sugerem que *las* seja regulado negativamente quando linhagens causadoras de BPF estão em contato com o hospedeiro. Já os genes *hadA* e *tabA*, específicos de Hae-BPF, parecem exercer um papel importante em infecções *in vitro* e *in vivo* por Hae 254/86. É possível que as diferenças de expressão relativa desses dois genes para as linhagens 254/86 e 258/86 sejam resultado de pequenas diferenças entre os genomas, já que os *primers* para as reações de qRT-PCR foram obtidos a partir do genoma de 254/86.

Mais estudos serão necessários para se determinar o papel desses genes na virulência de Hae-BPF, tanto em células humanas quanto em modelos animais. Já que o *knockout* gênico não foi possível nestas linhagens, a expressão heteróloga dos produtos codificados por *las*, *hadA* e *tabA* em outras bactérias, tal como foi feito em *E. coli* com HadA (Serruto et al., 2009) pode ser uma ferramenta útil no estudo da virulência de

linhagens causadoras de BPF. Entretanto, é provável que a virulência desse grupo de bactérias seja resultado de um conjunto de fatores, e não apenas um, assim, análises de transcriptoma ou RNASeq (baseados nos genomas disponíveis) seriam de grande valia na identificação de fatores únicos de Hae-BPF.

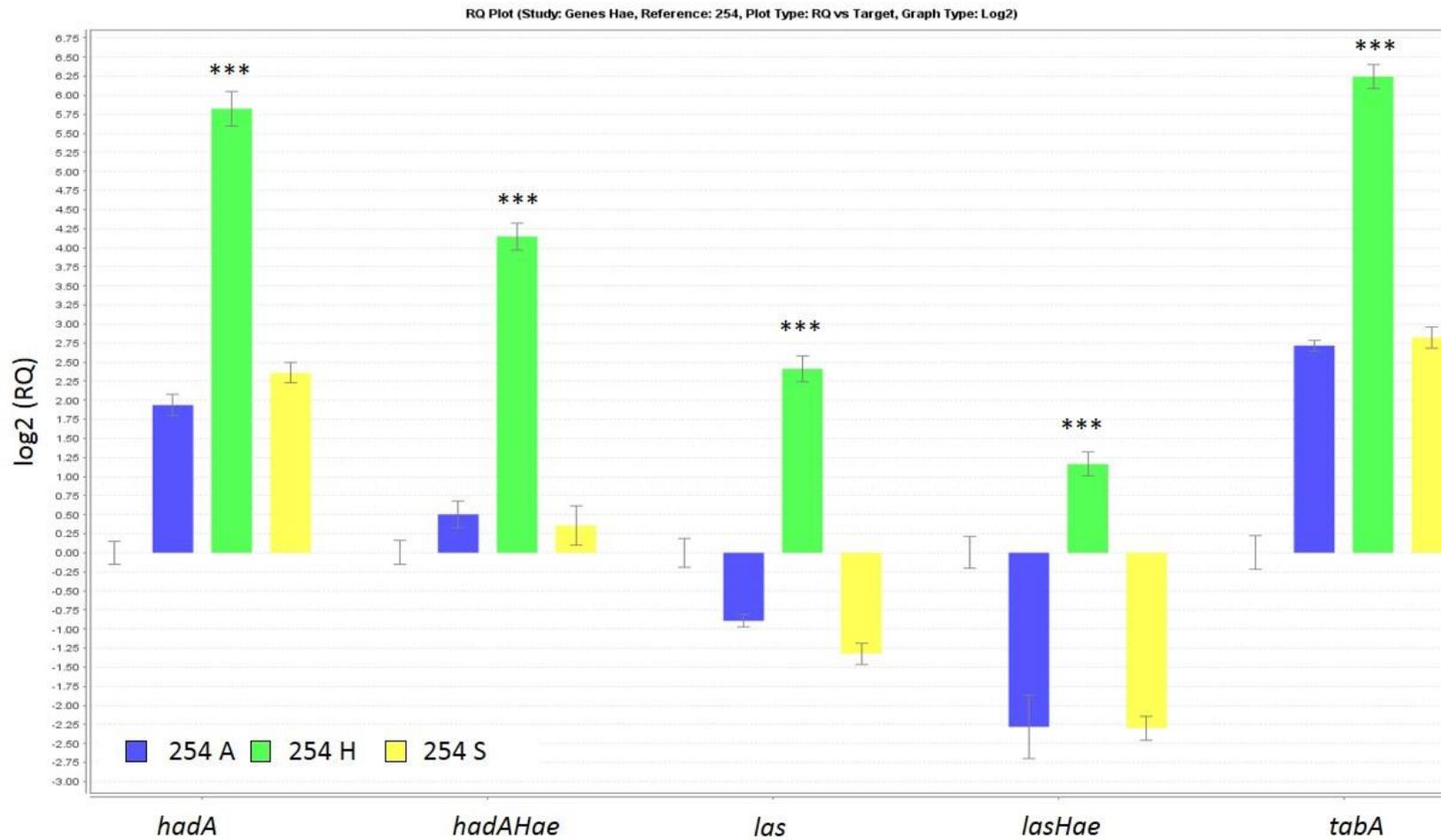


Figura 17. Expressão relativa de *hadA*, *las* e *tabA* em Hae 254/86. *gyrA*, *rpoC* e *rpoD* foram usados como genes de referência para normalização do dados e a amostra controle como calibradora (linhagem bacteriana sem passagem em célula ou animal). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por *one-way* ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação as demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

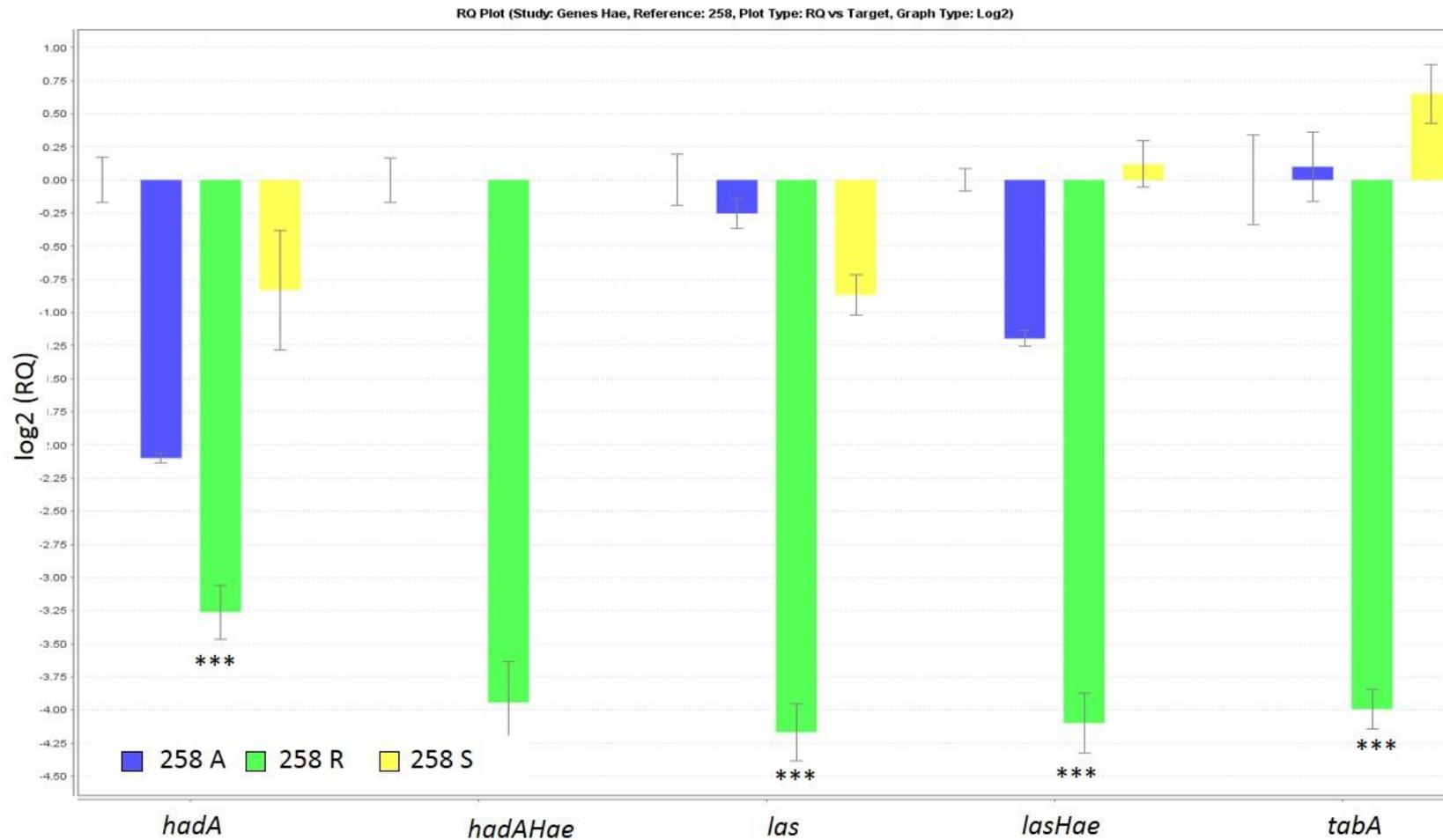


Figura 18. Expressão relativa de *hadA*, *las* e *tabA* em Hae 258/86. *gyrA*, *rpoC* e *rpoD* foram usados como genes de referência para normalização do dados e a amostra controle como calibradora (linhagem bacteriana sem passagem em célula ou animal). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por *one-way* ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação as demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

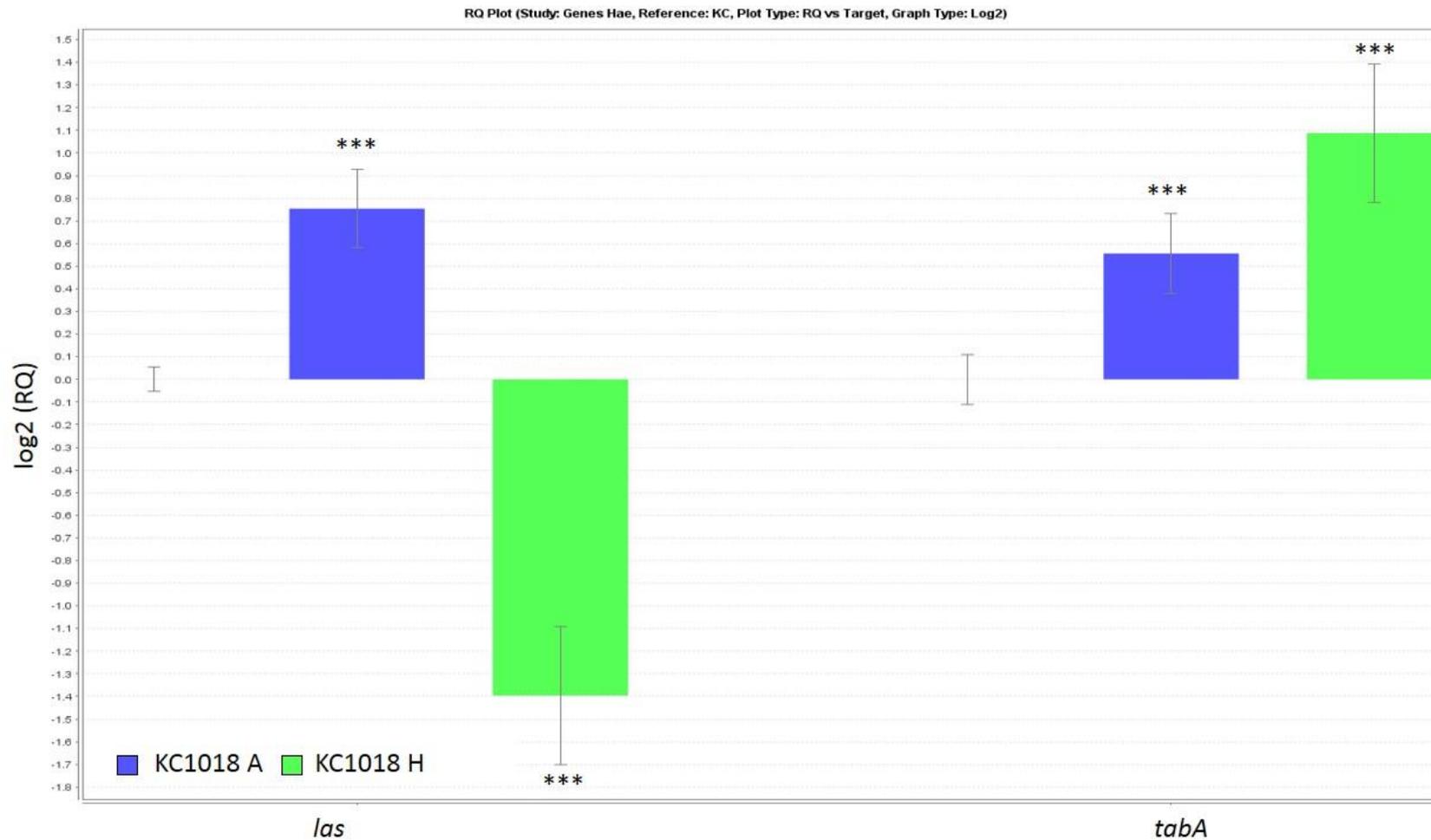


Figura 19. Expressão relativa de *las* e *tabA* em Hae KC1018. *gyrA*, *rpoC* e *rpoD* foram usados como genes de referência para normalização dos dados e a amostra controle como calibradora (linhagem bacteriana sem passagem em célula ou animal). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por *one-way* ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação as demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

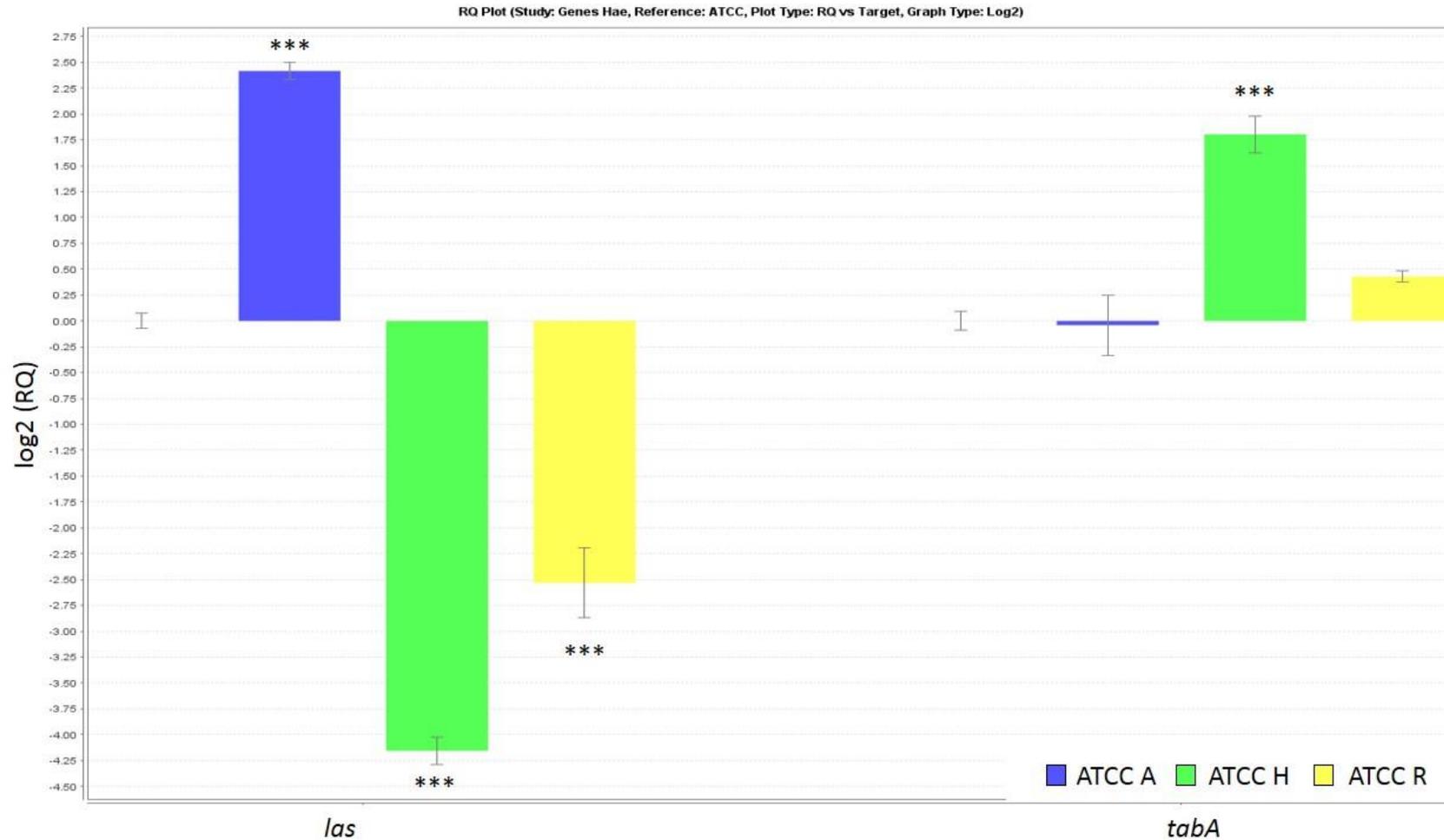


Figura 20. Expressão relativa de *las* e *tahA* em Hae ATCC 11116. *gyrA*, *rpoC* e *rpoD* foram usados como genes de referência para normalização do dados e a amostra controle como calibradora (linhagem bacteriana sem passagem em célula ou animal). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por *one-way* ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação as demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

8. Quantificação de citocinas inflamatórias por qRT-PCR

Avaliou-se a expressão de citocinas por qRT-PCR em células de origem pulmonar (A549) e epitelial (HaCaT) infectadas por linhagens de Hae e Hi β -lac. Tais linhagens celulares foram escolhidas devido à fisiopatologia de BPF e com o objetivo de analisar a resposta do hospedeiro *in vitro* e *in vivo*.

Os dados de qRT-PCR mostram que a expressão de IL-10 e TNF- α por A549 quando infectada por 254/86 é maior em relação às demais linhagens de Hae ($p < 0,001$). O perfil de expressão dessas citocinas na infecção por 258/86 e Karina foram similares. Porém, a expressão de IL-10 nestas células infectadas por KC1018 foi significativamente menor ($p < 0,001$) (Figura 21).

Células A549 quando infectadas por linhagens NTHi apresentam uma maior expressão de CXCL-8 e ICAM-1 a partir de 6 horas (Geelen et al., 2014), 12 e 24 horas de infecção. Níveis significativos de TNF- α não foram detectados em infecções por 12 e 24 horas (Frick et al., 2010). Porém, sabe-se que dados de DNA *microarray* sobre a resposta inflamatória a Hi causadores de otite média em modelo animal, moléculas inflamatórias como TNF- α e IL10 atuam no período inicial de infecção de 3 a 6 horas (Hernandez et al., 2015). Isso explicaria a discrepância entre os dados aqui apresentados e os de Frick e colaboradores (2010) para a mesma célula A549, já que foram infectadas por um período de 6 horas por linhagens de Hae.

Acredita-se que TNF- α seja uma das principais citocinas estimulantes da expressão de ICAM-1 por células epiteliais pulmonares (Look et al., 1992, Tosi et al., 1992, Bloemen et al., 1993). Assim, é possível que o aumento da expressão de TNF- α em A549 infectada por Hae esteja vinculada a um aumento da expressão de ICAM-1, como um mecanismo de defesa das células hospedeiras contra o patógeno.

Aparentemente, a regulação da inflamação de células pulmonares está vinculada

a diversos fatores de causa e efeito que hiperativam o sistema imune inato do hospedeiro, enquanto patógenos, tais como *Haemophilus*, podem desencadear a resposta inflamatória, a presença de outros indutores, como quimiocinas, fatores de crescimento e outros microrganismos, podem sinergicamente aumentar tal resposta e, conseqüentemente, causar danos ao hospedeiro (Wang et al., 2012). O quadro de hemorragia e necrose pulmonar da BPF pode ser resultado dessa inflamação exacerbada de células pulmonares em contato com Hae.

Já, a expressão relativa de IL-10 e TNF- α em HaCaT infectadas por Hae 254/86, 258/86 foram semelhantes ($p > 0,05$). Houve uma redução da expressão de IL-10 e TNF- α em comparação ao controle, exceto a expressão dessas duas citocinas na infecção por ATCC 1116 (Figura 22).

Os dados obtidos para HaCaT infectadas por Hae-BPF são semelhantes ao de Zaretzky & Kawula (1999), os quais não observaram a produção de TNF- α neste tipo celular após a infecção por *Haemophilus ducreyi* nos tempos de 2, 4, 8, 24, 48 e 72 horas. Desse modo, sugere-se que IL-10 e TNF- α não exerçam um papel principal na resposta inflamatória por HaCaT tanto a *H. ducreyi* quanto a Hae.

A expressão de IL-10 no sangue total de animais infectados por β -lac foi significativamente maior ($p < 0,001$) quando comparada às linhagens de Hae, causadoras ou não de BPF. Não foram observadas diferenças quantitativas na expressão de IL-10 em animais infectados por 24 ou 48 horas (Figura 23). Não foi possível determinar a expressão de TNF- α nos ensaios *in vivo*. Isso pode ser explicado pelo fato de se tratar de moléculas que atuam no período inicial de infecção (3 - 6 horas) (Hernandez et al., 2015).

De maneira geral, foi observada uma menor expressão de IL-10 nas infecções em células HaCaT e em animais, enquanto que em células A549, houve um aumento de IL-10 e TNF- α . Contudo, é importante ressaltar que para se obter uma análise mais precisa,

seria necessário o uso de outra técnica capaz de avaliar sua concentração protéica, tal como *Ezyme Linked Immuno Sorbent Assay* (ELISA).

Os dados aqui apresentados são de grande importância, inéditos e possibilitam o desenvolvimento de uma nova área de estudo sobre Hae-BPF, já que não foram encontrados na literatura dados similares da resposta inflamatória *in vitro* e *in vivo* da BPF. Não obstante, os dados da literatura se referem à infecção por *Haemophilus* não pertencentes ao grupo *aegyptius*. Como o perfil de citocinas pode variar de acordo com a linhagem celular e o patógeno, é possível que existam diferenças nas respostas inflamatórias entre linhagens Hae e linhagens Hi.

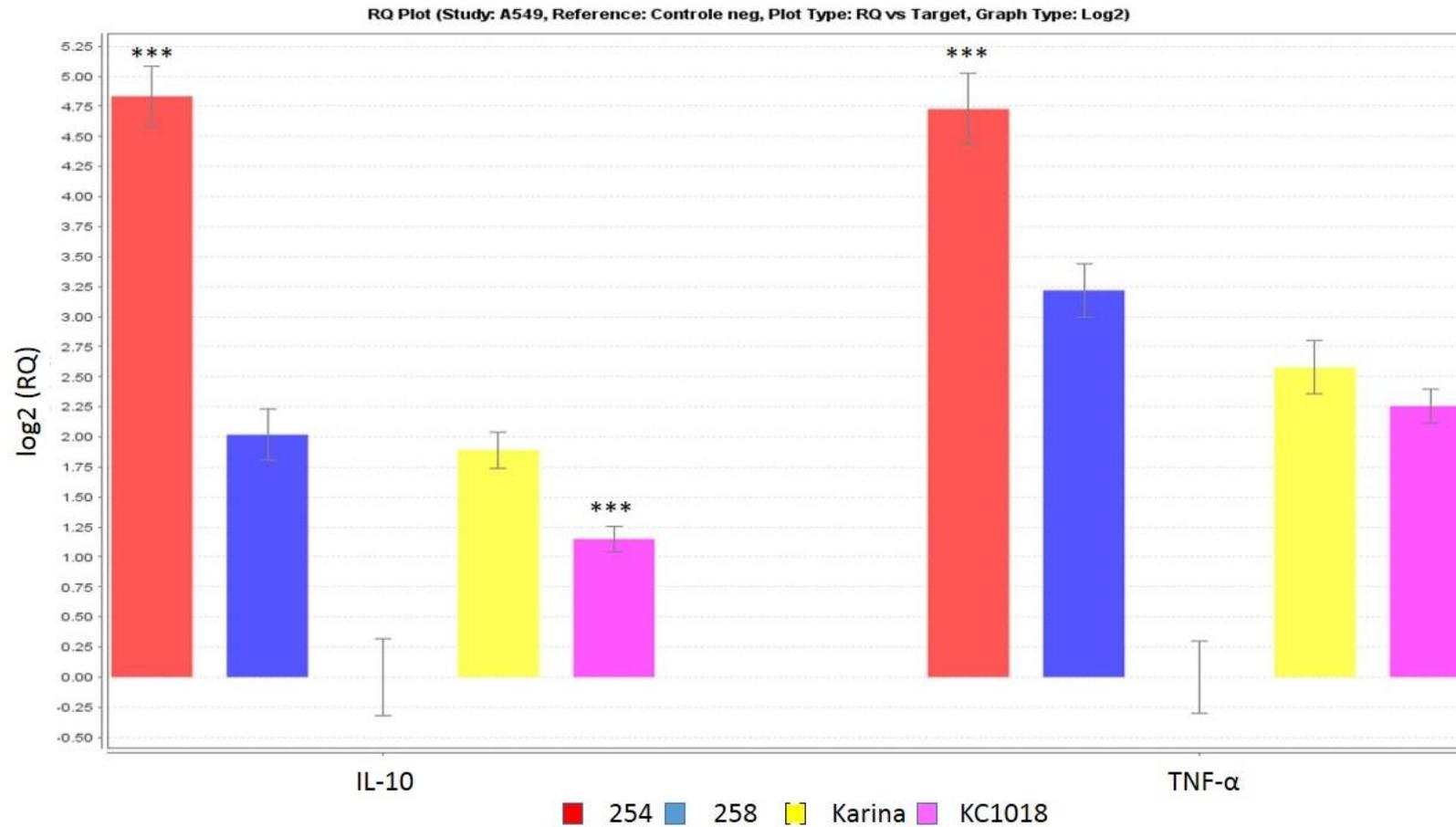


Figura 21. Expressão relativa de IL-10 e TNF- α em células A549 infectadas por Hae 254/86, 258/86, Karina e KC1018. *GAPDH* foi usado como gene de referência para normalização dos dados e a amostra controle como calibradora (células não infectadas). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por *one-way* ANOVA com as amostras comparadas entre si com o método de Tukey. *** $p < 0,001$ em relação às demais amostras para cada gene. As barras indicam o desvio padrão das amostras.

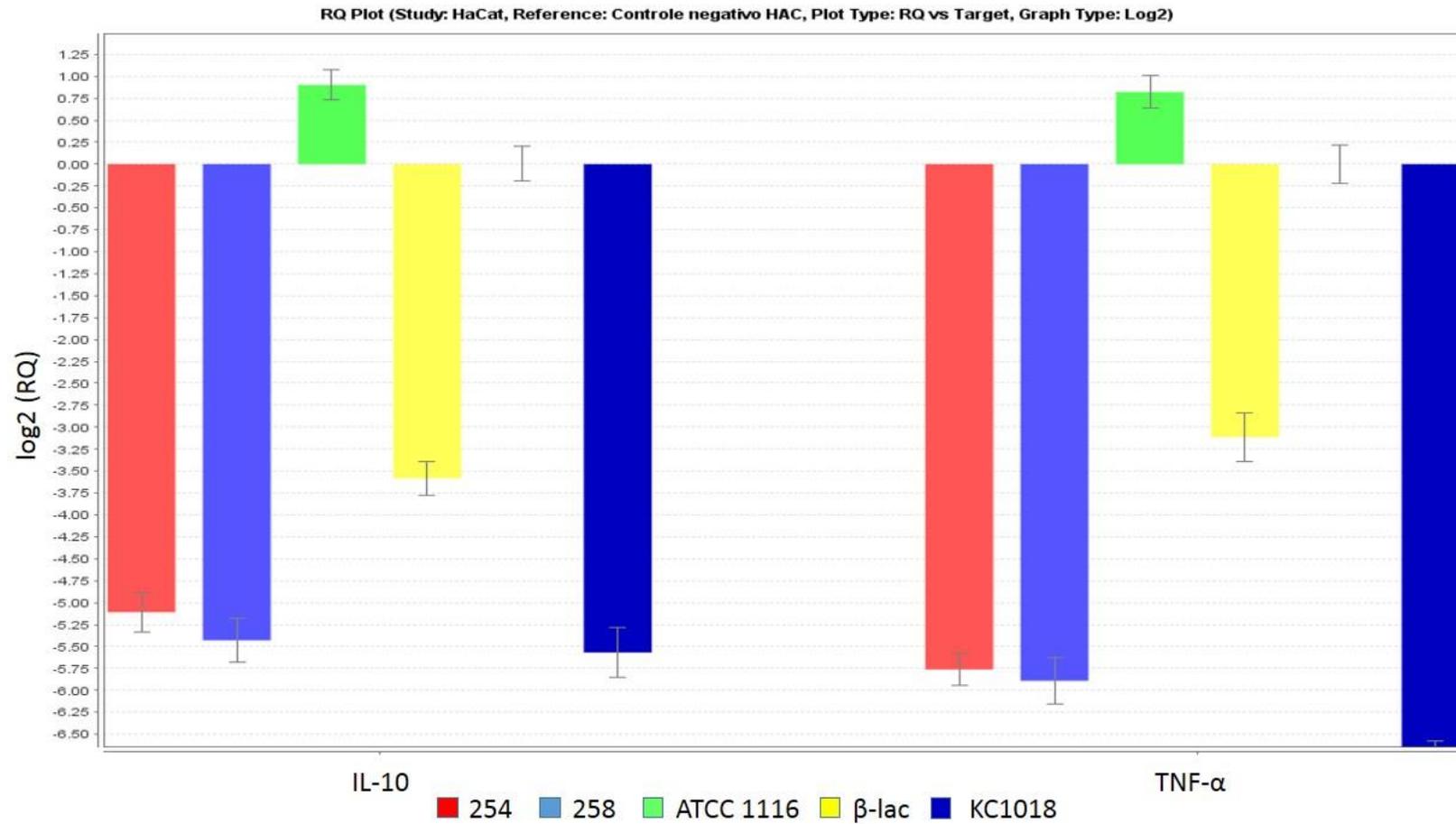


Figura 22. Expressão relativa de IL-10 e TNF- α em células HaCaT infectadas por Hae 254/86, 258/86, KC1018, ATCC 11116 e Hi β -lac. *GAPDH* foi usado como gene de referência para normalização do dados e a amostra controle como calibradora (células não infectadas). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por *one-way* ANOVA com as amostras comparadas entre si com pelo método de Tukey. $p < 0,05$. As barras indicam o desvio padrão das amostras.

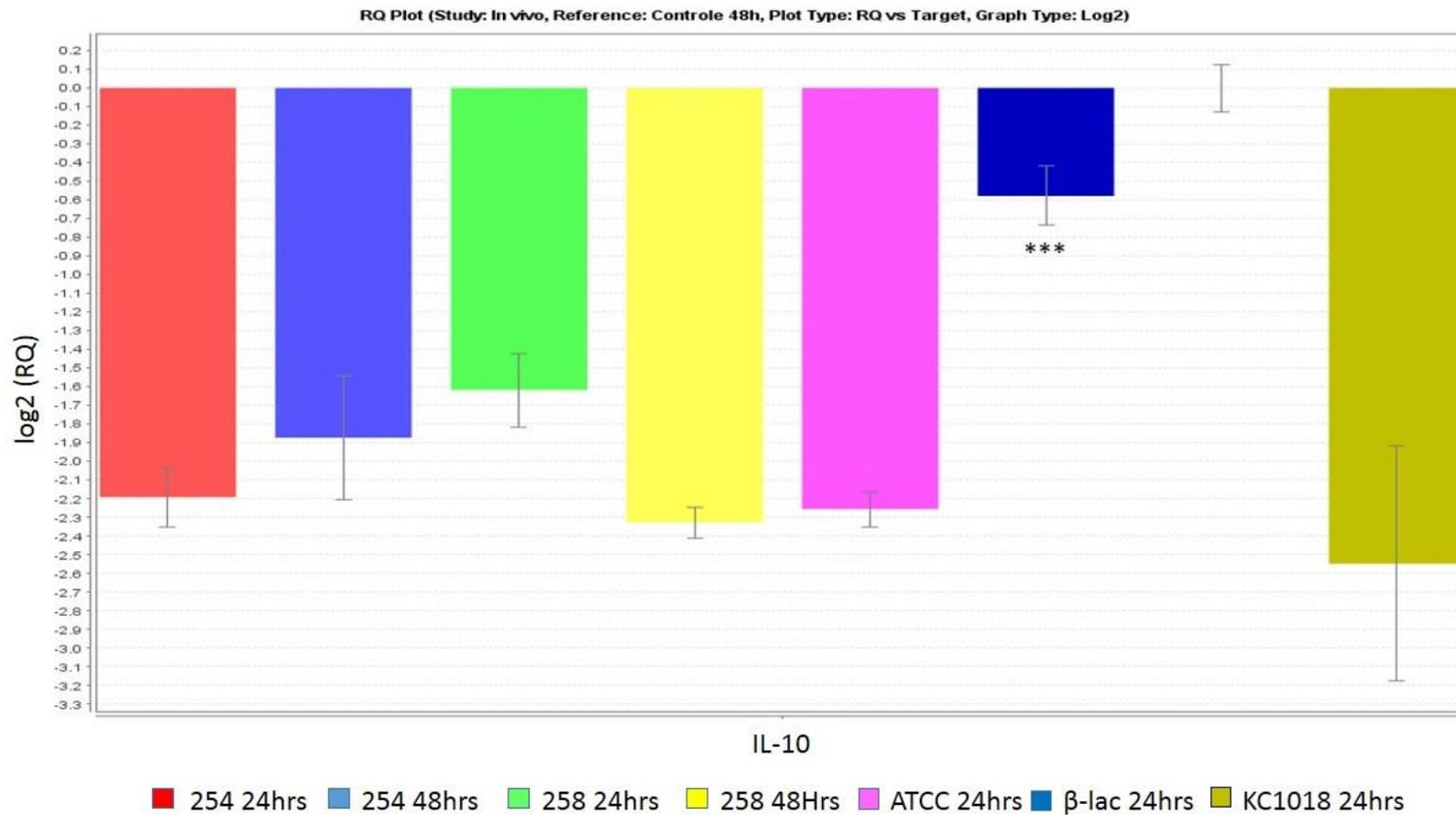


Figura 23. Expressão relativa de IL-10 em ratos Sprague-Dawley infectados por Hae 254/86, 258/86, KC1018, ATCC 11116 e Hi β-lac. *GAPDH* foi usado como gene de referência para normalização do dados e a amostra controle como calibradora (animais não infectados). Os ensaios foram feitos em triplicatas para cada amostra e a análise estatística foi feita por *one-way* ANOVA com as amostras comparadas entre si com pelo método de Tukey. *** $p < 0,001$ em relação as demais amostras. As barras indicam o desvio padrão das amostras.

9. Análise dos lipídios de Hae por Espectrometria de Massas

Inicialmente, na tentativa de identificar lipídios presentes nas amostras de Hae, buscou-se por aqueles que possuíam padrões comerciais disponíveis, mediante análise por *fingerprint* e cromatografia líquida de ultra-eficiência. Foi observado que nenhum dos padrões testados (cardiolipina, monosialotetrahexosilgangliosido, fosfatidilcolina, fosfatidiletanolamina, fosfatidilinositol, fosfatidilserina e esfingomiéline) se encontra nas amostras. Entretanto, é possível que tais lipídios estejam presentes em Hae, porém com estruturas diferentes a dos padrões analisados, como por exemplo, cadeias laterais distintas, e, assim, íons e espectros MS/MS também distintos aos dos padrões.

Com o intuito de caracterizar lipídios das linhagens Hae 254/86, Hae 258/86, Hae KC1018, Hae ATCC 11116 e Hi β -lac submetidas a diferentes tratamentos (estoque, infecção *in vivo* e *in vitro*) e relacioná-los à virulência de Hae, realizou-se uma análise por espectrometria de massas com inserção direta com ESI nos modos positivo e negativo. A faixa de massas adquirida foi de 100 - 2000 m/z, embora não foram observados sinais a partir de m/z 700.

Nas análises de *fingerprint* nas amostras de Hae, vários íons em comum foram encontrados em mais de uma amostra. Através de seus perfis de fragmentação por ESI-MS/MS, confirmou-se que eram de fato os mesmos constituintes presentes nas amostras.

Realizou-se posteriormente uma análise multivariada das amostras por PCA e HCA, a fim de se agrupar as amostras pelos íons que possuem em comum e identificar íons marcadores de cada tipo de amostra. Os dados indicam que as amostras apresentam muitos íons em comum, o que impossibilitou a identificação de íons específicos para Hae-BPF em comparação a Hae não-BPF e Hi, independentemente de se tratar de uma amostra estoque ou isolada de células ou animais. Assim, cada linhagem bacteriana foi analisada separadamente.

Apesar das semelhanças entre as amostras, a análise por HCA dos dados por ESI(+) de cada linhagem permitiu agrupá-las de acordo com sua origem: estoque, infecção em células A549 (amostras A), infecção em células Hec-1b (amostras H), infecção em células HaCaT (amostras R) e infecção em ratos Sprague-Dawley (amostras S). (Figuras 24 e 25). Ainda, no caso das linhagens Hae-BPF, pode-se observar a formação de 2 grandes *clusters*: um composto pelas amostras de origem animal (S) e outro, pelas demais amostras.

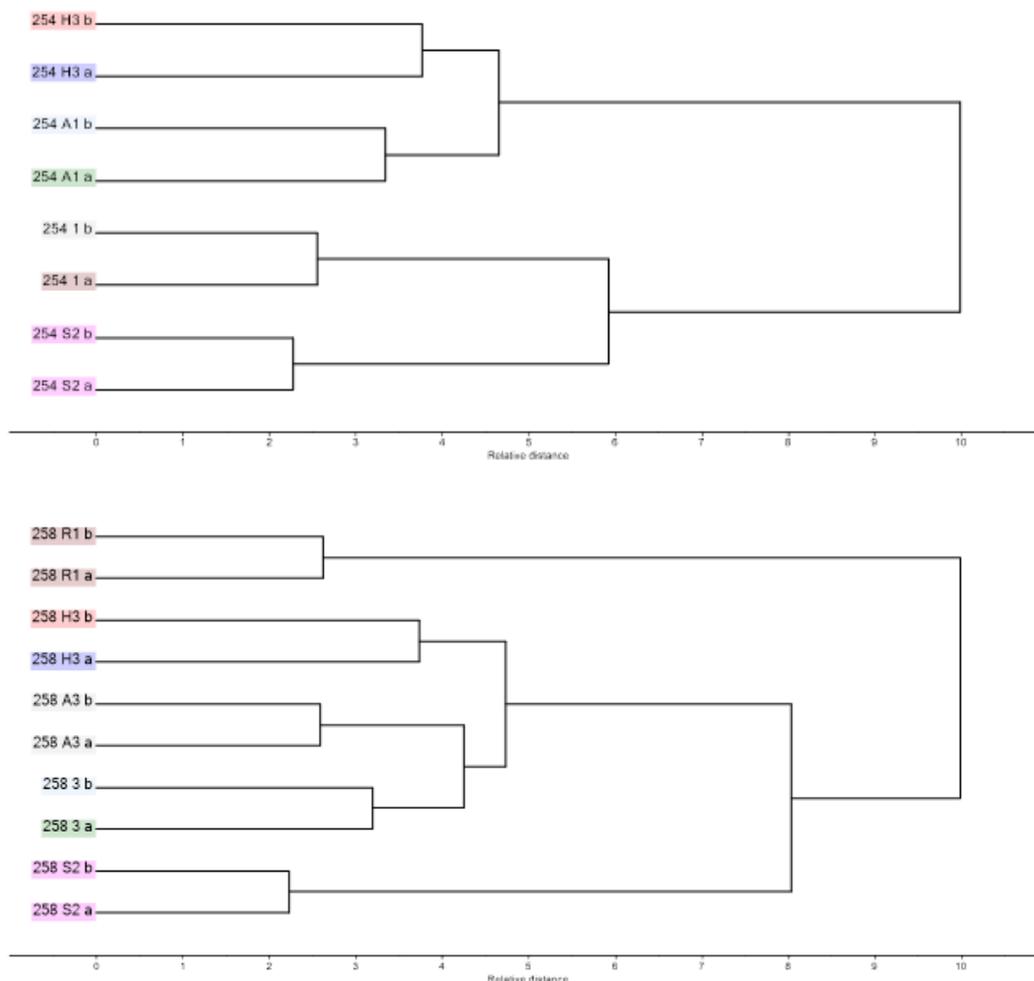


Figura 24. HCA dos dados de ESI(+)MS de Hae-BPF 254/86 e 258/86 segundo o método de Ward e distância Euclidiana ao quadrado. Pode-se observar o agrupamento das amostras em duplicata do estoque, da infecção em A549 (amostras A), da infecção em Hec-1b (amostras H), da infecção em HaCaT (amostras R) e da infecção em Sprague-Dawley (amostras S).

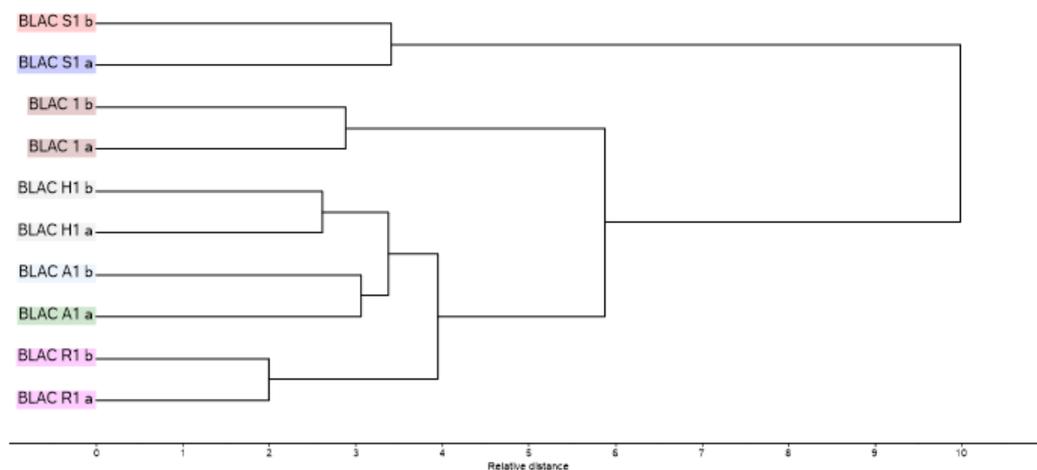
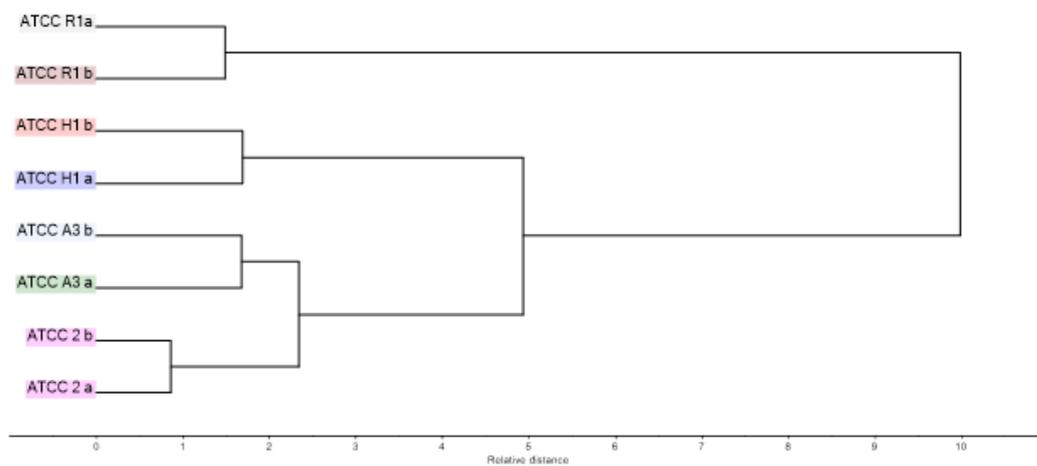
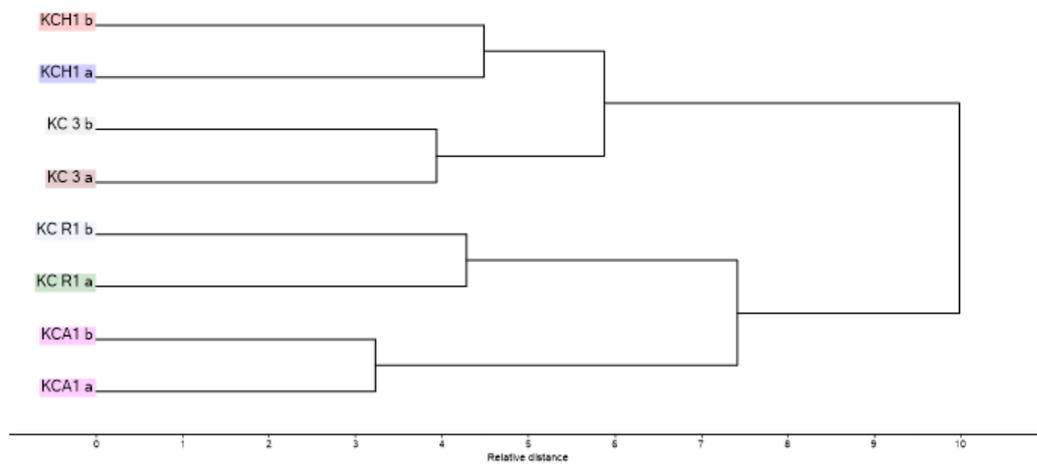


Figura 25. HCA dos dados de ESI(+)^{MS} de Hae KC1018 e ATCC 11116 e Hi β -lac segundo o método de Ward e distância Euclidiana ao quadrado. Pode-se observar o agrupamento das amostras em duplicata do estoque, da infecção em A549 (amostras A), da infecção em Hec-1b (amostras H) e da infecção em HaCaT (amostras R) e, no caso de Hi β -lac, também da infecção em Sprague-Dawley (amostras S).

A fim de agrupar as amostras e determinar os íons que influenciam nesse agrupamento, realizou-se um estudo de componentes principais (PCA) avaliando os *scores* e os *loadings* para cada linhagem.

Os *fingerprints* no modo positivo mostram diferenças qualitativas e quantitativas dos íons observados nas condições descritas na Tabela 2 em Materiais e Métodos. Nos gráficos dos *loadings* do PCA, que mostra os pesos das variáveis, é possível observar que as amostras possuem muitos íons em comum, o que dificulta identificar aqueles que mais influenciam no agrupamento das amostras no gráfico dos *scores* (Figuras 26 - 35 em Anexos).

No PCA das amostras de 254/86, pode ser observado que a amostra S (proveniente de infecção animal), destaca-se da demais e possui uma grande quantidade de íons diferentes (Figura 26 em Anexos). Nos *fingerprints* destas amostras, pode ser observado que o íon mais intenso na amostra S, é o m/z 239,8, que embora esteja presente nas outras amostras, a sua intensidade é bem menor. O íon m/z 292,7 é característico da amostra sem passagem em animal ou em células (linhagem estoque) e não está presente no resto das amostras (Figura 27 em Anexos). É possível que esse fosfolípido sofra alguma modificação após a bactéria entrar em contato com o hospedeiro. Os *fingerprints* das demais amostras 254/86 são muito semelhantes.

Nas amostras de 258/86, observa-se um comportamento semelhante, no qual a amostra S (animal) destaca-se das outras. Porém, neste caso, não foi possível identificar algum íon responsável por esta separação. Ainda, os *fingerprints* destas amostras, são de fato muito semelhantes, e a distribuição dos íons no gráfico dos *loadings* (Figura 28 em Anexos) sugere que as amostras possuem uma composição semelhante, inclusive a amostra S, cujas diferenças possivelmente sejam mais quantitativas. Ainda, o íon m/z 409,9 é menos intenso nas amostras S e R, e se observa que a amostra R possui um íon

m/z 602,5 e outro m/z 620,3 são característicos desta amostra (Figura 29 em Anexos).

O fato das amostras isoladas de animais (S de 254/86 e 258/86) destacarem-se das demais nas análises dos íons por HCA e PCA (Figuras 24, 26 e 28) sugerem que, após a passagem *in vivo*, Hae sofre alterações em sua composição lipídica, ou dos lipídios das membranas interna e externa ou até da porção lipídica do LOS. Rubin & St Geme (1993) observaram que amostras de Hae recuperadas de animais são virulentas e apresentam alterações no perfil de LOS. Assim, é possível que tal fenótipo esteja relacionado com a variação do perfil de lipídios desta bactéria, como no caso de *M. tuberculosis* que torna-se mais virulenta nesta situação (Khuller et al., 1982).

Já no caso de Hae KC1018, observa-se que as amostras são muito semelhantes entre si, com todos os íons em comum. Assim, a separação das amostras se dá através das intensidades dos picos dos *fingerprints*, o que caracteriza um agrupamento por diferenças quantitativas e não qualitativas (Figuras 30 e 31 em Anexos).

Nas amostras da linhagem ATCC 11116, observa-se que as amostras R são diferentes das outras e o íon m/z 438,0 é o maioritário no *fingerprint* desta amostra, apesar desta molécula também ser presente nas demais amostras, porém com uma intensidade menor. Além disso, os íons m/z 174,6 e m/z 183,6 são menos intensos na amostra R do que nas amostras A, H e na amostra sem passagem em célula (Figura 32 e 33 em Anexos).

Nas amostras de Hi β -lac, por seu turno, a amostra H aparenta ser a mais diferente das demais (Figura 34 em Anexos). No *fingerprint* desta amostra, observa-se um íon mais intenso m/z 239,8, presente nas outras amostras mas com uma intensidade muito menor. Os íons maioritários nas outras amostras m/z 174,6 e m/z 183,6 são menos intensos na amostra H (Figura 35 em Anexos).

Foi observado que todas as linhagens, independente da origem (estoque original,

infecção *in vitro* ou *in vivo*), possuem alguns íons em comum: m/z 141,6; m/z 174,6; m/z 183,6; m/z 199,6; m/z 227,7; m/z 239,7 e m/z 409,9, cujas intensidades mudam com os tratamentos (Figura 27, 29, 31, 33 e 35 em Anexos).

Já nos *fingerprints* no modo negativo, não foram advertidas diferenças qualitativas dos íons observados no espectro, apenas diferenças quantitativas. Dessa forma, os respectivos dados de por *fingerprint* e as análises HCA e PCA estão anexados no final do texto (Figuras 36 - 47 em Anexos).

10. Considerações finais

Acredita-se que BPF seja uma doença erradicada. Porém, pouco se sabe sobre os mecanismos de virulência de Hae que permitiram que este evoluísse de uma bactéria causadora de conjuntivite para um patógeno extremamente invasivo.

Os resultados apresentados neste trabalho, resumidos na tabela a seguir, indicam que as linhagens de Hae-BPF são semelhantes entre si quanto ao seu genomas e apresentam características específicas entre as demais linhagens de Hae estudadas. Contudo, tais linhagens de Hae não apresentam comportamento clonal como descrito anteriormente na literatura. Como Hae-BPF é um exemplo notável da evolução de um patógeno humano, aspectos sobre seu genoma, competência à transformação e alterações após o contato com hospedeiro foram abordados de modo a enriquecer o conhecimento, embasar e incentivar futuros estudos sobre este patógeno único, relativamente “novo” e negligenciado. Como proposta, sugere-se a expressão das proteínas Las, TabA e CrgA de Hae, objetivando aprofundar o papel destes no contato com o hospedeiro. Ainda, ensaios para transcriptoma e RNASeq tanto de genes de Hae quanto do hospedeiro seriam interessantes no estudo da virulência do patógeno e sua fisiopatologia na BPF.

Tabela 11. Resumo dos principais resultados obtidos neste trabalho.

	Karina	219/86	254/86	258/86	284/86	321/86	406/86	KC1018	ATCC 11116	Rd KW20	Hi β -lac
Genomas											
Número de <i>contigs</i>	53	34	-	39	35	39	89	91	-	-	-
Tamanho do genoma montado (Mb)	1.95	1.96	1.99	1.83	1.97	1.96	1.96	1.91	1.96	1.83	-
Comparação pan-genômica (maior semelhança com)	219/86, 254/86 e 406/86	Karina, 254/86, 321/86 e 406/86	Karina, 219/86, 321/86, 406/86 e F3047	RdKW20, R2846 e 10810	321/86 e 406/86	219/86, 254/86 e 406/86	Karina, 219/86, 254/86 e 321/86	Nenhuma	-	258/86, R2846 e 10810	-
Presença de <i>crgA</i>	Sim	Sim	Sim	-	Sim	Sim	Sim	Não	Não	Não	-
Presença de <i>mdaB</i>	Sim	Sim	Sim	-	Sim	Sim	Sim	Sim	Sim	Sim	-
Tamanho do plasmídio fechado (pb)	31849	31889	32433	32625	31903	31897	31886	-	-	-	-
Transformação e Competência											
<i>Knockout gênico de las, hadA e tabA</i>	Não	-	Não	Não	-	-	-	-	-	-	-
Transformação com pLG2 (número de UFC)	-	-	0	0	-	-	-	0	0	19	4
Presença de todos os genes relacionados à competência de Hi	Não (18/21)	Não (18/21)	Não (18/21)	-	Não (18/21)	Não (18/21)	Não (18/21)	Não (19/21)	Não (19/21)	Sim (21/21)	-
Infecção <i>in vivo</i>											
Recuperação de bactérias após passagem em animal	-	-	Sim	Sim	-	-	-	Não	Não	-	Sim

Expressão de genes por qRT-PCR**

Expressão de <i>las</i>	-	-	254/86 A (-)	258/86 A (-)	-	-	-	KC A (-)	ATCC A (-)	-	-
			254/86 S (-)	258/86 R (-)				KC H (-)	ATCC H (+)		
				258/86 S (-)					ATCC R (+)		
Expressão de <i>hadA</i>	-	-	254/86 A (+)	258/86 A (-)	-	-	-	-	-	-	-
			254/86 H (+)	258/86 R (-)							
			254/86 S (+)	258/86 S (-)							
Expressão de <i>tabA</i>	-	-	254/86 A (+)	258/86 R (-)	-	-	-	KC A (+)	ATCC H (+)	-	-
			254/86 H (+)					KC H (+)			
			254/86 S (+)								

Lipídios por ESI(+)-MS

Linhagens que se destacaram nas análises dos lipídios	-	-	254/86 S	258/86 S	-	-	-	Nenhuma	ATCC R	-	Hi β -lac H
---	---	---	----------	----------	---	---	---	---------	--------	---	----------------------

** Linhagens que tiveram expressão significativamente maior (+) ou menor (-) em relação à amostra controle ($p < 0,001$).

CONCLUSÕES

Os resultados apresentados sugerem que linhagens Hae são semelhantes entre si quanto ao seu genoma, porém não clonais. Ainda, Hae-BPF apresenta fatores específicos, ausentes em linhagens não causadoras da doença, os quais possivelmente estão correlacionados à virulência deste patógeno, sendo candidatos promissores para estudos futuros sobre BPF.

Com base nos dados, pode-se concluir que:

- Hae-BPF apresenta os genes *crgA*, *hadA* e *tabA*, ausentes em linhagens Hae não invasivas analisadas.
- Hae, independente de causar BPF ou não, não foi competente ao processo de transformação e não apresenta todos os genes relacionados à competência do gênero *Haemophilus*.
- Hae-BPF apresenta determinantes específicos de virulência que auxiliam na sobrevivência *in vivo*, pois não foram recuperadas linhagens de Hae não causadoras de BPF em animais infectados.
- Os genes *gyrA*, *rpoD* e *rpoC* apresentaram expressão mais estável em relação aos demais genes analisados em Hae e Hi.
- Hae apresentou menor expressão do gene *las* em contato com o hospedeiro *in vitro* e *in vivo*.
- O contato com o hospedeiro animal induz alterações no perfil de lipídios de Hae-BPF.

REFERÊNCIAS

- Albritton, W.L., J.K. Setlow, M. Thomas, F. Sottnek & A.G. Steigerwalt, (1984) Heterospecific transformation in the genus *Haemophilus*. *Molecular & general genetics* : MGG 193: 358-363.
- Altschul, S.F., W. Gish, W. Miller, E.W. Myers & D.J. Lipman, (1990) Basic local alignment search tool. *Journal of molecular biology* 215: 403-410.
- Ausubel, F.M. & I. Wiley, (2001) *Current protocols in molecular biology*. In. New York: J. Wiley, pp.
- Aziz, R.K., D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, F. Meyer, G.J. Olsen, R. Olson, A.L. Osterman, R.A. Overbeek, L.K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G.D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke & O. Zagnitko, (2008) The RAST Server: rapid annotations using subsystems technology. *BMC genomics* 9: 75.
- Bartlett, J.G., (2006) Narrative review: the new epidemic of *Clostridium difficile*-associated enteric disease. *Annals of internal medicine* 145: 758-764.
- Beutler, B., (2002) TLR4 as the mammalian endotoxin sensor. *Current topics in microbiology and immunology* 270: 109-120.
- Birnboim, H.C. & J. Doly, (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* 7: 1513-1523.
- Bligh, E.G. & W.J. Dyer, (1959) A rapid method of total lipid extraction and purification. *Canadian journal of biochemistry and physiology* 37: 911-917.
- Bloemen, P.G., M.C. van den Tweel, P.A. Henricks, F. Engels, S.S. Wagenaar, A.A. Rutten & F.P. Nijkamp, (1993) Expression and modulation of adhesion molecules on human bronchial epithelial cells. *Am J Respir Cell Mol Biol* 9: 586-593.
- Brenner, D.J., L.W. Mayer, G.M. Carlone, L.H. Harrison, W.F. Bibb, M.C. Brandileone, F.O. Sottnek, K. Irino, M.W. Reeves, J.M. Swenson & et al., (1988) Biochemical, genetic, and epidemiologic characterization of *Haemophilus influenzae* biogroup aegyptius (*Haemophilus aegyptius*) strains associated with Brazilian purpuric fever. *Journal of clinical microbiology* 26: 1524-1534.
- Bustin, S.A., V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele & C.T. Wittwer, (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry* 55: 611-622.
- Carlone, G.M., L. Gorelkin, L.L. Gheesling, S.K. Hoiseth, M.H. Mulks, S.P. O'Connor, R.S. Weyant, J.E. Myrick, L.W. Mayer & R.J. Arko, (1989) Potential virulence factors of *Haemophilus influenzae* biogroup aegyptius in Brazilian purpuric fever. The Brazilian Purpuric Fever Study Group. *The Pediatric infectious disease journal* 8: 245-247.

- Carlone, G.M., F.O. Sottnek & B.D. Plikaytis, (1985) Comparison of outer membrane protein and biochemical profiles of *Haemophilus aegyptius* and *Haemophilus influenzae* biotype III. *Journal of clinical microbiology* 22: 708-713.
- Casin, I., F. Grimont & P.A. Grimont, (1986) Deoxyribonucleic acid relatedness between *Haemophilus aegyptius* and *Haemophilus influenzae*. *Annales de l'Institut Pasteur. Microbiology* 137B: 155-163.
- CDC (1987a) Brazilian purpuric fever: epidemic purpura fulminans associated with antecedent purulent conjunctivitis. Brazilian Purpuric Fever Study Group. *Lancet* 2: 757-761.
- CDC (1987b) *Haemophilus aegyptius* bacteraemia in Brazilian purpuric fever. Brazilian Purpuric Fever Study Group. *Lancet* 2: 761-763.
- CDC, (1986) Centers for Disease Control. Brazilian purpuric fever: *Haemophilus aegyptius* bacteremia complicating purulent conjunctivitis. *MMWR. Morbidity and mortality weekly report* 35: 553-554.
- CDC, (1985) Centers for Disease Control. Preliminary report: epidemic fatal purpuric fever among children--Brazil. *MMWR. Morbidity and mortality weekly report* 34: 217-219.
- Chen, I., P.J. Christie & D. Dubnau, (2005) The ins and outs of DNA transfer in bacteria. *Science* 310: 1456-1460.
- Chomczynski, P., (1993) A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *BioTechniques* 15: 532-534, 536-537.
- Cohen, J., (2002) The immunopathogenesis of sepsis. *Nature* 420: 885-891.
- Cury, G.C., R.F.C. Pereira, L.M. de Hollanda & M. Lancellotti, (2014) Inflammatory response of *Haemophilus influenzae* biotype *aegyptius* causing Brazilian Purpuric Fever. *Braz J Microbiol* 45: 1449-1454.
- Darling, A.E., B. Mau & N.T. Perna, (2010) progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PloS one* 5: e11147.
- Darling, A.C., B. Mau, F.R. Blattner & N.T. Perna, (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome research* 14: 1394-1403.
- Davis, J., A.L. Smith, W.R. Hughes & M. Golomb, (2001) Evolution of an autotransporter: domain shuffling and lateral transfer from pathogenic *Haemophilus* to *Neisseria*. *Journal of bacteriology* 183: 4626-4635.
- Deghmane, A.E., S. Petit, A. Topilko, Y. Pereira, D. Giorgini, M. Larribe & M.K. Taha, (2000) Intimate adhesion of *Neisseria meningitidis* to human epithelial cells is under the control of the *crgA* gene, a novel LysR-type transcriptional regulator. *EMBO J* 19: 1068-1078.
- Diep, B.A., H.A. Carleton, R.F. Chang, G.F. Sensabaugh & F. Perdreau-Remington, (2006) Roles of 34 virulence genes in the evolution of hospital- and community-associated strains of methicillin-resistant *Staphylococcus aureus*. *The Journal of infectious diseases* 193: 1495-1503.

- Dobson, S.R., J.S. Kroll & E.R. Moxon, (1992) Insertion sequence IS1016 and absence of Haemophilus capsulation genes in the Brazilian purpuric fever clone of Haemophilus influenzae biogroup aegyptius. *Infection and immunity* 60: 618-622.
- Drozdz, M., A. Piekarowicz, J.M. Bujnicki & M. Radlinska, (2012) Novel non-specific DNA adenine methyltransferases. *Nucleic acids research* 40: 2119-2130.
- Ernst, R.K., T. Guina & S.I. Miller, (2001) Salmonella typhimurium outer membrane remodeling: role in resistance to host innate immunity. *Microbes and infection / Institut Pasteur* 3: 1327-1334.
- Erwin, A.L. & R.S. Munford, (1989) Comparison of lipopolysaccharides from Brazilian purpuric fever isolates and conjunctivitis isolates of Haemophilus influenzae biogroup aegyptius. Brazilian Purpuric Fever Study Group. *Journal of clinical microbiology* 27: 762-767.
- Farley, M.M., D.S. Stephens, P.S. Brachman, Jr., R.C. Harvey, J.D. Smith & J.D. Wenger, (1992) Invasive Haemophilus influenzae disease in adults. A prospective, population-based surveillance. CDC Meningitis Surveillance Group. *Annals of internal medicine* 116: 806-812.
- Frick, A.G., T.D. Joseph, L. Pang, A.M. Rabe, J.W. St Geme, 3rd & D.C. Look, (2000) Haemophilus influenzae stimulates ICAM-1 expression on respiratory epithelial cells. *J Immunol* 164: 4185-4196.
- Geelen, T.H., G.R. Gaajetaan, E.F. Wouters, G.G. Rohde, F.M. Franssen, G.E. Grauls, E.E. Stobberingh, C.A. Bruggeman & F.R. Stassen, (2014) The host immune response contributes to Haemophilus influenzae virulence. *Respir Med* 108: 144-152.
- Gladman, S., (2009) *VelvetOptimiser*. US Patent, 2.2.5. CSIRO & Monash University.
- Guo, L., K.B. Lim, J.S. Gunn, B. Bainbridge, R.P. Darveau, M. Hackett & S.I. Miller, (1997) Regulation of lipid A modifications by Salmonella typhimurium virulence genes phoP-phoQ. *Science* 276: 250-253.
- Harrison, L.H., V. Simonsen & E.A. Waldman, (2008) Emergence and disappearance of a virulent clone of Haemophilus influenzae biogroup aegyptius, cause of Brazilian purpuric fever. *Clinical microbiology reviews* 21: 594-605.
- Hayek, N., (2013) Lateral transfer and GC content of bacterial resistant genes. *Front Microbiol* 4: 41.
- Helms, B., (2006) Host-Pathogen interaction: Lipids grease the way. *Eur. J. Lipid Sci. Technol.* 108: 3.
- Hempel, R.J., D.J. Morton, T.W. Seale, P.W. Whitby & T.L. Stull, (2013) The role of the RNA chaperone Hfq in Haemophilus influenzae pathogenesis. *BMC Microbiol* 13: 134.
- Hernandez, M., A. Leichtle, K. Pak, N.J. Webster, S.I. Wasserman & A.F. Ryan, (2015) The transcriptome of a complete episode of acute otitis media. *BMC Genomics* 16: 259.

Herriott, R.M., E.M. Meyer & M. Vogt, (1970) Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*. *Journal of bacteriology* 101: 517-524.

Hollanda, L.M., G.C. Cury, R.F. Pereira, G.A. Ferreira, A. Sousa, E.M. Sousa & M. Lancellotti, (2011) Effect of mesoporous silica under *Neisseria meningitidis* transformation process: environmental effects under meningococci transformation. *Journal of nanobiotechnology* 9: 28.

Hong, Y., G. Wang & R.J. Maier, (2008) The NADPH quinone reductase MdaB confers oxidative stress resistance to *Helicobacter hepaticus*. *Microb Pathog* 44: 169-174.

Hood, D.W., M.E. Deadman, M.P. Jennings, M. Bisercic, R.D. Fleischmann, J.C. Venter & E.R. Moxon, (1996) DNA repeats identify novel virulence genes in *Haemophilus influenzae*. *Proceedings of the National Academy of Sciences of the United States of America* 93: 11121-11125.

Hou, T., Q. Du, L. Wang, H. Zhou, Y. Xi, Z. Chen & Z. Wang, (2014) Cross reactivity of *Neisseria meningitidis* *crgA* diagnostic PCR primers with nontypeable *haemophilus influenzae*. *Clin Lab* 60: 1425-1429.

Ieva, R., C. Alaimo, I. Delany, G. Spohn, R. Rappuoli & V. Scarlato, (2005) *CrgA* is an inducible LysR-type regulator of *Neisseria meningitidis*, acting both as a repressor and as an activator of gene transcription. *J Bacteriol* 187: 3421-3430.

Irino, K., F. Grimont, I. Casin & P.A. Grimont, (1988) rRNA gene restriction patterns of *Haemophilus influenzae* biogroup *aegyptius* strains associated with Brazilian purpuric fever. *Journal of clinical microbiology* 26: 1535-1538.

Kesanopoulos, K., G. Tzanakaki, S. Levidiotou, C. Blackwell & J. Kremastinou, (2005) Evaluation of touch-down real-time PCR based on SYBR Green I fluorescent dye for the detection of *Neisseria meningitidis* in clinical samples. *FEMS Immunol Med Microbiol* 43: 419-424.

Khuller, G.K., R. Taneja, S. Kaur & J.N. Verma, (1982) Lipid composition and virulence of *Mycobacterium tuberculosis* H37Rv. *The Australian journal of experimental biology and medical science* 60 (Pt 5): 541-547.

Kilian, M., I. Sorensen & W. Frederiksen, (1979) Biochemical characteristics of 130 recent isolates from *Haemophilus influenzae* meningitis. *Journal of clinical microbiology* 9: 409-412.

Kilian, M., (1976) A taxonomic study of the genus *Haemophilus*, with the proposal of a new species. *Journal of general microbiology* 93: 9-62.

Kostman, J.R., B.L. Sherry, C.L. Fligner, S. Egaas, P. Sheeran, L. Baken, J.E. Bauwens, C. Clausen, D.M. Sherer, J.J. Plorde & et al., (1993) Invasive *Haemophilus influenzae* infections in older children and adults in Seattle. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 17: 389-396.

Kroll, J.S., J.L. Farrant, S. Tyler, M.B. Coulthart & P.R. Langford, (2002) Characterisation and genetic organisation of a 24-MDa plasmid from the Brazilian Purpuric Fever clone of *Haemophilus influenzae* biogroup *aegyptius*. *Plasmid* 48: 38-48.

- Kroll, J.S., K.E. Wilks, J.L. Farrant & P.R. Langford, (1998) Natural genetic exchange between *Haemophilus* and *Neisseria*: intergeneric transfer of chromosomal genes between major human pathogens. *Proceedings of the National Academy of Sciences of the United States of America* 95: 12381-12385.
- Leaves, N.I., E.C. Anderson & S.J. Toy, (1994) Outer membrane protein profiling to distinguish between *Haemophilus aegyptius* and non-capsulate *Haemophilus influenzae* biotype III. *British journal of biomedical science* 51: 307-311.
- Li, M.S., J.L. Farrant, P.R. Langford & J.S. Kroll, (2003) Identification and characterization of genomic loci unique to the Brazilian purpuric fever clonal group of *H. influenzae* biogroup *aegyptius*: functionality explored using meningococcal homology. *Molecular microbiology* 47: 1101-1111.
- Look, D.C., S.R. Rapp, B.T. Keller & M.J. Holtzman, (1992) Selective induction of intercellular adhesion molecule-1 by interferon-gamma in human airway epithelial cells. *Am J Physiol* 263: L79-87.
- Mazloun, H.A., M. Kilian, Z.M. Mohamed & M.D. Said, (1982) Differentiation of *Haemophilus aegyptius* and *Haemophilus influenzae*. *Acta pathologica, microbiologica, et immunologica Scandinavica. Section B, Microbiology* 90: 109-112.
- McIntyre, P., G. Wheaton, J. Erlich & D. Hansman, (1987) Brazilian purpuric fever in central Australia. *Lancet* 2: 112.
- Mell, J.C., S. Shumilina, I.M. Hall & R.J. Redfield, (2011) Transformation of natural genetic variation into *Haemophilus influenzae* genomes. *PLoS pathogens* 7: e1002151.
- Miller, S.I., R.K. Ernst & M.W. Bader, (2005) LPS, TLR4 and infectious disease diversity. *Nature reviews. Microbiology* 3: 36-46.
- New, R.R.C., (1990) *Liposomes: a practical approach*. IRL Press.
- Parra, E., E. Castaneda & J. Moreno, (2007) [Identification of *Haemophilus influenzae*, *Streptococcus pneumoniae* and *Neisseria meningitidis* by polymerase chain reaction]. *Biomedica* 27: 454-460.
- Peak, I.R., M.P. Jennings, D.W. Hood & E.R. Moxon, (1999) Tetranucleotide repeats identify novel virulence determinant homologues in *Neisseria meningitidis*. *Microbial pathogenesis* 26: 13-23.
- Pereira, R.F.C., D.A. Alves, R.K. Jacinto, L.M.d. Hollanda, L.M.C. Verinaud, C.M.L. Machado & M. Lancellotti, (2011) Effects of *Neisseria meningitidis* Infection in Tumor Glioblastoma Cell Line NG97: Respiratory Pathogen Inducing Apoptosis. *J Bacteriol Parasitol* 2: 1-5.
- Peters, V.B. & L.G. Rubin, (1992) Antibodies to lipooligosaccharide of a Brazilian purpuric fever isolate of *Haemophilus influenzae* biogroup *aegyptius* lack bactericidal and protective activity. *Infection and immunity* 60: 3423-3427.
- Pittman, M. & D.J. Davis, (1950) Identification of the Koch-Weeks bacillus (*Haemophilus aegyptius*). *Journal of bacteriology* 59: 413-426.

- Preston, A., R.E. Mandrell, B.W. Gibson & M.A. Apicella, (1996) The lipooligosaccharides of pathogenic gram-negative bacteria. *Critical reviews in microbiology* 22: 139-180.
- Quinn, F.D., R.S. Weyant, M.J. Worley, E.H. White, E.A. Utt & E.A. Ades, (1995) Human microvascular endothelial tissue culture cell model for studying pathogenesis of Brazilian purpuric fever. *Infection and immunity* 63: 2317-2322.
- Quinn, F.D., R.S. Weyant, F.J. Candal & E.W. Ades, (1994) Destruction of human microvascular endothelial cell capillary-like microtubules by Brazilian purpuric fever-associated *Haemophilus influenzae* biogroup *aegyptius*. *Pathobiology: journal of immunopathology, molecular and cellular biology* 62: 109-112.
- Rebeil, R., R.K. Ernst, B.B. Gowen, S.I. Miller & B.J. Hinnebusch, (2004) Variation in lipid A structure in the pathogenic yersiniae. *Molecular microbiology* 52: 1363-1373.
- Rissman, A.I., B. Mau, B.S. Biehl, A.E. Darling, J.D. Glasner & N.T. Perna, (2009) Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics* 25: 2071-2073.
- Rubin, L.G., (1995) Role of the 145-kilodalton surface protein in virulence of the Brazilian purpuric fever clone of *Haemophilus influenzae* biogroup *aegyptius* for infant rats. *Infection and immunity* 63: 3555-3558.
- Rubin, L.G. & J.W. St Geme, 3rd, (1993) Role of lipooligosaccharide in virulence of the Brazilian purpuric fever clone of *Haemophilus influenzae* biogroup *aegyptius* for infant rats. *Infection and immunity* 61: 650-655.
- Rubin, L.G. & A. Rizvi, (1991) Antibody to a 145-kilodalton outer membrane protein has bactericidal activity and protective activity against experimental bacteremia caused by a Brazilian purpuric fever isolate of *Haemophilus influenzae* biogroup *aegyptius*. The Brazilian Purpuric Fever Study Group. *Infect Immun* 59: 4576-4582.
- Rubin, L.G. & G.M. Carlone, (1989) An infant rat model of bacteremia with Brazilian purpuric fever isolates of *Haemophilus influenzae* biogroup *aegyptius* (*Haemophilus aegyptius*). The Brazilian Purpuric Fever Study Group. *The Pediatric infectious disease journal* 8: 247-248.
- Santana-Porto, E.A., A.A. Oliveira, M.R. da-Costa, A. Pinheiro, C. Oliveira, M.L. Lopes, L.E. Pereira, C. Sacchi, W.N. Araujo & J. Sobel, (2009) Suspected Brazilian purpuric fever, Brazilian Amazon region. *Emerging infectious diseases* 15: 675-676.
- Schmittgen, T.D. & K.J. Livak, (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nature protocols* 3: 1101-1108.
- Segada, L.M. & A.J. Lesse, (1997) Creation of an isogenic P1-deficient mutant of *Haemophilus influenzae* biogroup *aegyptius*. *Gene* 204: 185-194.
- Serruto, D., T. Spadafina, M. Scarselli, S. Bambini, M. Comanducci, S. Hohle, M. Kilian, E. Veiga, P. Cossart, M.R. Oggioni, S. Savino, I. Ferlenghi, A.R. Taddei, R. Rappuoli, M. Pizza, V. Maignani & B. Arico, (2009) HadA is an atypical new multifunctional trimeric coiled-coil adhesin of *Haemophilus influenzae* biogroup *aegyptius*, which promotes entry into host cells. *Cellular microbiology* 11: 1044-1063.

- Sinha, S., J.C. Mell & R.J. Redfield, (2012) Seventeen Sxy-dependent cyclic AMP receptor protein site-regulated genes are needed for natural transformation in *Haemophilus influenzae*. *Journal of bacteriology* 194: 5245-5254.
- Smith, A.L., D.H. Smith, D.R. Averill, Jr., J. Marino & E.R. Moxon, (1973) Production of *Haemophilus influenzae* b meningitis in infant rats by intraperitoneal inoculation. *Infection and immunity* 8: 278-290.
- Strouts, F.R., P. Power, N.J. Croucher, N. Corton, A. van Tonder, M.A. Quail, P.R. Langford, M.J. Hudson, J. Parkhill, J.S. Kroll & S.D. Bentley, (2012) Lineage-specific virulence determinants of *Haemophilus influenzae* biogroup aegyptius. *Emerging infectious diseases* 18: 449-457.
- Taha, M.K., J.M. Alonso, M. Cafferkey, D.A. Caugant, S.C. Clarke, M.A. Diggle, A. Fox, M. Frosch, S.J. Gray, M. Guiver, S. Heuberger, J. Kalmusova, K. Kesanopoulos, A.M. Klem, P. Kriz, J. Marsh, P. Molling, K. Murphy, P. Olcen, O. Sanou, G. Tzanakaki & U. Vogel, (2005) Interlaboratory comparison of PCR-based identification and genogrouping of *Neisseria meningitidis*. *J Clin Microbiol* 43: 144-149.
- Takala, A.K., J. Eskola & L. van Alphen, (1990) Spectrum of invasive *Haemophilus influenzae* type b disease in adults. *Archives of internal medicine* 150: 2573-2576.
- Tondella, M.L., F.D. Quinn & B.A. Perkins, (1995) Brazilian purpuric fever caused by *Haemophilus influenzae* biogroup aegyptius strains lacking the 3031 plasmid. *The Journal of infectious diseases* 171: 209-212.
- Tosi, M.F., J.M. Stark, C.W. Smith, A. Hamedani, D.C. Gruenert & M.D. Infeld, (1992) Induction of ICAM-1 expression on human airway epithelial cells by inflammatory cytokines: effects on neutrophil-epithelial cell adhesion. *Am J Respir Cell Mol Biol* 7: 214-221.
- Tsao, D., K.L. Nelson, D. Kim & A.L. Smith, (2012) Infant rat infection modifies phenotypic properties of an invasive nontypeable *Haemophilus influenzae*. *Microbes Infect* 14: 509-516.
- Tuyama, M., R.F. Boente, M.C. Rebelo, R.P. Igreja & D.E. Barroso, (2008) The utility of the polymerase chain reaction assay for aetiologic definition of unspecified bacterial meningitis cases. *Mem Inst Oswaldo Cruz* 103: 138-142.
- van der Meer-Janssen, Y.P., J. van Galen, J.J. Batenburg & J.B. Helms, (2010) Lipids in host-pathogen interactions: pathogens exploit the complexity of the host cell lipidome. *Progress in lipid research* 49: 1-26.
- Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe & F. Speleman, (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* 3: RESEARCH0034.
- Varela, J.N., M.C. Amstalden, R.F. Pereira, L.M. de Hollanda, H.J. Ceragioli, V. Baranauskas & M. Lancellotti, (2014) *Haemophilus influenzae* porine ompP2 gene transfer mediated by graphene oxide nanoparticles with effects on transformation process and virulence bacterial capacity. *Journal of nanobiotechnology* 12: 14.

Virata, M., N.E. Rosenstein, J.L. Hadler, N.L. Barrett, M.L. Tondella, L.W. Mayer, R.S. Weyant, B. Hill & B.A. Perkins, (1998) Suspected Brazilian purpuric fever in a toddler with overwhelming Epstein-Barr virus infection. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* 27: 1238-1240.

Wang, W.Y., J.H. Lim & J.D. Li, (2012) Synergistic and feedback signaling mechanisms in the regulation of inflammation in respiratory infections. *Cell Mol Immunol* 9: 131-135.

Weyant, R.S., W.F. Bibb, D.S. Stephens, B.P. Holloway, W.F. Moo-Penn, K.A. Birkness, L.O. Hesel & L.W. Mayer, (1990) Purification and characterization of a pilin specific for Brazilian purpuric fever-associated *Haemophilus influenzae* biogroup *aegyptius* (*H. aegyptius*) strains. *Journal of clinical microbiology* 28: 756-763.

Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen & T.L. Madden, (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13: 134.

Zaretsky, F.R. & T.H. Kawula, (1999) Examination of early interactions between *Haemophilus ducreyi* and host cells by using cocultured HaCaT keratinocytes and foreskin fibroblasts. *Infect Immun* 67: 5352-5360.

Zerbino, D.R. & E. Birney, (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome research* 18: 821-829.

ANEXOS

1. Lista de *primers*

Tabela 12. Lista de *primers* utilizados neste trabalho, suas funções e principais características.

Concatenação dos <i>contigs</i> de Hae 258/86		
Nome do <i>primer</i>	Sequência 5' – 3'	Características
25 – 08 <i>Forward</i>	CGATCCTGATGTGTTACTCG	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
25 – 08 <i>Reverse</i>	GCAAATTGTATGTCCGATTGTGG	
08 – 13 <i>Forward</i>	CACTACCTGCTCAAGGTG	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
08 – 13 <i>Reverse</i>	GCAAATTGTATGTCCGATTGTGG	
13 – 11 <i>Forward</i>	GAAGCAGGTCACGCCATTGTGG	Obtido a partir do genoma de Hae 258/86; TA = 62 °C
13 – 11 <i>Reverse</i>	GAAGCTGGTCATGCGATTGTGG	
11 – 28 <i>Forward</i>	GGCAACCAGCAGAGAAACCC	Obtido a partir do genoma de Hae 258/86; TA = 61 °C
11 – 28 <i>Reverse</i>	GAATCCCCATAATTTGAGGCTCGC	
06-23 <i>Forward</i>	TAGCAACACAACAAGTTTACAAGG	Obtido a partir do genoma de Hae 258/86; TA = 56 °C
06-23 <i>Reverse</i>	GAGCAAGGCGTTATTTTCAGG	
26 –24 <i>Forward</i>	GCAAGGTAAGCAATGAAACGC	Obtido a partir do genoma de Hae 258/86; TA = 53 °C
26 –24 <i>Reverse</i>	GTGATATGGAAGATAAATGGTAGC	
24 – 17 <i>Forward</i>	CTGACTCACAACTGGCAAG	Obtido a partir do genoma de Hae 258/86; TA = 56 °C
24 – 17 <i>Reverse</i>	GCTGAATTAGAAACGGTTGAGAAG	
17 – 18 <i>Forward</i>	CCTGTAGGGCTAGGAGCA	Obtido a partir do genoma de Hae 258/86; TA = 59 °C

17 – 18 <i>Reverse</i>	GAAAATAAAGTGGTGGGCGATAACC	
18 – 04 <i>Forward</i>	CGCAACAGCCGCAATGTATGC	Obtido a partir do genoma de Hae 258/86; TA = 63 °C
18 – 04 <i>Reverse</i>	ACGCTGCCTGTCTCTTTTAAAGCC	
16 – 25 <i>Forward</i>	CAATAATCGTCCAGCCTTCG	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
16 – 25 <i>Reverse</i>	ACGCAGGAAGATGTAGAGG	
21 – 31 <i>Forward</i>	CGCCTTTGACTAAGCGAGTGTTG	Obtido a partir do genoma de Hae 258/86; TA = 62 °C
21 – 31 <i>Reverse</i>	TAGCATTGTGGCAAGAGCTGC	
02 – 01 <i>Forward</i>	TCATTAACTTTGCTCCACCCA	Obtido a partir do genoma de Hae 258/86; TA = 58 °C
02 – 01 <i>Reverse</i>	TCGCTGGATTTGGATTGGCG	
01 – 05 <i>Forward</i>	GAGCCTAAAATATGGTCAGCACC	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
01 – 05 <i>Reverse</i>	CTCTAATGTGGGAGTGTTAGCA	
05 – 15 <i>Forward</i>	CATATCTTTCGACACGGGGTTC	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
05 – 15 <i>Reverse</i>	GGAAGGAGAAGGAATAGCGG	
15 – 10 <i>Forward</i>	TATGAATTGGCGGAATGGACG	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
15 – 10 <i>Reverse</i>	GTTTTGCTATGATCTGTGGACG	
10 – 12 <i>Forward</i>	CGCATTATAGGAAACAAGCACC	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
10 – 12 <i>Reverse</i>	TCAGCAATTTTTGGTGGGG	
12 – 09 <i>Forward</i>	ACCCCATATCTTTCGACACG	Obtido a partir do genoma de Hae 258/86; TA = 55 °C
12 – 09 <i>Reverse</i>	TAGTCAGATGTTAAGTCAAGCTG	
09 – 03 <i>Forward</i>	TCCGATGATGTGAATGGTCG	Obtido a partir do genoma de Hae 258/86; TA = 56 °C
09 – 03 <i>Reverse</i>	CTTCACTGGTTTTTGATTGTGC	
03 – 07 <i>Forward</i>	CATAATCACGGATTTGTTTGGTGGA	Obtido a partir do genoma de Hae 258/86; TA = 58 °C
03 – 07 <i>Reverse</i>	ATCAGACAATCTGTGTGGGCAC	
07 – 32 <i>Forward</i>	CTTTCGACACGGGGTTCTTCTG	Obtido a partir do genoma de Hae 258/86; TA = 62 °C
07 – 32 <i>Reverse</i>	ACTCGTCAGCAAGAAAGCAAGC	
32 – 29 <i>Forward</i>	CTGGGGTGAAGTCGTAACAAGG	Obtido a partir do genoma de Hae 258/86; TA = 60 °C

32 – 29 <i>Reverse</i>	TCCTGTCTTGCCTACTCAGAC	
39 – 38 <i>Forward</i>	CAAAGAATTATCCTGGCGGC	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
39 – 38 <i>Reverse</i>	GAGGAAAGACTGAGTAACGAGAG	
23 – 22 <i>Forward</i>	TGACTGTTACTCATCGTCTATTTCG	Obtido a partir do genoma de Hae 258/86; TA = 56 °C
23 – 22 <i>Reverse</i>	TGGCGTGCGTTGTAAGC	
22 – 27 <i>Forward</i>	CATTTGATTGCCTTGCCAGAG	Obtido a partir do genoma de Hae 258/86; TA = 58 °C
22 – 27 <i>Reverse</i>	GATACGGATACGCACATTGC	
04 – 20 <i>Forward</i>	GGTTGAATTGTGGAAGCAATCG	Obtido a partir do genoma de Hae 258/86; TA = 58 °C
04 – 20 <i>Reverse</i>	GACTGATGAGGTGGATTTAATCGTG	
20 – 16 <i>Forward</i>	CGTTCTCCATTTGAAAAAAGGACTG	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
20 – 16 <i>Reverse</i>	ACCACAACCTTCTTACCAAGC	
Fechamento do plasmídeo de Hae		
Nome do <i>primer</i>	Sequência 5' – 3'	Características
plasm19 <i>Forward</i>	ACGCAAAAATCGTATGACGC	Obtido a partir do genoma de Hae 258/86; TA = 57 °C
plasm19 <i>Reverse</i>	TCGCAATGAAACTTGAATCAGCC	
<i>Knockout</i> de genes em Hae		
Nome do <i>primer</i>	Sequência 5' – 3'	Características
<i>lasA</i> <i>Forward</i>	CAGGCTGTATGACGTCTCCATC	Obtido a partir do genoma de Hae 254/86; TA = 55 °C
<i>lasA</i> <i>Reverse</i>	GAACAGACTGGAGATTACTGGC	
<i>hadAΔ</i> <i>Forward</i>	GCACAAGTCAAAAAAGATGAACTTAGTGAG	Obtido a partir do genoma de Hae 254/86; TA = 55 °C
<i>hadAΔ</i> <i>Reverse</i>	GCAAGACCACGTTTTAAGATCTTTATTAAGG	
<i>tabAΔ</i> <i>Forward</i>	GGTTGTTTCAGAATTAACGAAAGGTG	Obtido a partir do genoma de Hae 254/86; TA = 55 °C
<i>tabAΔ</i> <i>Reverse</i>	CAGTTTACGGTGCTATTATTACCAATAG	
ermAM <i>Forward</i>	AAGCTTGCCGTCTGAATGGGACCTCTTACTTCTTGG	Amplificação do cassete de resistência à eritromicina ermAM
ermAM <i>Reverse</i>	GCAAACCTTAAGAGTGTGTTGATAG	

Seleção de genes de referência de Hae para qRT-PCR		
Nome do primer	Sequência 5' – 3'	Características
<i>alaS Forward</i>	GCACCTTACGCATCAGACAAC	Obtido a partir do genoma de Hae 254/86, alanil-tRNA sintetase, 137 pb
<i>alaS Reverse</i>	TAACGGTCGCCATCTTCTTC	
<i>era Forward</i>	CGGTGACGGGTAGTTTGTG	Obtido a partir do genoma de Hae 254/86, proteína Era de ligação ao GTP, 114 pb
<i>era Reverse</i>	CGGCTTTATTGCTATCGTAGGTC	
<i>gmk Forward</i>	TCAATCGCAGGTAAAGAGGTTTC	Obtido a partir do genoma de Hae 254/86, guanilato quinase 142 pb
<i>gmk Reverse</i>	CACGCCAGGTGAAGTTG	
<i>gyrA Forward</i>	AAGCAGAGGATGGTATCGTGAG	Obtido a partir do genoma de Hae 254/86, DNA girase subunidade A, 94 pb
<i>gyrA Reverse</i>	GTGGCGGTAAAGGCAAATC	
<i>map Forward</i>	CTTCCGATGTGCTTGTGATG	Obtido a partir do genoma de Hae 254/86, metionina aminopeptidase, 147 pb
<i>map Reverse</i>	AGGTCGCCTTTGGGAAAC	
<i>primase Forward</i>	TTGGCTAACTGTGAAGGATGG	Obtido a partir do genoma de Hae 254/86, DNA primase 144 pb
<i>primase Reverse</i>	CGCCCCTTTATTGATGATTTG	
<i>recA Forward</i>	GGTAATCCTGAAACCACCACAG	Obtido a partir do genoma de Hae 254/86, recombinase A, 72 pb
<i>recA Reverse</i>	ACGGCGAATATCTAAGCGAAC	
<i>recF Forward</i>	TTACGACGAACCGCACTTTAC	Obtido a partir do genoma de Hae 254/86, proteína de replicação e reparo de DNA RecF, 94 pb
<i>recF Reverse</i>	GGTGTTCCTTGACGTAGTTTC	
<i>rho Forward</i>	AAAACAGCACGCCAAAAGTG	Obtido a partir do genoma de Hae 254/86, fator de terminação de transcrição Rho, 82 pb
<i>rho Reverse</i>	ACGAAGGAAACCAAACCATC	
<i>rpoA Forward</i>	AGCACAACCTGGCATTGAAG	Obtido a partir do genoma de Hae 254/86, RNA polimerase, subunidade alfa, 72 pb
<i>rpoA Reverse</i>	GAGCGTGGTTTTGGTCATACTC	
<i>rpoC Forward</i>	GCGGCTAAGAAAATGGTTGAG	Obtido a partir do genoma de Hae 254/86, RNA polimerase, subunidade beta, 114 pb
<i>rpoC Reverse</i>	ACCAAACGGTGAAGTGTTG	
<i>rpoD Forward</i>	TGCTGTGGCAGAATCAAGAG	Obtido a partir do genoma de Hae 254/86, RNA polimerase, fator sigma, 85 pb
<i>rpoD Reverse</i>	ACCTATTGGCGATGACGATG	
Expressão de genes de Hae por qRT-PCR		

Nome do <i>primer</i>	Sequência 5' – 3'	Características
<i>las Forward</i>	CTTGGTGGCGATGTTTTCTC	Obtido a partir do genoma de Hae 254/86, 82 pb
<i>las Reverse</i>	TGCGTTGTTCTGCTTGTC	
<i>las Hae Forward</i>	GGCTGCTGGTGATGTTCC	Obtido a partir do genoma de Hae 254/86, 78 pb
<i>las Hae Reverse</i>	TGCCAGTAATCTCCAGTCTGTTC	
<i>hadA Forward</i>	AGCTTGAGCAGCAAGACCAC	Obtido a partir do genoma de Hae 254/86, 94 pb
<i>hadA Reverse</i>	GGCAATCGCAACTCACACTC	
<i>hadA Hae Forward</i>	TTTTGACTTGTGCTTGTGCTTG	Obtido a partir do genoma de Hae 254/86, 80 pb
<i>hadA Hae Reverse</i>	TGTAATCGCTGTGTTGATAGGTG	
<i>tabA1 Forward</i>	TTCCAAGCCACCAATCAC	Obtido a partir do genoma de Hae 254/86, 85 pb
<i>tabA1 Reverse</i>	GCAGCAAATGGCGATAGAAC	

Expressão de citocinas em células humanas e ratos Sprague-Dawley por qRT-PCR

Nome do <i>primer</i>	Sequência 5' – 3'	Características
<i>GAPDH Forward</i>	TGCACCACCAACTGCTTAGC	Obtidos a partir do genoma de <i>Homo sapiens sapiens</i> e <i>Rattus norvegicus</i>
<i>GAPDH Reverse</i>	GGCATGGACTGTGGTCATGAG	
<i>IL-10 Forward</i>	GTGATGCCCCAAGCTGAGA	Obtidos a partir do genoma de <i>H. sapiens sapiens</i>
<i>IL-10 Reverse</i>	CACGGCCTTGCTCTTGTTT	
<i>TNF-α Forward</i>	TCTTCTCGAACCCCGAGTGA	Obtidos a partir do genoma de <i>H. sapiens sapiens</i>
<i>TNF-α Reverse</i>	CCTCTGATGGCACCACCAG	
<i>IL-10m Forward</i>	GGTTGCCAAGCCTTATCGGA	Obtidos a partir do genoma de <i>R. norvegicus</i>
<i>IL-10m Reverse</i>	ACCTGCTCCACTGCCTTGCT	
<i>TNF-αm Forward</i>	CATCTTCTCAAATTCGAGTGACAA	Obtidos a partir do genoma de <i>R. norvegicus</i>
<i>TNF-αm Reverse</i>	TGGGAGTAGACAAGGTACAACCC	

*Todos os *primers* foram avaliados quanto à sua especificidade através da plataforma Primer-BLAST (Ye et al., 2012) e quanto à formação de estruturas secundárias, dímeros, percentual de GC, temperatura de anelamento e de *melting* pelo *software* Primer3Plus (Untergasser et al., 2007).

2. Linhagens bacterianas utilizadas nos alinhamentos de sequências de DNA e proteínas

Tabela 13. Linhagens bacterianas utilizadas nas buscas e comparações de sequências.

Linhagem	Bactéria	Doença	Número de acesso (NCBI)	Pan-genômica	<i>crgA</i>	<i>mdaB</i>	Genes competência
Karina	Hae	BPF	-				
219/86	Hae	BPF	-				
254/86	Hae	BPF	FQ670178				
258/86	Hae	BPF	-				
284/86	Hae	BPF	-				
321/86	Hae	BPF	-				
406/86	Hae	BPF	-				
KC1018	Hae	Conjuntivite	-				
ATCC 11116	Hae	Conjuntivite	AFBC01000036				
RdKW20	Hi d	Cepa padrão	L42023				
F3047	Hae	Conjuntivite	FQ670204				
10810	Hi b	Meningite	NC_016809.1				
R2846	Hi NTHi	Otite média	NC_017452.1				
KR494	Hi	Miosite necrotizante	CP005967.1				
CGSHiCZ412602	Hi	Otite média	CP007805.1				
C486	Hi	Isolado clínico	CP007471.1				
PittEE	Hi	Otite média	CP000671.1				
Hi375	Hi	Otite média	CP009610.1				
053442	Nm	Meningite e sepse	CP000381.1				
FAM18	Nm	Sepse	AM421808.1				
G2136	Nm	Isolado clínico	CP002419.1				
NM3686	Nm	Isolado clínico	CP009418.1				
M7124	Nm	Isolado clínico	CP009419.1				
NM3682	Nm	Isolado clínico	CP009420.1				
NM3683	Nm	Isolado clínico	CP009421.1				
M10208	Nm	Isolado clínico	CP009422.1				

Alpha14	Nm	Isolado clínico	AM889136.1			
M04-240196	Nm	Isolado clínico	CP002423.1			
8013	Nm	Isolado clínico	FM999788.1			
Alpha522	Nm	Isolado clínico	FR845715.1			
Alpha710	Nm	Meningite e sepse	NC_017505.1			
M01-240355	Nm	Isolado clínico	CP002422.1			
510612	Nm	Meningite	CP007524.1			
WUE 2594	Nm	Isolado Clínico	FR774048.1			
LNP21362	Nm	Isolado Clínico	CP006869.1			
NZ-05/33	Nm	Isolado Clínico	CP002424.1			
M01-240149	Nm	Isolado Clínico	CP002421.1			
Z2491	Nm	Meningite e sepse	NC_003116.1			
H44/76	Nm	Linhagem de laboratório	CP002420.1			
MC58	Nm	Meningite e sepse	NC_003112.2			
020-06	NI	Isolado clínico	FN995097.1			
FA 1090	Ng	-	AE004969.1			
MS11	Ng	-	CP003909.1			
NCCP11945	Ng	Isolado clínico	CP001050.1			
FGI 57	Eb	Ambiental	CP003938.1			

Hae – *H. influenzae* biótipo aegyptius; Hi – *H. influenzae*; Hi d – Hi sorotipo d; Hi b – Hi sorotipo b; Hi NTHi – Hi não tipável; Nm – *N. meningitidis*, NI – *Neisseria lactamica*; Ng - *Neisseria gonorrhoeae*, Eb – *Enterobacteriaceae bacterium*.

*Somente as linhagens que apresentam as sequências de interesse foram utilizadas na comparação das mesmas.

3. Aprovação do CEUA



CEUA/Unicamp

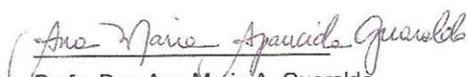
Comissão de Ética no Uso de Animais
CEUA/Unicamp

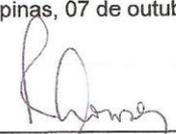
CERTIFICADO

Certificamos que o projeto "Estudo do genoma e virulência de linhagens de Haemophilus influenzae biotipo aegyptius causadoras de Febre Purpúrica Brasileira" (protocolo nº 3197-1), sob a responsabilidade de Prof. Dr. Marcelo Lancellotti / Rafaella Fabiana Carneiro Pereira, está de acordo com os Princípios Éticos na Experimentação Animal adotados pela Sociedade Brasileira de Ciência em Animais de Laboratório (SBCAL) e com a legislação vigente, LEI Nº 11.794, DE 8 DE OUTUBRO DE 2008, que estabelece procedimentos para o uso científico de animais, e o DECRETO Nº 6.899, DE 15 DE JULHO DE 2009.

O projeto foi aprovado pela Comissão de Ética no Uso de Animais da Universidade Estadual de Campinas - CEUA/UNICAMP - em 07 de outubro de 2013.

Campinas, 07 de outubro de 2013.


Prof. Dra. Ana Maria A. Guaraldo
Presidente


Fátima Alonso
Secretária Executiva

DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação de Mestrado/tese de Doutorado intitulada " ESTUDO DO GENOMA E VIRULÊNCIA DE LINHAGENS DE *HAEMOPHILUS INFLUENZAE* BIOTIPO AEGYPTIUS CAUSADORAS DE FEBRE PURPÚRICA BRASILEIRA"

() não se enquadra no § 4º do Artigo 1º da Informação CCPG 002/13, referente a bioética e biossegurança.

Tem autorização da(s) seguinte(s) Comissão(ões):

(X) CIBio – Comissão Interna de Biossegurança , projeto No. 069/98 – NB2,

Próton- 4280/10, Extrato prévio 2250/10 de 01/03/10 ,

Instituição: CIBio – IB - UNICAMP.

(X) CEUA – Comissão de Ética no Uso de Animais , projeto N° 3197-1, Instituição: CEUA - UNICAMP.

() CEP - Comissão de Ética em Pesquisa, protocolo No Instituição:

** Caso a Comissão seja externa ao IB/UNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.*

Rafaela Fabiana Carneiro Pereira

Rafaela Fabiana Carneiro Pereira
Aluno: (nome completo)

Marcelo Lancellotti

Marcelo Lancellotti
Orientador: (nome completo)

Para uso da Comissão ou Comitê pertinente:

(X) Deferido () Indeferido

[Assinatura]

Carimbo e assinatura

Para uso da Comissão ou Comitê pertinente:

(X) Deferido () Indeferido

[Assinatura]

Carimbo e assinatura

Prof.ª Dr.ª HELENA COSTINHO FRANCO DE OLIVEIRA
Presidente
CIBio/IB-UNICAMP

4. Dados ESI(+)-MS

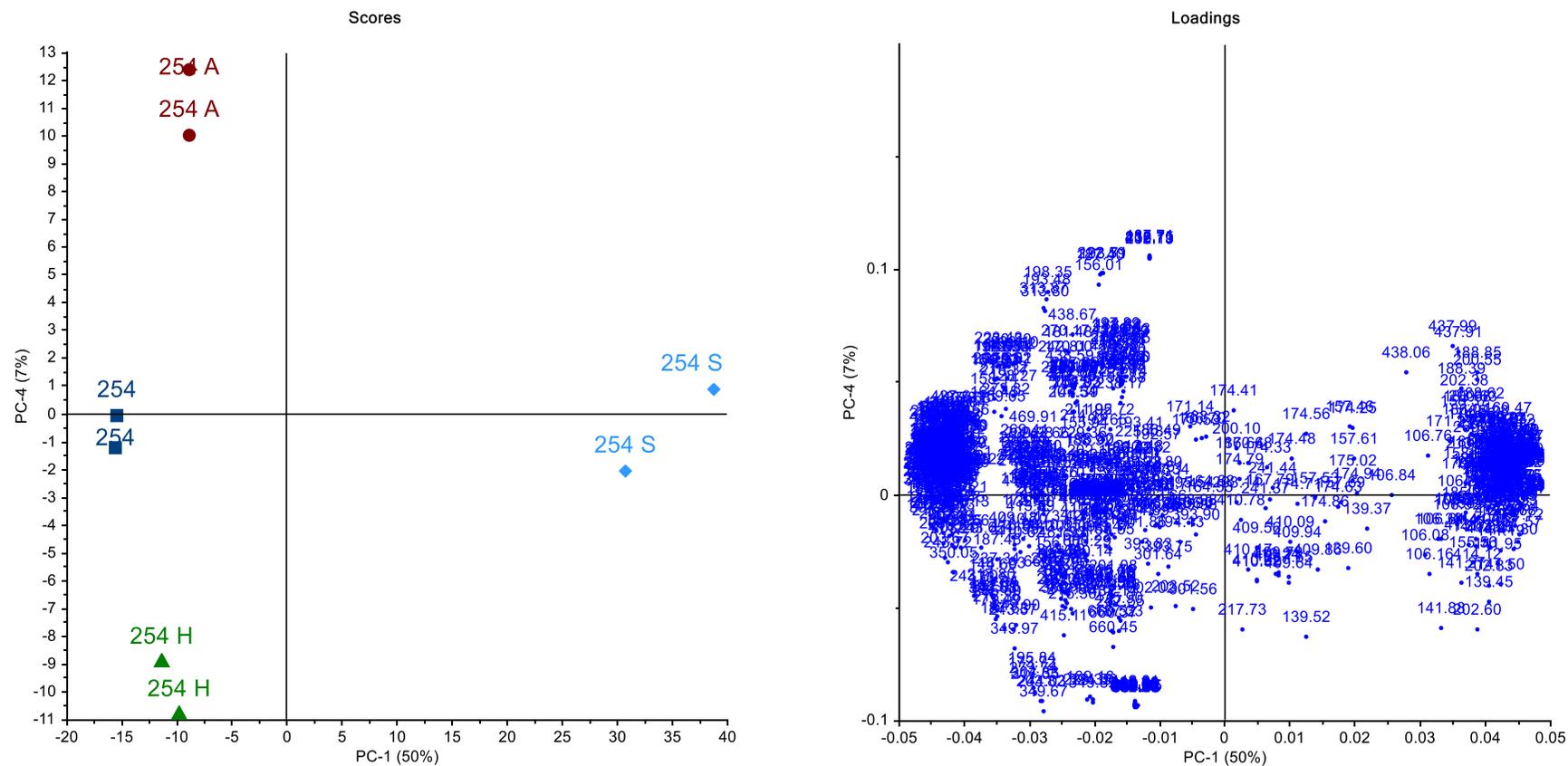


Figura 26. PCA dos dados de ESI(+)-MS das amostras de 254/86: *Scores* – amostras e *loadings* – íons. Pode-se observar a separação das amostras, principalmente as de origem animal (S) em relação às demais.

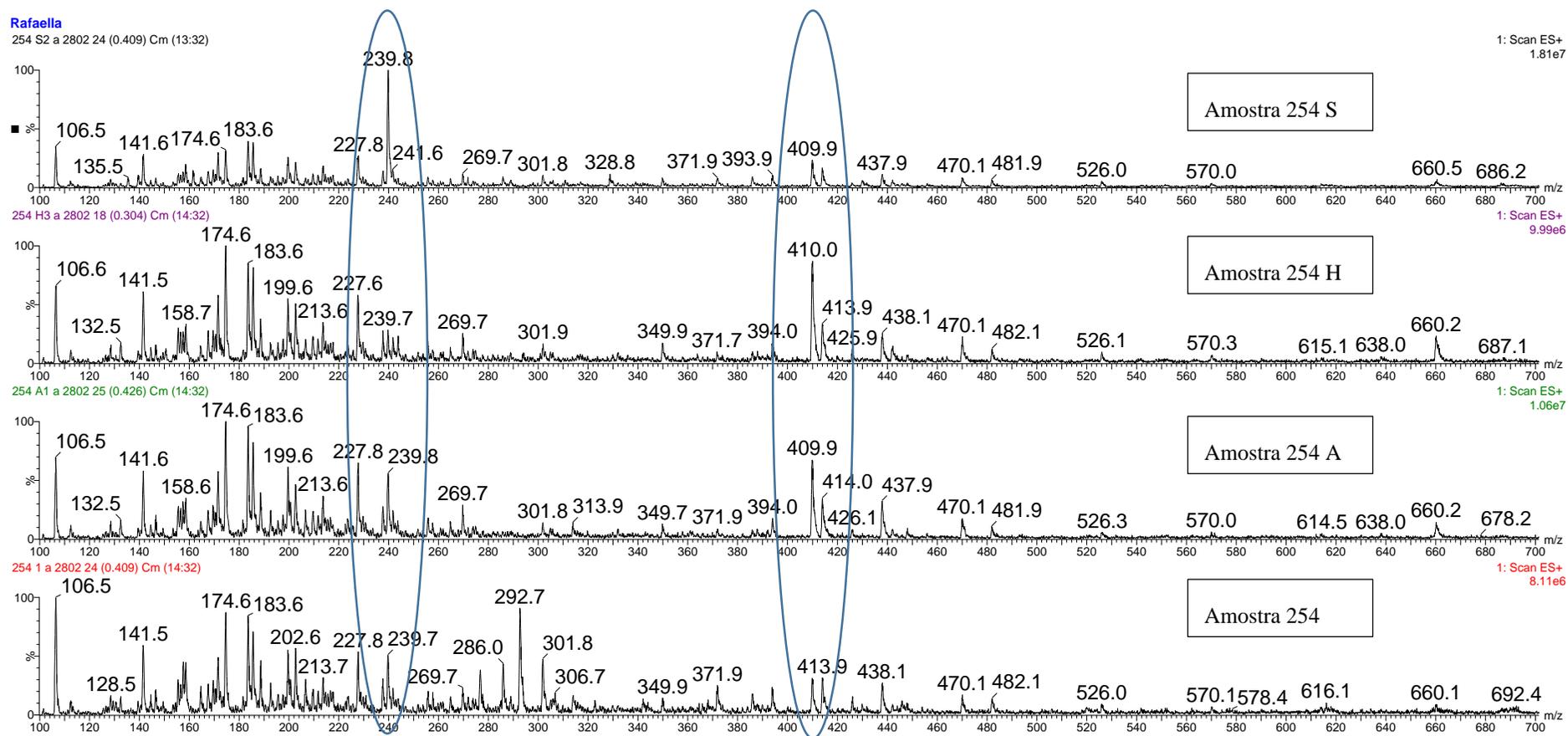


Figura 27. ESI(+)-MS fingerprints das amostras de 254/86. Os íons que se destacaram nestas amostras estão em evidência.

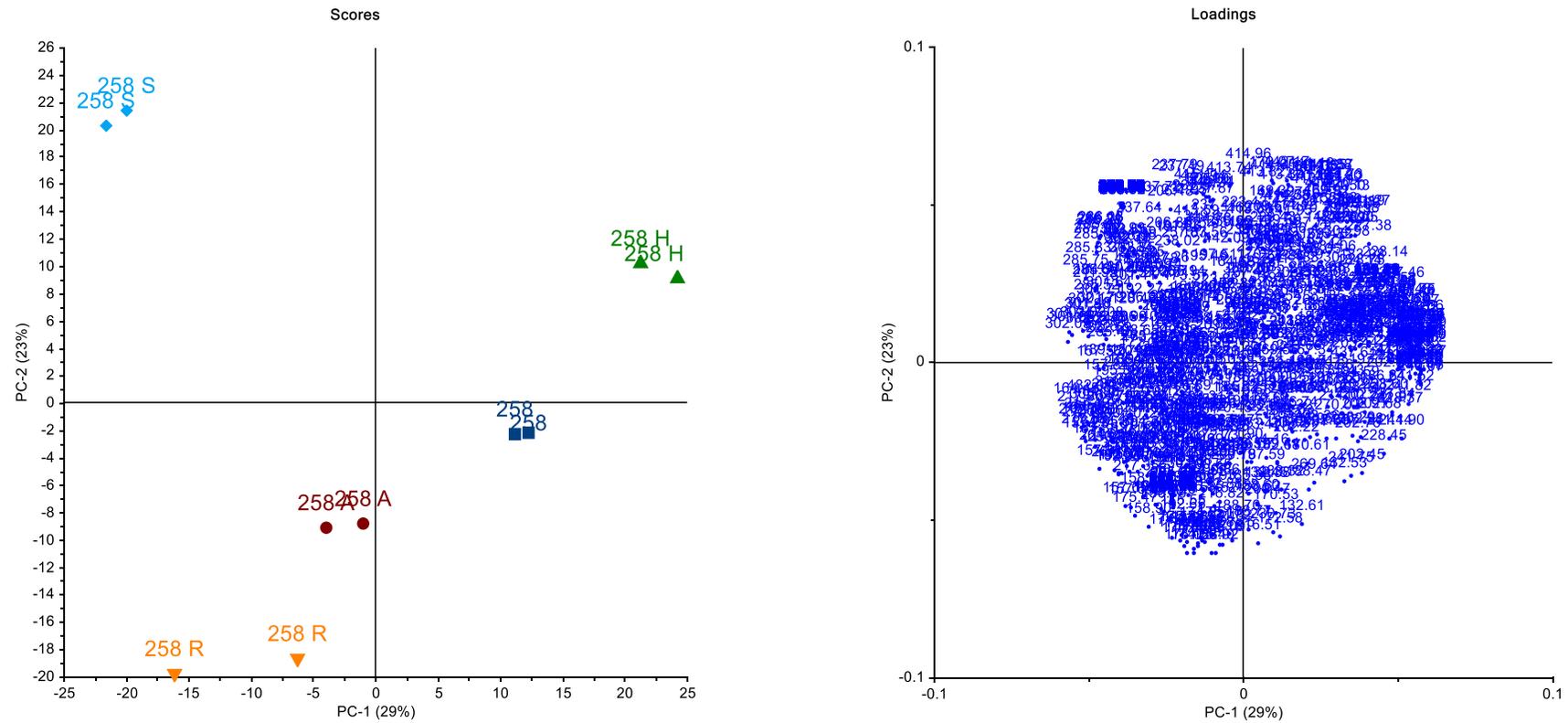


Figura 28. PCA dos dados de ESI(+)-MS das amostras de 258/86: *Scores* – amostras e *loadings* – íons. Pode-se observar a separação das amostras, principalmente as de origem animal (S) em relação às demais.

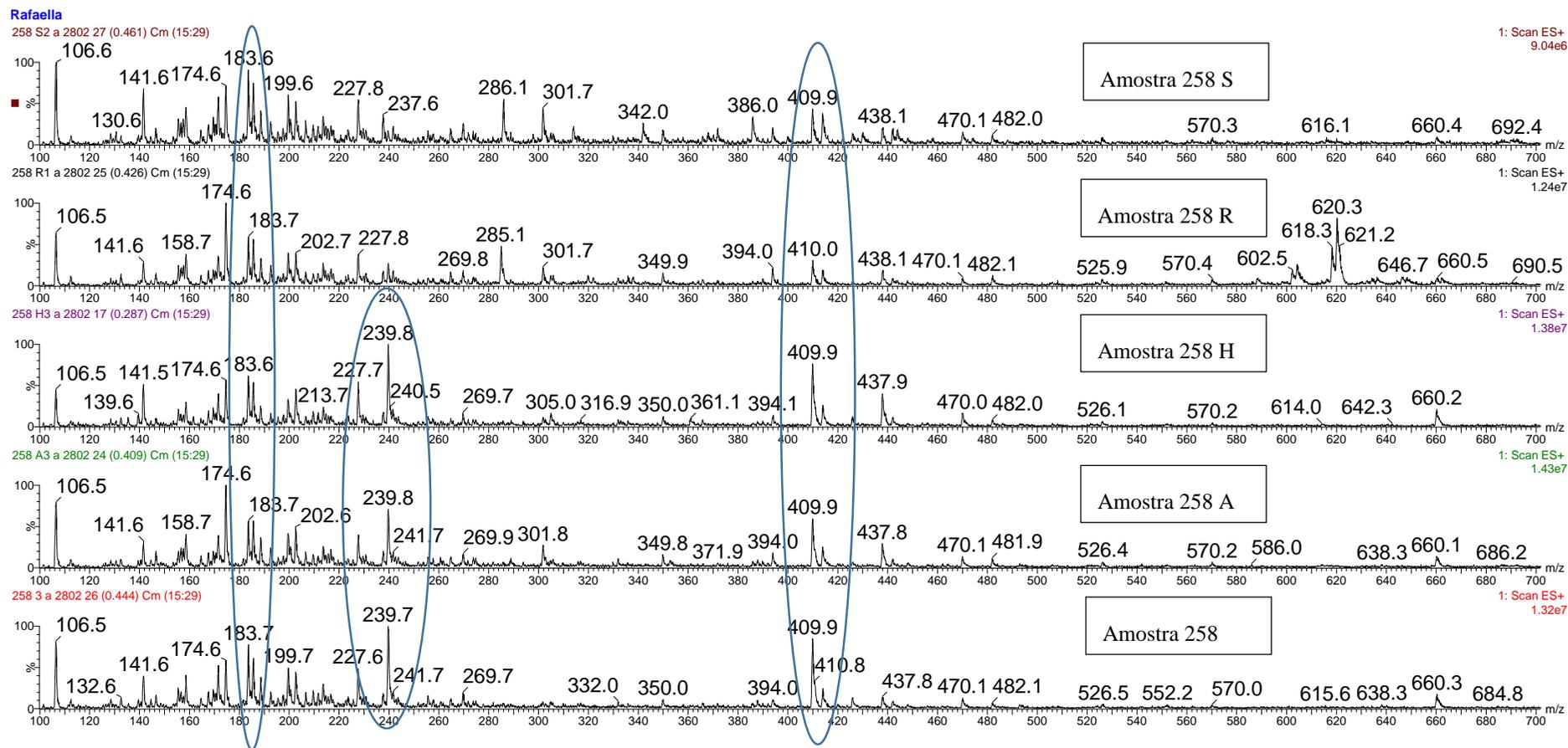


Figura 29. ESI(+)-MS *fingerprints* das amostras de 258/86. Os íons que se destacaram nestas amostras estão em evidência

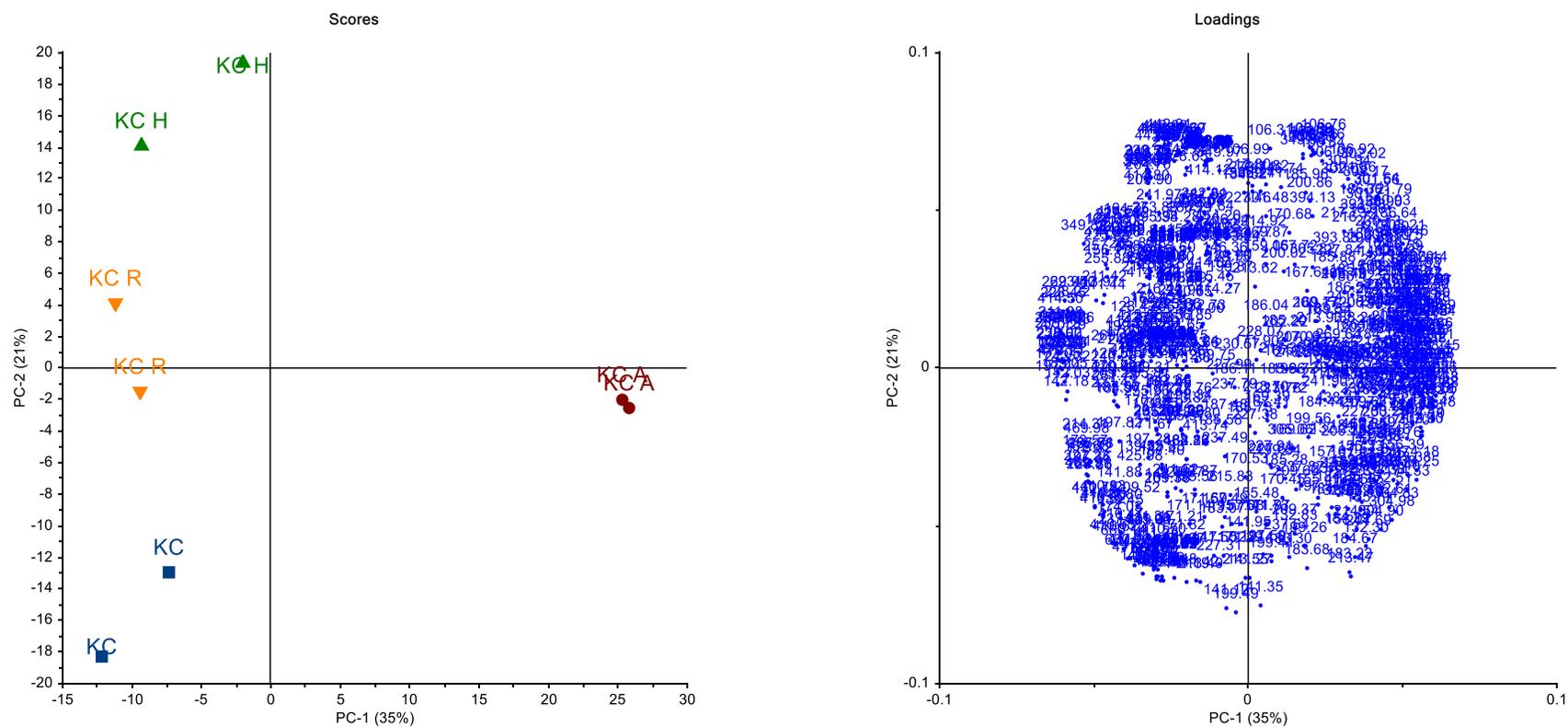


Figura 30. PCA dos dados de ESI(+)-MS das amostras de KC1018: *Scores* – amostras e *loadings* – íons. Não é possível observar a separação das amostras, exceto as de origem de infecção em células A549 (A).

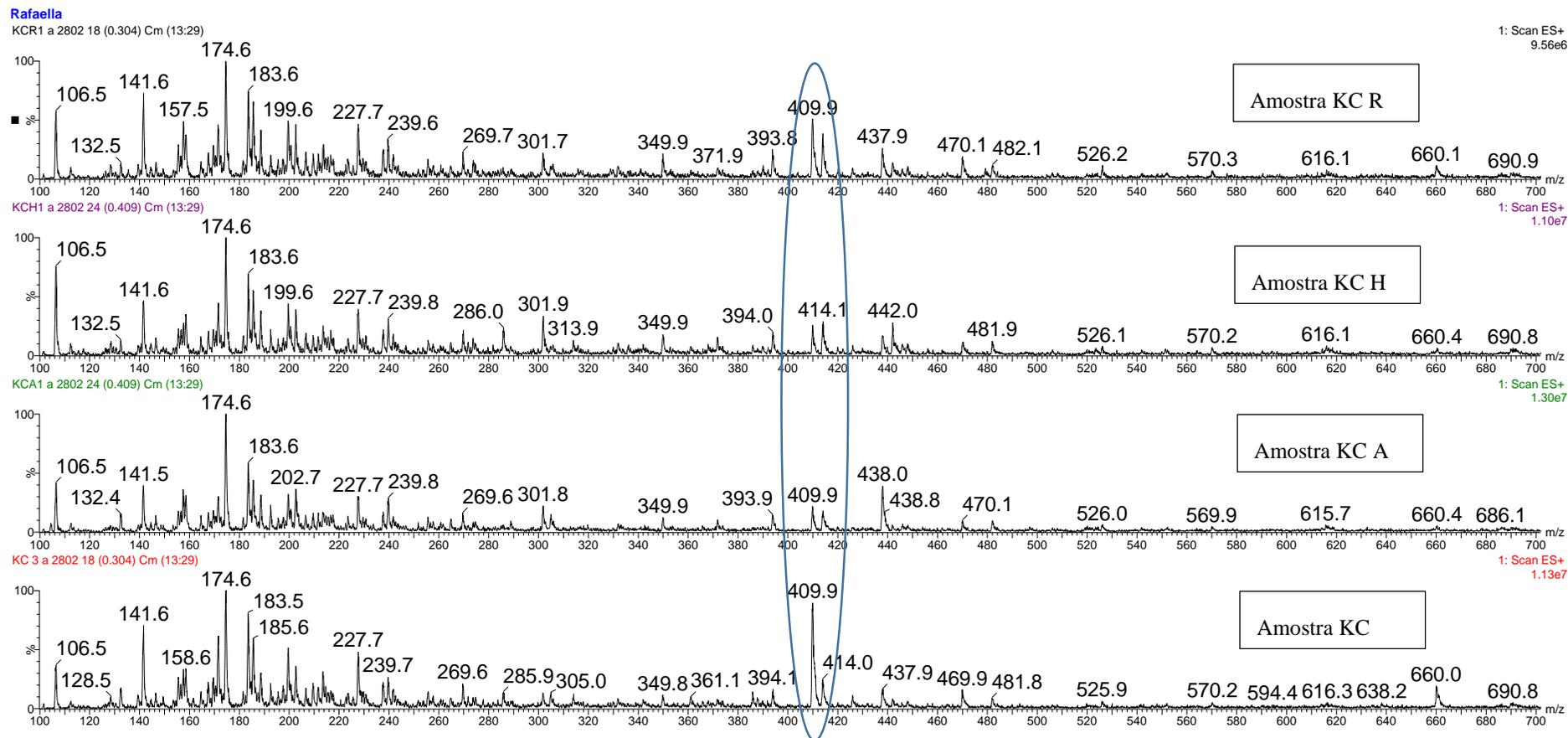


Figura 31. ESI(+)-MS fingerprints das amostras de KC1018. Os íons que se destacaram nestas amostras estão em evidência.

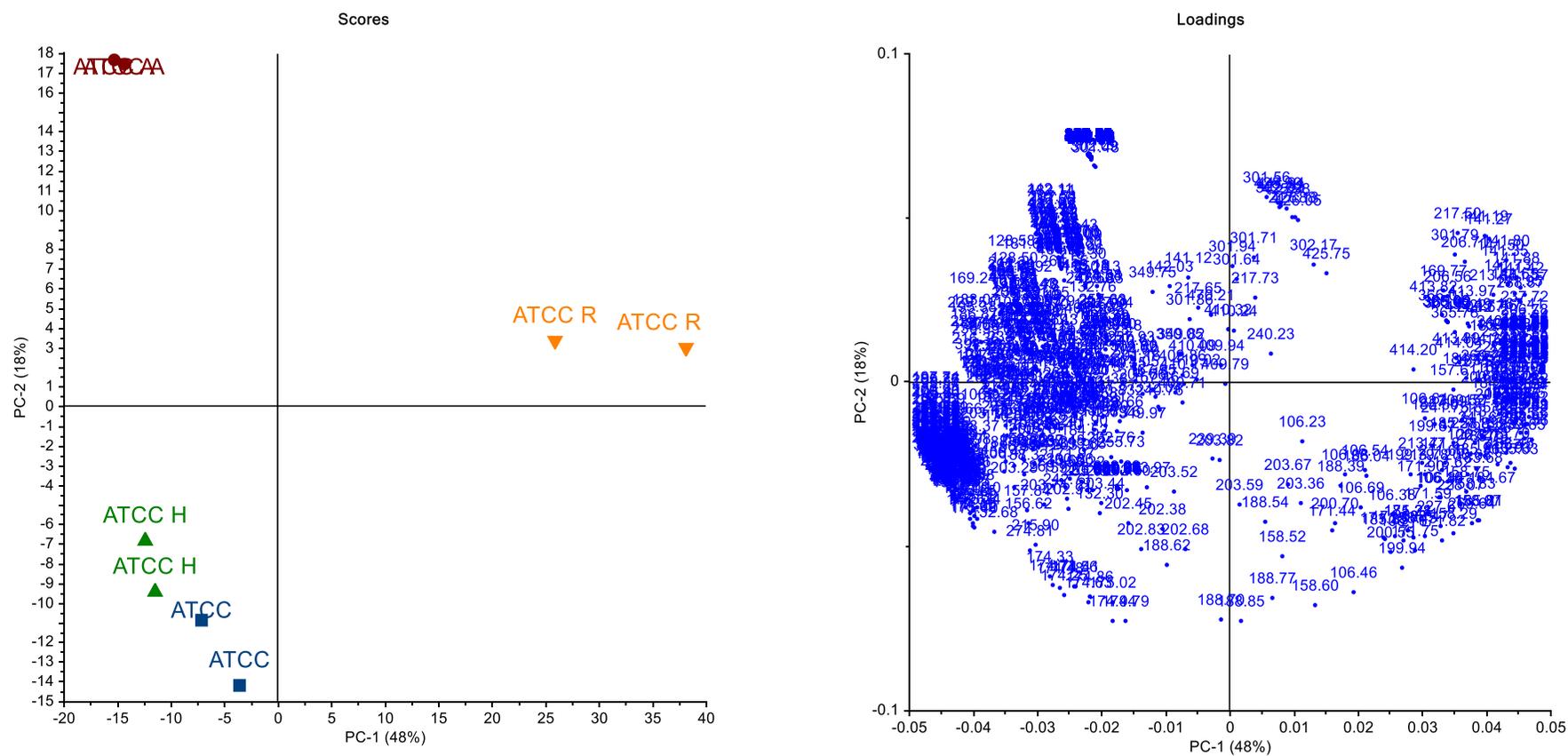


Figura 32. PCA dos dados de ESI(+)-MS das amostras de ATCC 11116: *Scores* – amostras e *loadings*- íons. É possível observar a separação das amostras de origem de infecção em células A549 (A) e HaCaT (R).

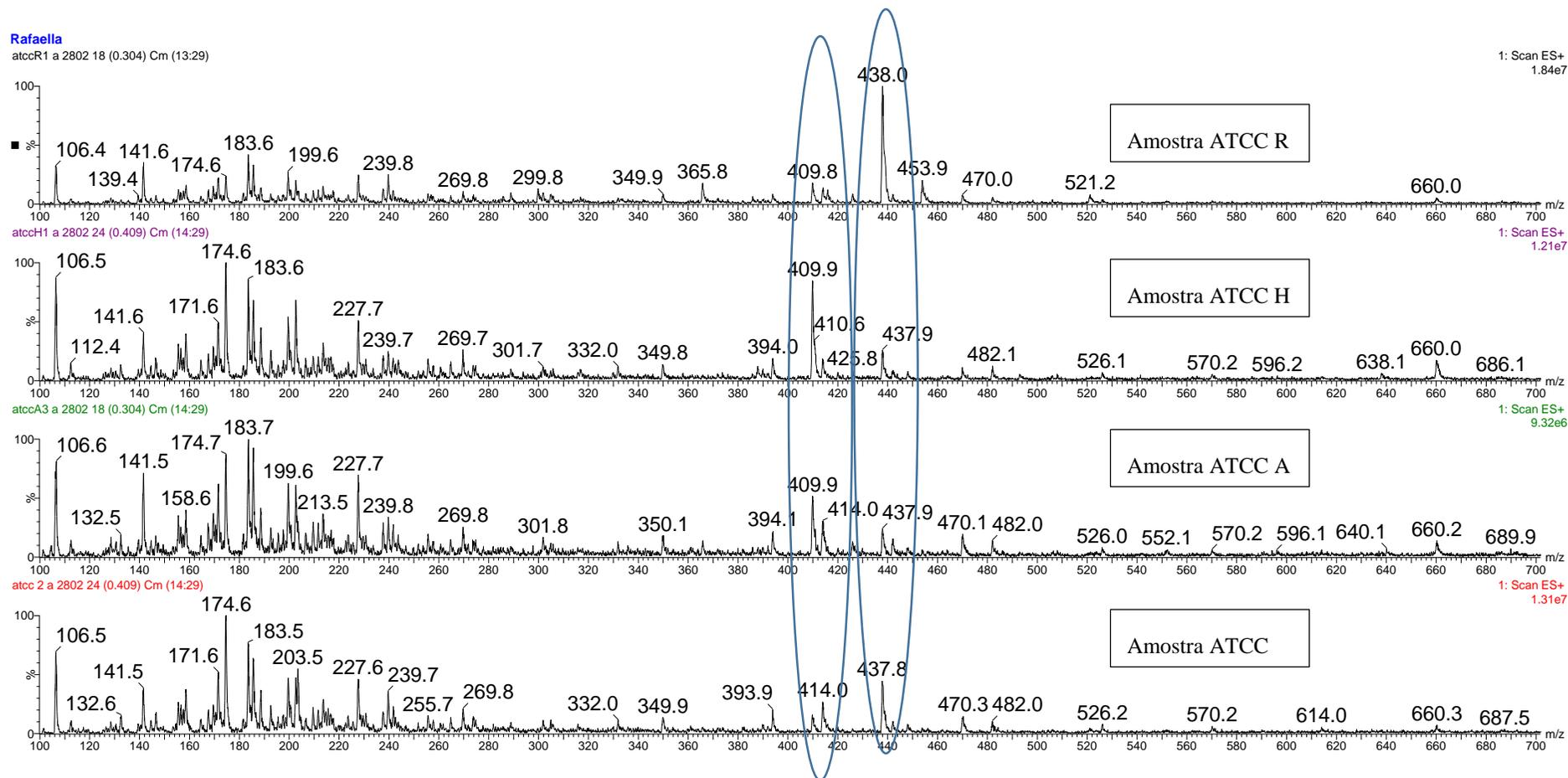


Figura 33. ESI(+)-MS *fingerprints* das amostras de ATCC 11116. Os íons que se destacaram nestas amostras estão em evidência.

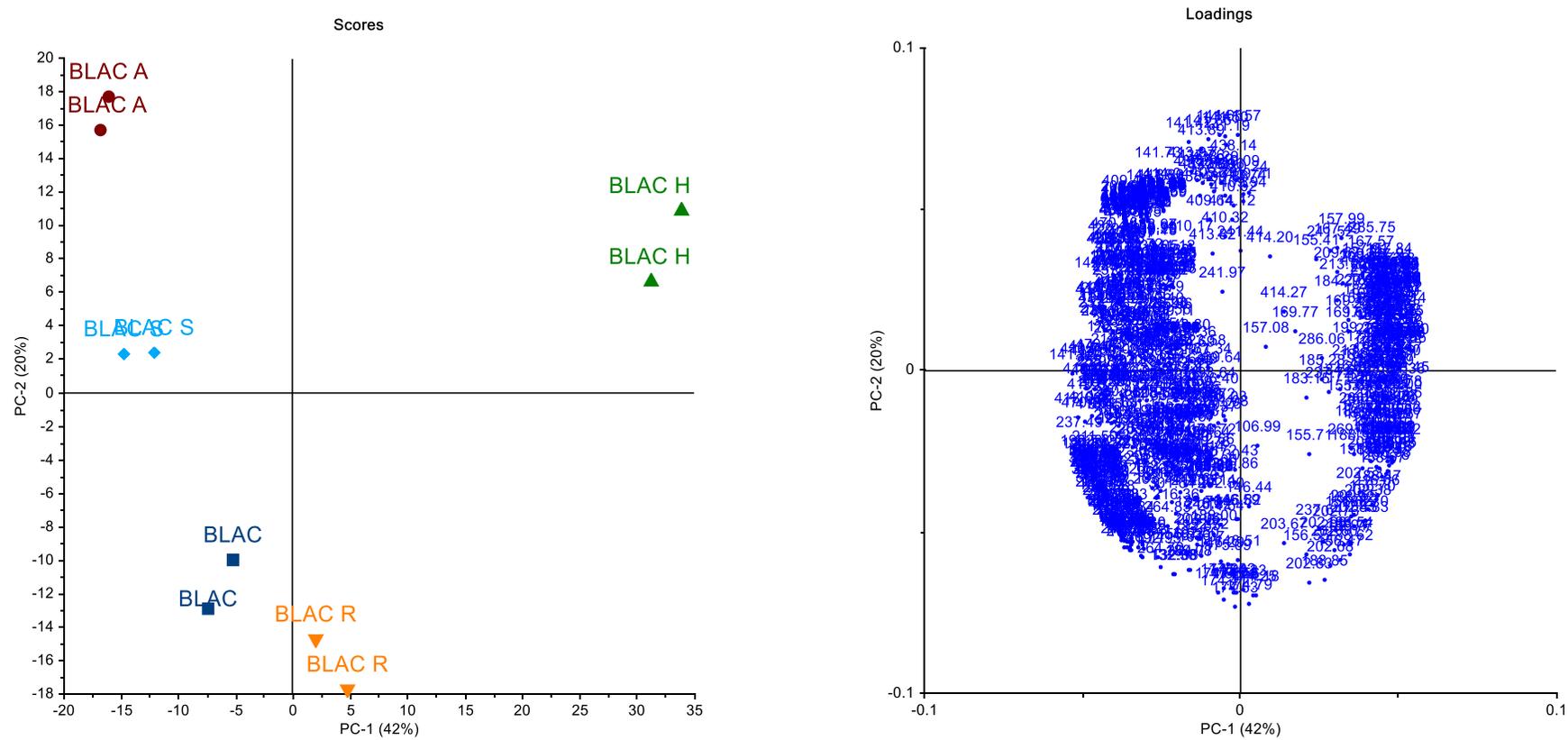


Figura 34. PCA dos dados de ESI(+)-MS das amostras de β -lac: *Scores* – amostras e *loadings* – íons. Não é possível observar a separação das amostras, exceto as de origem de infecção em células Hec-1b (H).

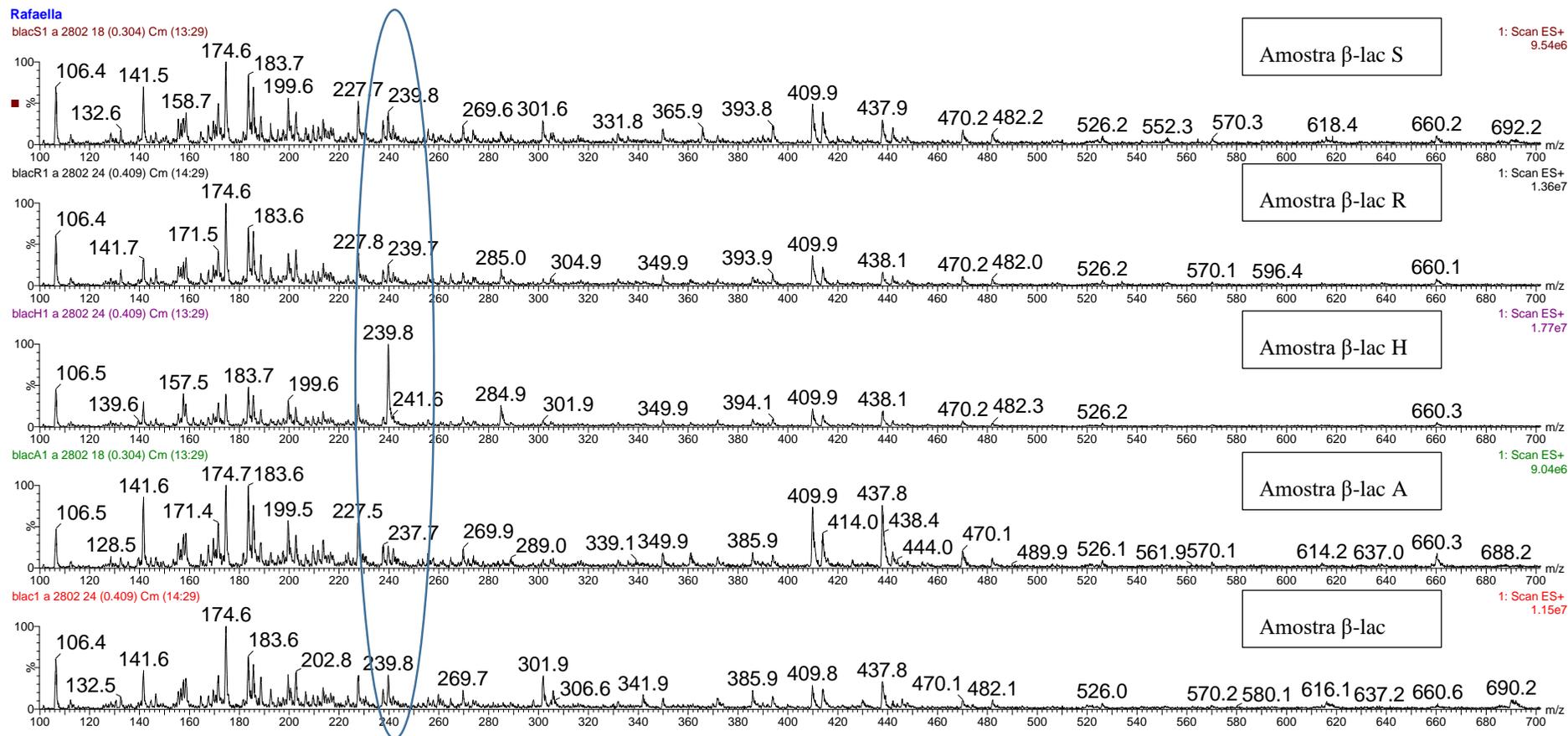


Figura 35. ESI(+)-MS *fingerprints* das amostras de β -lac. Os íons que se destacaram nestas amostras estão em evidência.

5. Dados ESI (-) MS

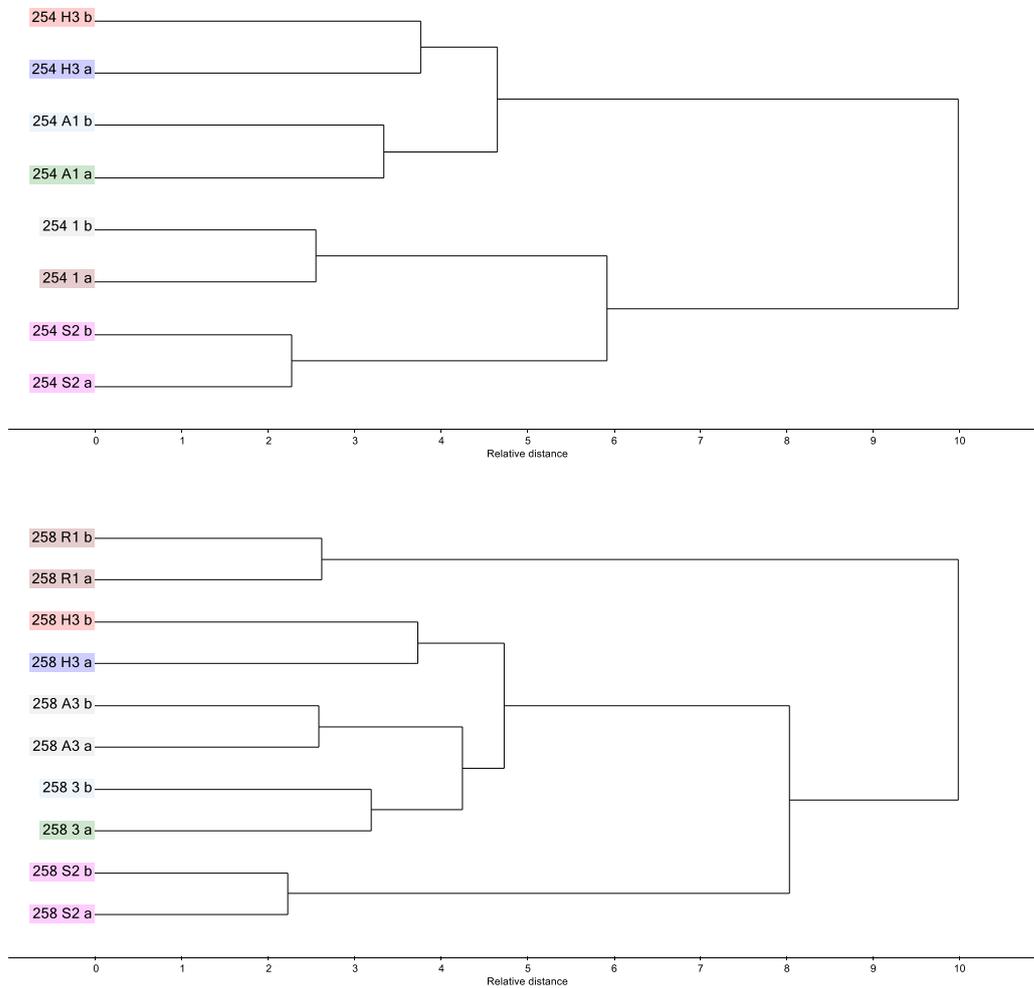


Figura 36. HCA dos dados de ESI(-)MS de Hae-BPF 254/86 e 258/86 segundo o método de Ward e distância Euclidiana ao quadrado.. Pode-se observar o agrupamento das amostras em duplicata do estoque, da infecção em A549 (amostras A), da infecção em Hec-1b (amostras H), da infecção em HaCaT (amostras R) e da infecção em Sprague-Dawley (amostras S).

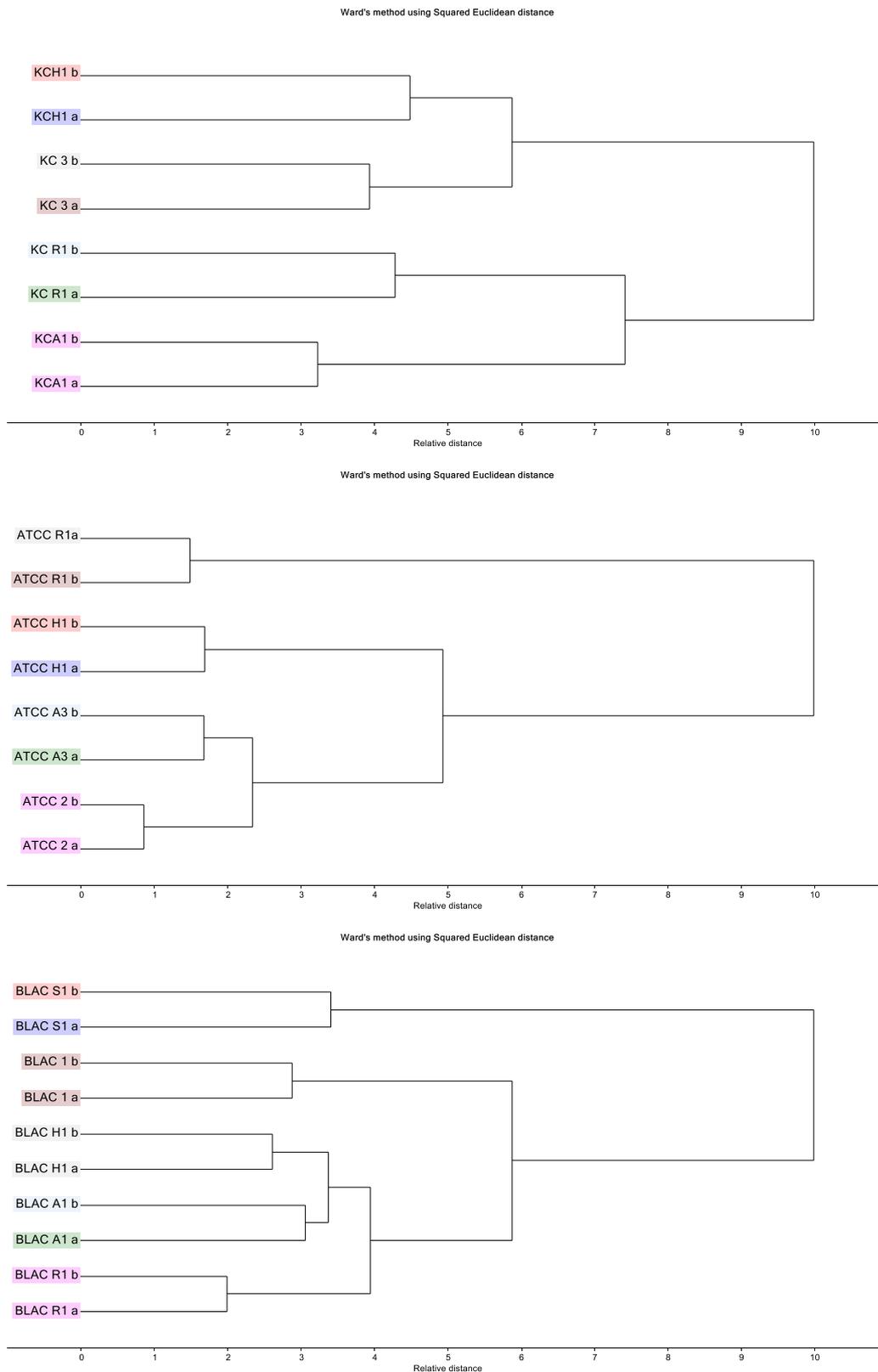


Figura 37. HCA dos dados de ESI(-)MS de Hae KC1018 e ATCC 11116 e Hi β -lac segundo o método de Ward e distância Euclidiana ao quadrado. Pode-se observar o agrupamento das amostras em duplicata do estoque, da infecção em A549 (amostras A), da infecção em Hec-1b (amostras H) e da infecção em HaCaT (amostras R) e, no caso de Hi β -lac, também da infecção em Sprague-Dawley (amostras S).

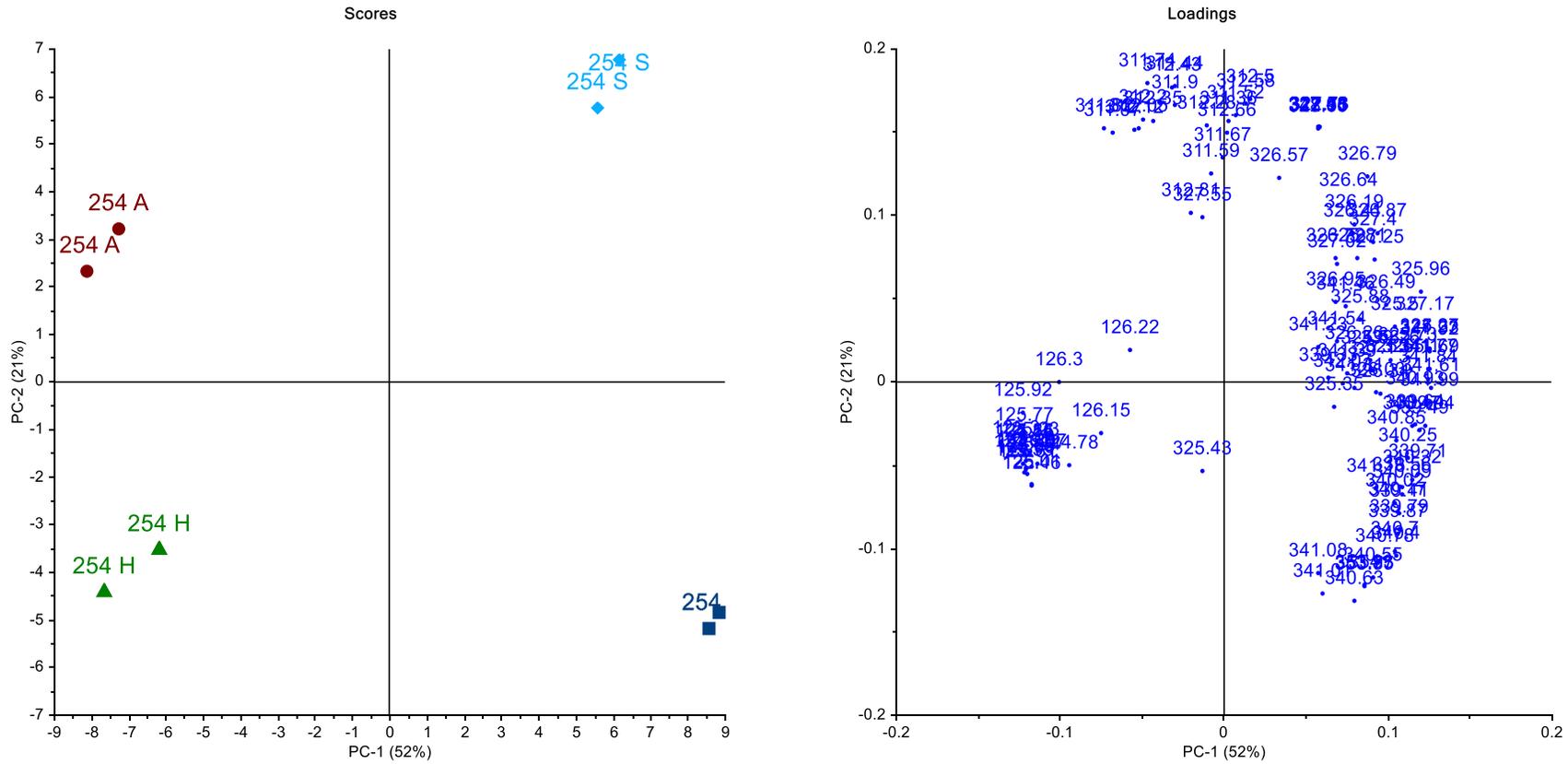


Figura 38. PCA dos dados de ESI(-)MS das amostras de 254/86: *Scores* – amostras e *loadings* – íons. Pode-se observar a separação das amostras.

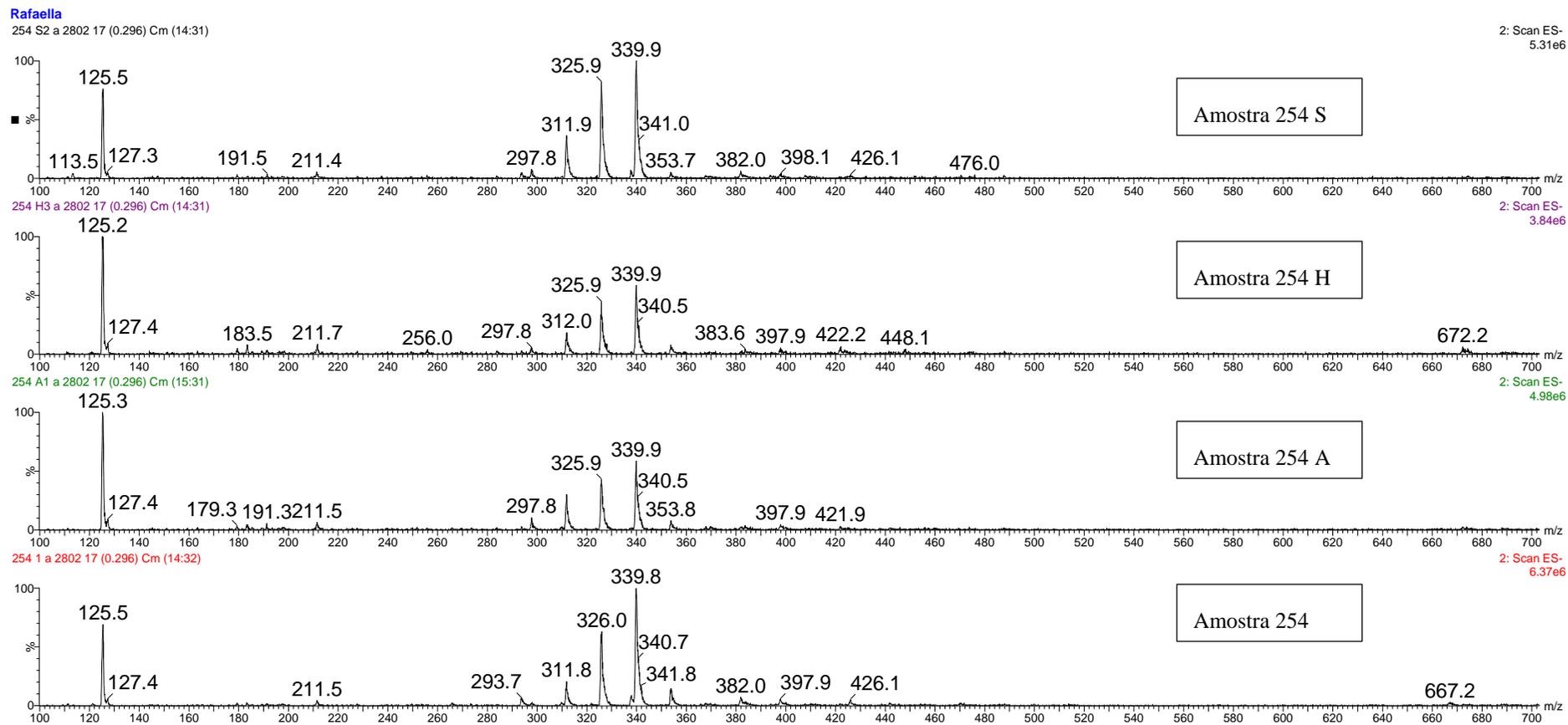


Figura 39. ESI(-)MS fingerprints das amostras de 254/86.

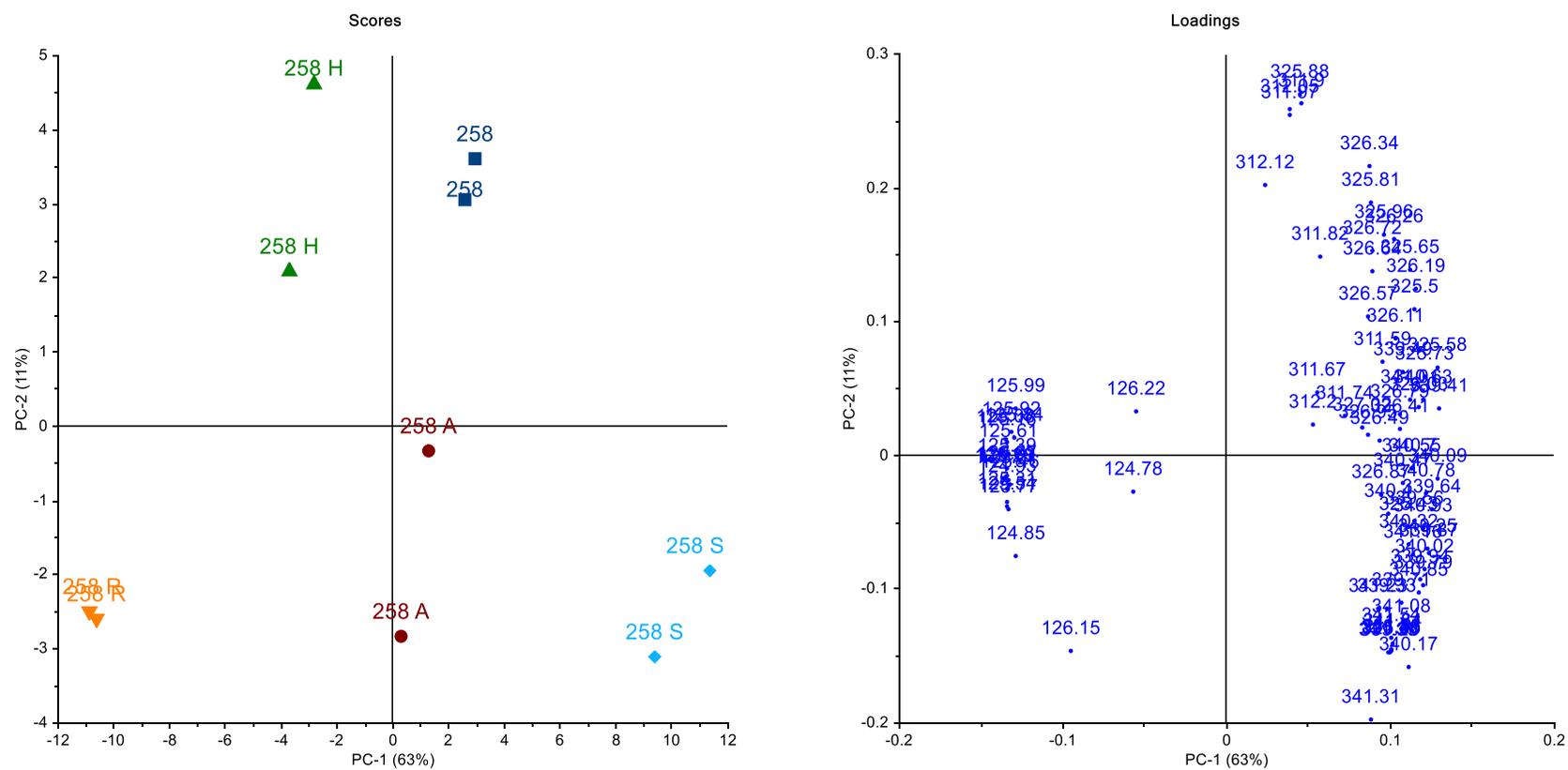


Figura 40. PCA dos dados de ESI(-)MS das amostras de 258/86: *Scores* – amostras e *loadings* – íons.

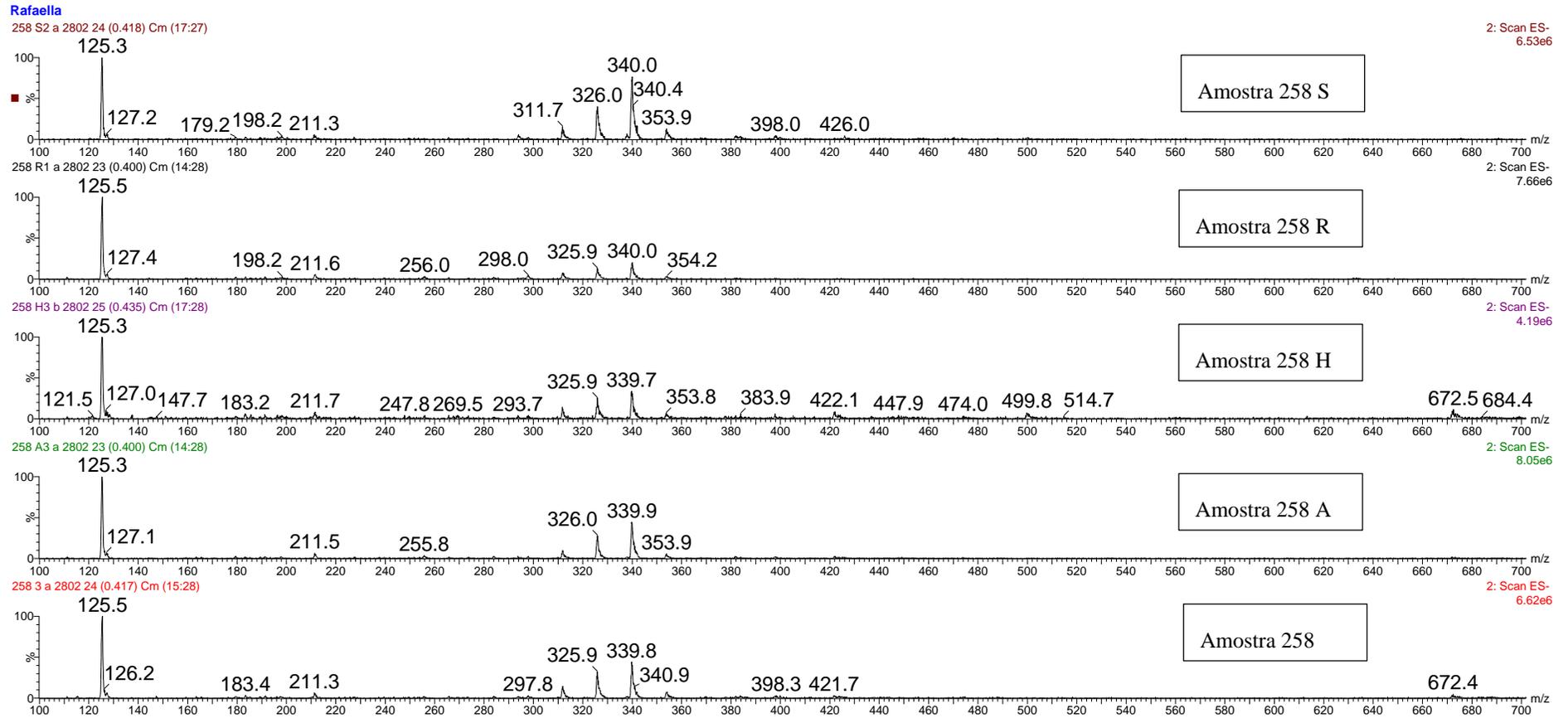


Figura 41. ESI(-)MS *fingerprints* das amostras de 258/86.

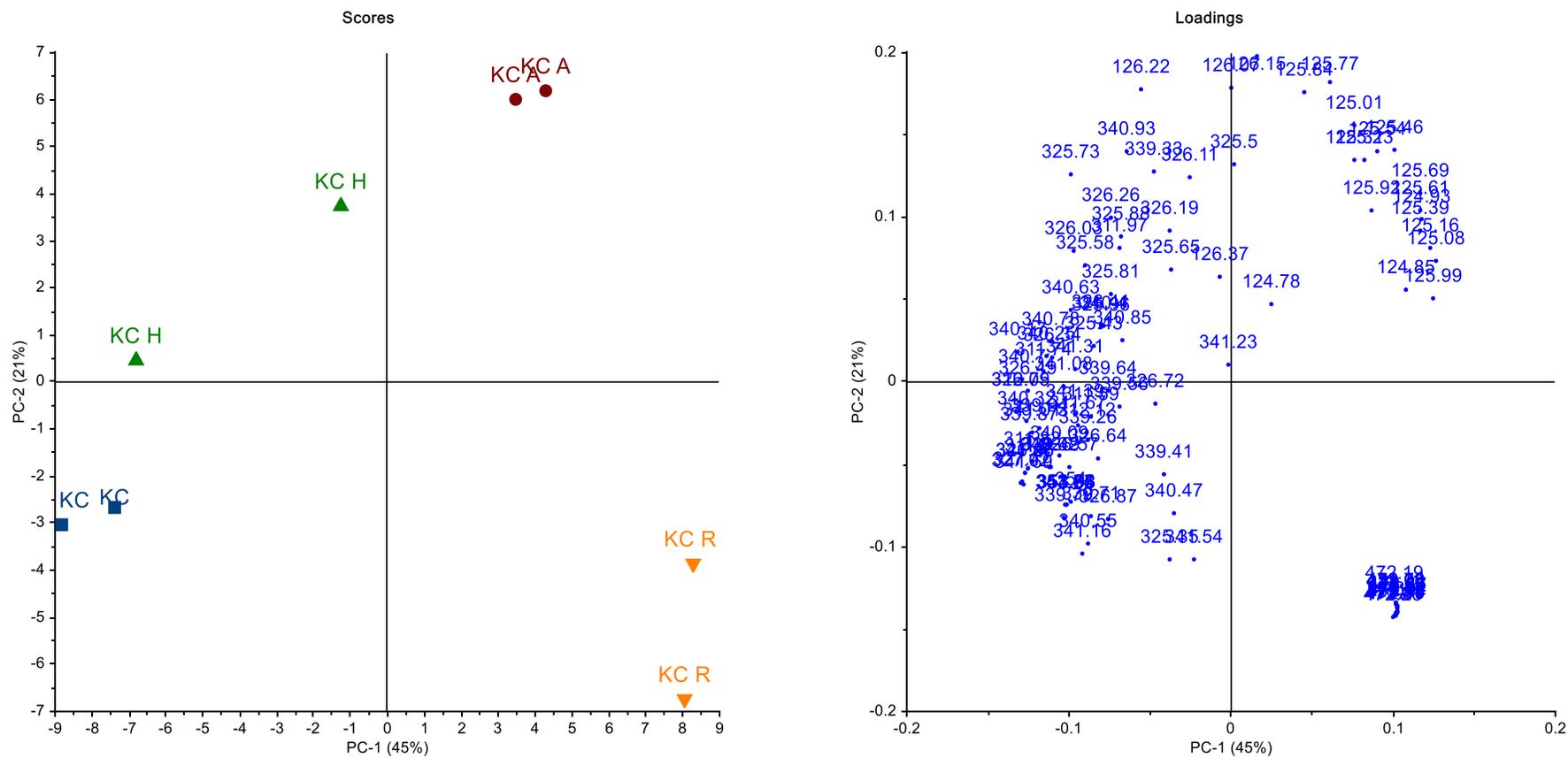


Figura 42. PCA dos dados de ESI(-)MS das amostras de KC1018: *Scores* – amostras e *loadings* – íons.

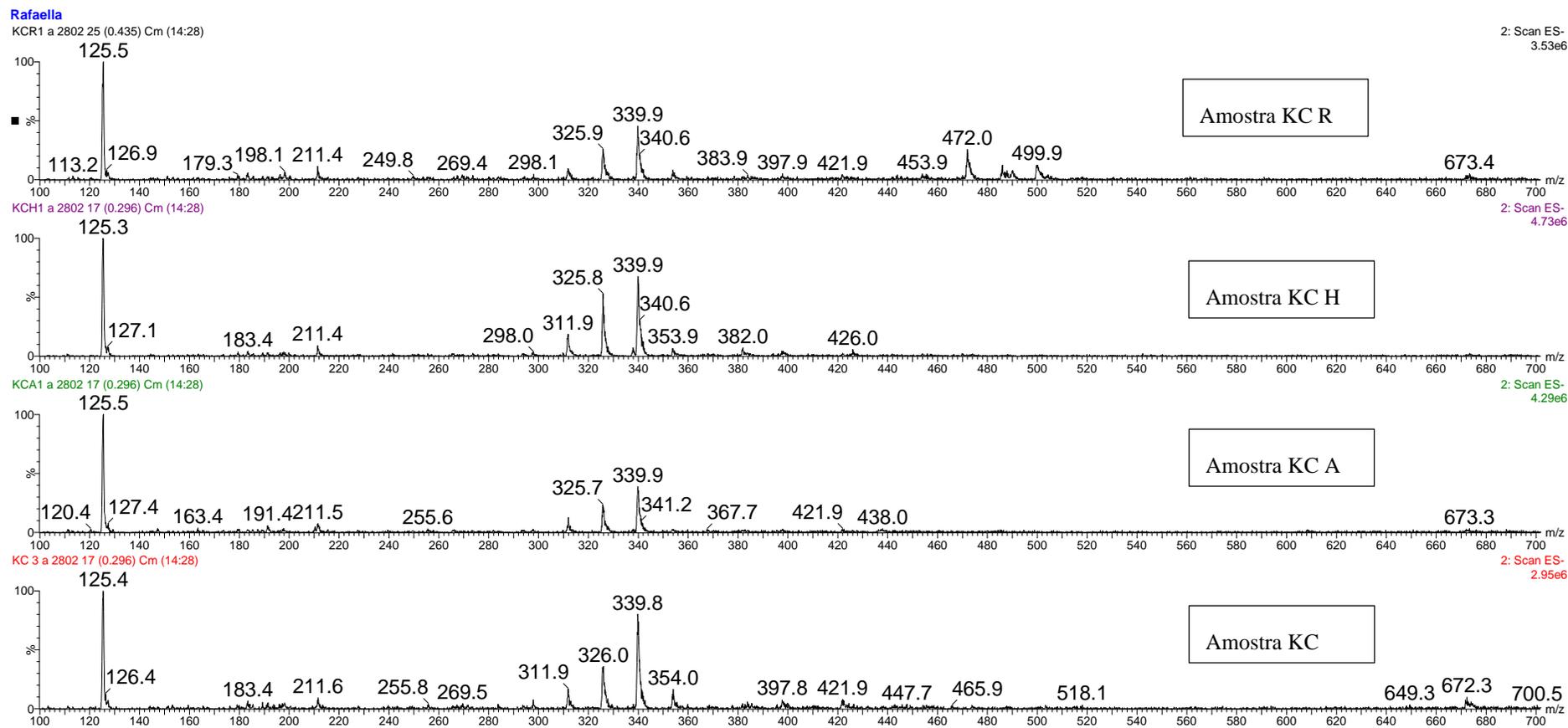


Figura 43. ESI(-)-MS fingerprints das amostras de KC1018.

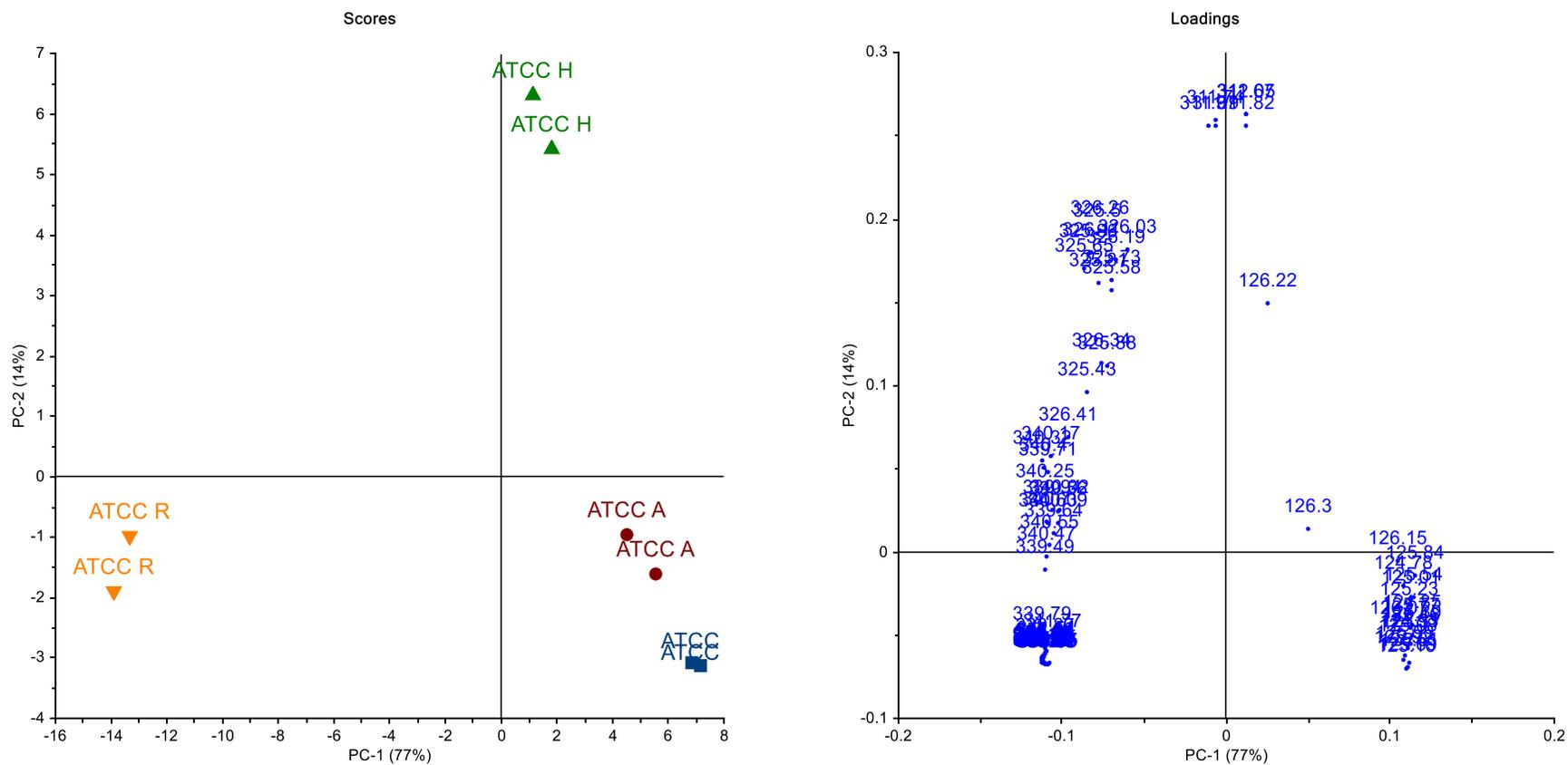


Figura 44. PCA dos dados de ESI(-)MS das amostras de ATCC 11116: *Scores* – amostras e *loadings* – íons. Pode-se observar a separação das amostras de infecção de células Hec-1b (H).

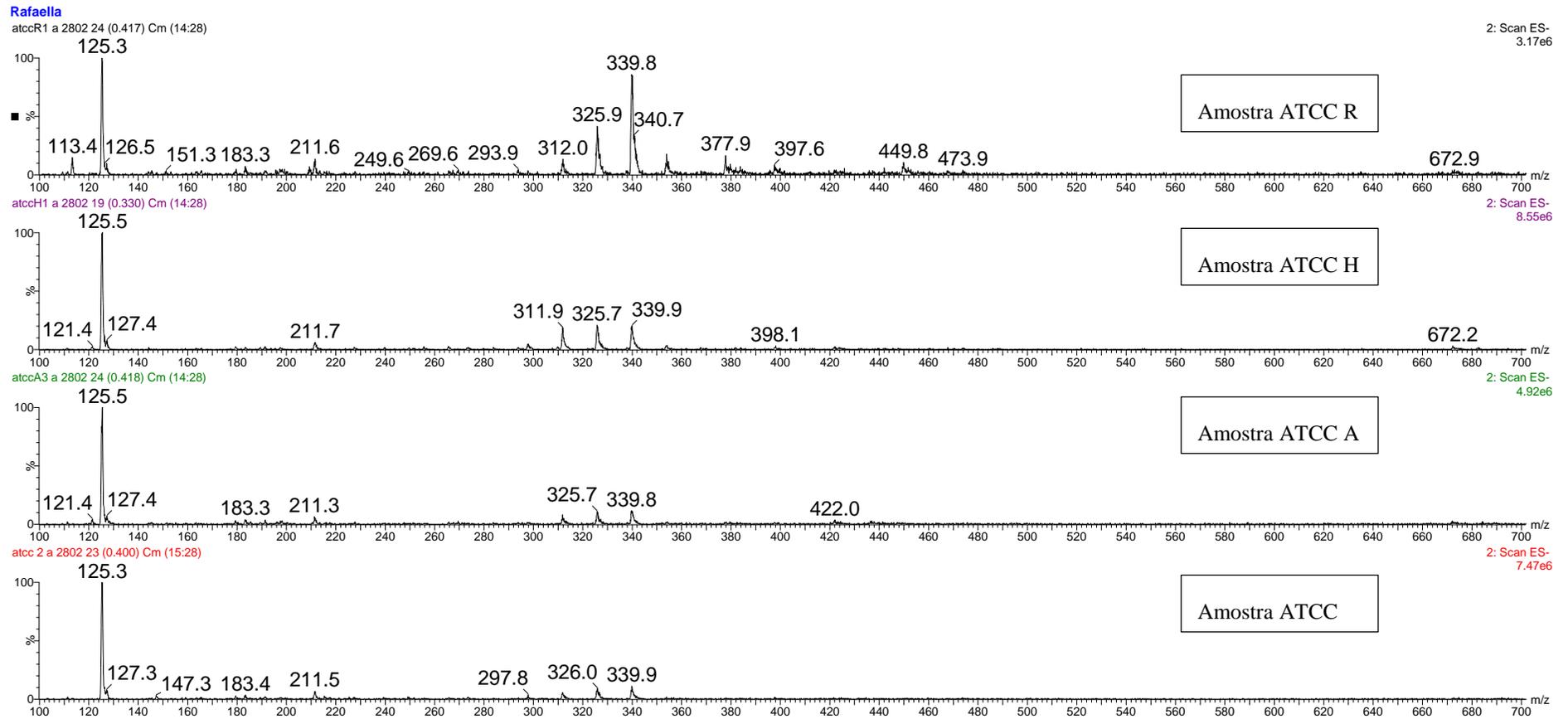


Figura 45. ESI(-)MS *fingerprints* das amostras de ATCC 11116.

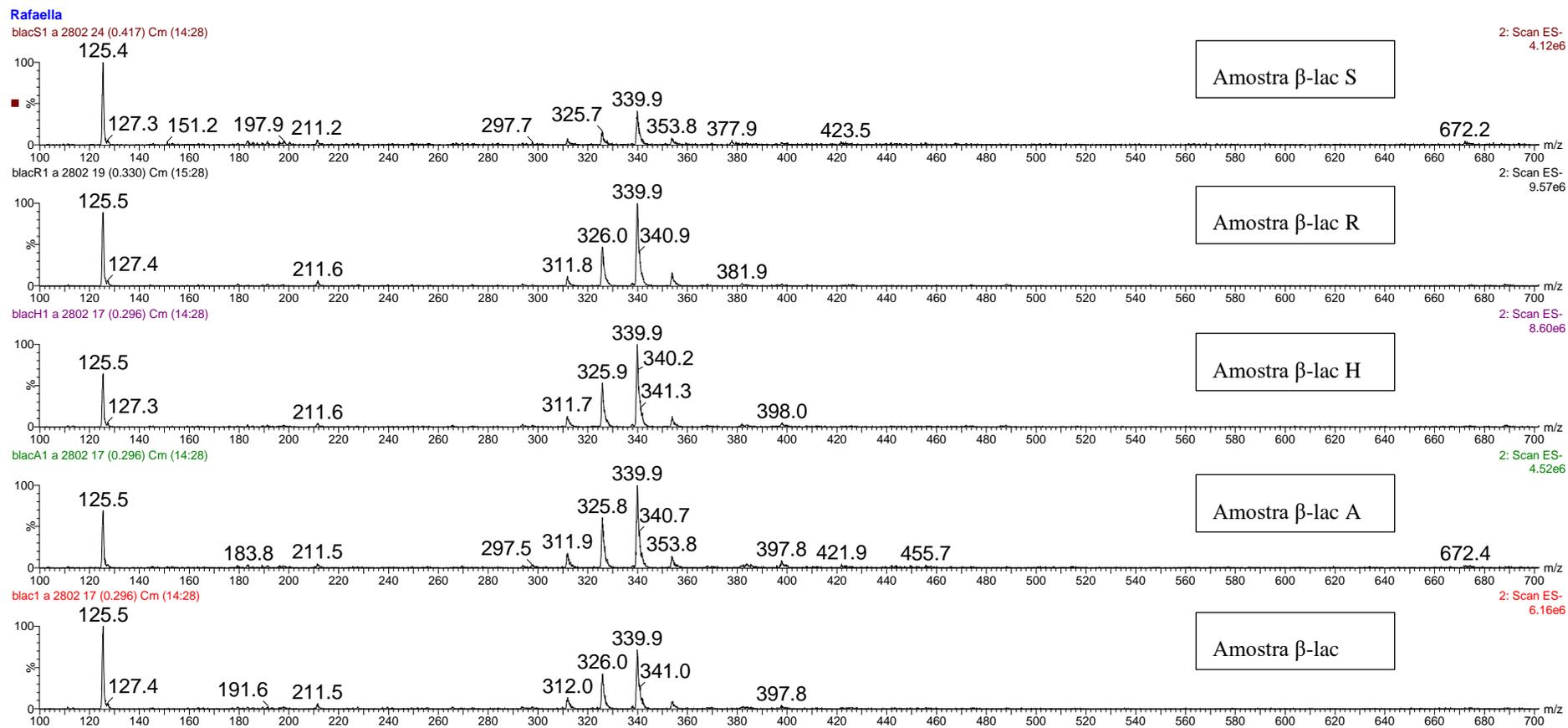


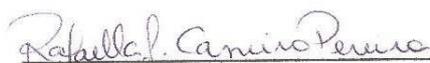
Figura 47. ESI(-)-MS fingerprints das amostras de β -lac.

6. Declaração

Profa. Dra. Rachel Meneguello
Presidente
Comissão Central de Pós-Graduação
Declaração

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada "Estudo do genoma e virulência de linhagens de *Haemophilus influenzae* causadoras de Febre Purpúrica Brasileira", não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 19/05/2015



Autor

RG n.º 44083773-2



Orientador

RG n.º 26217183-1

7. *Curriculum* durante o doutorado

Publicações

- Varela, J.N., Amstalden, M.C.K., Pereira, R.F.C., Hollanda, L.M., Ceragioli, H.J. Baranauskas, V., Lancellotti, M. Haemophilus influenzae porine ompP2 gene transfer mediated by graphene oxide nanoparticles with effects on transformation process and virulence bacterial capacity. *Journal of Nanobiotechnology* v. 2, p. 14, 2014.
- Cury, G.C.G., Pereira, R.F.C., Hollanda, L.M., Lancellotti, M. Inflammatory response of Haemophilus influenzae biotype aegyptius causing Brazilian Purpuric Fever. *Brazilian Journal of Microbiology (Impresso)*, v. 45, p. 1449-1454, 2014.
- Izidoro Júnior, M.S., Varela, J.N., Alves, D.A., Pereira, R.F.C., Brocchi, M., Lancellotti, M., Hollanda, L.M. Effects of Salmonella enteritidis serovar typhimurium Infection in Adenocarcinomic Human Alveolar Basal Epithelial Cells A549 In vitro: Bacteria Induce Apoptosis in Adenocarcinomic Cell. *Journal of Bacteriology & Parasitology*, v. 3, p. 1000158, 2012.
- Pereira, R.F.C., Alves, D.A., Jacinto, R. K., Hollanda, L.M., Verinaud, L.M.C., Lancellotti, M. Effects of Neisseria meningitidis Infection in Tumor Glioblastoma Cell Line NG97: Respiratory Pathogen Inducing Apoptosis. *Journal of Bacteriology & Parasitology*, v. 02, p. 2-6, 2011.
- Hollanda, L.M., Cury, G.C.G., Pereira, R.F.C., Ferreira, G.A, Sousa, A., Sousa, E. M.B., Lancellotti, M. Effect of mesoporous silica under Neisseria meningitidis transformation process: environmental effects under meningococci transformation. *Journal of Nanobiotechnology*, v. 9, p. 28, 2011.

Capítulo de livro

- Shishido, S. M., Cury, G.C.G., Pereira, R.F.C., Mattos, I.B., Alves, D.A., Varela, J. N., Izidoro Júnior, M.S., Amstalden, M. C. K., Esteves, C.Z., Lancellotti, M., Hollanda, L.M., NANOTECHNOLOGY: Diagnostics and Therapeutics. *NANOTECHNOLOGY*. 1ed., 2013, v. , p. 4500-.

Patente

- Lancellotti, M.; Pereira, R.F.C. Kit de mimetismo de Barreira Hemato-encefálica Humana, 2011, Brasil. Patente: Privilégio de Inovação. Número do registro: 018110022861, data de depósito: 17/06/2011. FAPESP 2008/56778-1.

Programa Estágio Docente (PED)

- Participação no PED na disciplina de Biologia- Microbiologia Básica (BM320) para o curso de Engenharia de Alimentos da UNICAMP sob a supervisão da profa. Clarice Weis Arns, no primeiro semestre de 2014.

Experiência Didática

- Professora Auxiliar nas disciplinas de Bioquímica Clínica, Genética Molecular, Microbiologia e Parasitologia Clínica do curso de Farmácia da Faculdade de Jaguariúna Campus II (Fevereiro - Agosto de 2014).
- Cultura de Células: Conceitos Básicos e Aplicações. Curso de 04 horas na Universidade de Marília. 2011.

- Aula para o curso de Farmácia “PCR e Real Time PCR”. 2013. FR605 – Biologia Molecular.
- Aula para o curso de Farmácia “Introdução à Genômica e Bioinformática”. 2013. FR725 – Biotecnologia.
- Aula para o curso de Farmácia “Aplicações da Genômica à Biotecnologia Farmacêutica”. 2013. FR900 – Biotecnologia Microbiana na Indústria Farmacêutica.

Cursos

- Fundamentos da PCR Quantitativa em Tempo Real. Life Technologies Brasil. 2014.
- Bioinformática para o Microbiologista. Sociedade Brasileira de Microbiologia. 2013
- Transcriptional Regulation and Transcriptome Analy. Sociedade Brasileira de Biologia Celular. 2012.
- Genomas de Procaríotos: Next Gen e Bioinformática. Sociedade Brasileira de Microbiologia. 2011.
- Tuberculose: Atualidade e Perspectivas. Sociedade Brasileira de Microbiologia. 2011.

Participação em congressos

- 10th International Congress on Cell Biology and the XVI Meeting of the Brazilian Society for Cell Biology. 2012. Rio de Janeiro, Brasil.
- 22nd IUBMB & 37th FEBS Congress. 2012. Sevilha, Espanha.
- Congresso Brasileiro de Microbiologia. 2011. Foz do Iguaçu, Brasil.

Participação em Banca de TCC

- Duarte, M. C. T.; Figueira, G. M.; Pereira, R.F.C. Participação em banca de Beatriz Degam Klemz. Comparação entre métodos qualitativo e quantitativo para determinação de atividade antimicrobiana de óleos essenciais. 2012. Trabalho de Conclusão de Curso (Graduação em Farmácia) - Universidade Estadual de Campinas.
- Lancellotti, Marcelo; Hollanda, Luciana Maria de; Pereira, R.F.C. Participação em banca de Maria Cecília Krähenbühl Amstalden. Construção do vetor plasmidial para expressão e deleção do gene luxS de *Streptococcus pneumoniae*. 2011. Trabalho de Conclusão de Curso (Graduação em Farmácia) - Universidade Estadual de Campinas.

Participação em outros projetos

- Ana Carolina Afonso Silva (Graduanda em Farmácia). Análise *in vitro* do papel do plasmídeo constitutivo de *Haemophilus influenzae* biótipo aegyptius causador de Febre Purpúrica Brasileira. 2013. Iniciação Científica. - Universidade Estadual de Campinas, Conselho Nacional de Desenvolvimento Científico e Tecnológico. Orientador: Marcelo Lancellotti.