

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

Monique Ribeiro Tiba



**“IDENTIFICAÇÃO DE NOVOS ANTÍGENOS FLAGELARES
DE *ESCHERICHIA COLI* DE ORIGEM HUMANA”**

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato (a)
Monique Ribeiro Tiba
[Assinatura]
e aprovada pela Comissão Julgadora.

Tese apresentada ao Instituto de
Biologia para obtenção do Título
de Doutor em Genética e Biologia
Molecular, na área de
Microbiologia.

Orientador: Prof. Dr. Domingos da Silva Leite

Campinas, 2009

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP

T432i	<p>Tiba, Monique Ribeiro Identificação de novos antígenos flagelares de <i>Escherichia coli</i> de origem humana / Monique Ribeiro Tiba. – Campinas, SP: [s.n.], 2009.</p> <p>Orientador: Domingos da Silva Leite. Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.</p> <p>1. <i>Escherichia coli</i>. 2. Flagelos (Microbiologia) - Antígenos. 3. Sorologia. 4. Sequenciamento de DNA. I. Leite, Domingos da Silva, 1960-. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.</p> <p>(rcdi/ib)</p>
-------	---

Título em inglês: Identification of new *Escherichia coli* flagellar antigen from human origin.

Palavras-chave em inglês: *Escherichia coli*; Serology; Flagella (Microbiology) - Antigens; DNA sequencing.

Área de concentração: Microbiologia.

Titulação: Doutora em Genética e Biologia Molecular.

Banca examinadora: Domingos da Silva Leite, Fabiana Fantinatti Garbognini, Marcelo Lancelloti, Gleize Villela Carbonell, Mario Paulo Amante Penatti.

Data da defesa: 24/08/2009.

Programa de Pós-Graduação: Genética e Biologia Molecular.

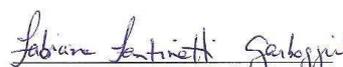
Campinas, 24 de agosto de 2009.

BANCA EXAMINADORA

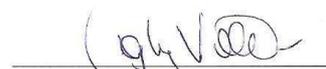
Prof. Dr. Domingos da Silva Leite. (Orientador)


Assinatura

Profa. Dra. Fabiana Fantinatti Garboggini


Assinatura

Profa. Dra. Gleize Vilella Carbonell.


Assinatura

Prof. Dr. Marcelo Brocchi.

Assinatura

Prof. Dr. Mario Paulo Amante Penatti.


Assinatura

Prof. Dr. Marcelo Lancelloti.


Assinatura

Profa. Dra. Elisabeth Pelosi Teixeira.

Assinatura

Profa. Dra. Cleide Ferreira Catani.

Assinatura

"Para realizar grandes conquistas, devemos não apenas agir, mas também sonhar; não apenas planejar, mas também acreditar."

(Anatole France)

À André Casas

"O amor é sofredor, é benigno; o amor não é invejoso; o amor não se vangloria, não se ensoberbece, não se porta inconvenientemente, não busca os seus próprios interesses, não se irrita, não suspeita mal; não se regozija com a injustiça, mas se regozija com a verdade; tudo sofre, tudo crê, tudo espera, tudo suporta..."

(Coríntios 13: 4-7)

Agradecimentos

A Deus por iluminar meu caminho e me fazer seguir sempre em frente.

Aos meus pais Dario e Regina e irmãos Flávio e Kenny, pelo amor, incentivo e colaboração fundamental para o desenvolvimento e conclusão desta etapa de minha vida.

A minha avó pelas orações.

Aos queridos Dário, Clarice, João e Ana que tanto me ajudaram neste começo de carreira, pelo carinho, amor e apoio incondicional.

A Lisa Casas por sempre fazer parte de minha vida, pelo carinho, amor incondicional. Agradeço também a Biba, Ozzyto e Bono.

Ao meu orientador, Prof. Dr. Domingos da Silva Leite, por ter me recebido como sua orientada e ter confiado no meu trabalho.

À Prof. Dra. Maria Sílvia Viccari Gatti, pelo apoio, paciência e incentivo durante a realização deste trabalho e participação no PED.

Ao Prof. Dr. Marcelo Brocchi pelo apoio e ensinamento das técnicas deste trabalho e pela participação na pré-banca.

Ao Prof. Dr. Tomomasa Yano pelo apoio e incentivo.

Aos meus amigos do laboratório que nunca esquecerei Dani, Cláudia, Geórgio, Mônica, Juliana, Maria Clara, Dan, Lus, Tati, Gerson, Mário, Mi, Alinne, Alex, Carol, pelos incentivos, auxílios e companheirismo.

À bióloga e amiga Mirtis, pelo apoio, ensinamento e disponibilidade durante o desenvolvimento deste trabalho.

Ao Erivaldo, pela disponibilidade e apoio técnico fundamental para a execução deste trabalho.

À Ana, pelo apoio técnico, ensinamentos e alegrias.

Ao Marcelo Falsarella Carazzolle do Laboratório de Genômica pelo auxílio no sequenciamento.

Aos meus antigos e novos amigos Alinão, Michelle, Roberta, Gabi, Davi, Li, Gui, Lilian, Rê, Murphy, Biju, Bio, que mesmo à distância, sempre me incentivaram.

A CAPES pela concessão da bolsa de doutorado e a FAPESP pelo apoio financeiro concedido.

A todos os professores e funcionários do Departamento de Microbiologia e Imunologia e do curso de pós-graduação em Genética e Biologia Molecular.

Agradeço a todos que diretamente ou indiretamente contribuíram para a realização deste trabalho.

I- Introdução	17
1-<i>Escherichia coli</i>	18
2-Estruturas flagelares	21
2.1-Filamento	22
2.2-Gancho	25
2.3- Corpo basal	26
3- O sistema gênico do flagelo	29
4- Variação de fase flagelar	33
5- Sorologia	39
6- PCR-RFLP	41
II-Objetivos	45
1-Geral	46
III- Material e Métodos	47
3.1-Amostras de <i>Escherichia coli</i>	48
3.1.1: Amostras de <i>E. coli</i> padrões para antígeno flagelar “H”	48
3.1.2: Amostras de <i>E. coli</i> com antígeno H não-tipável (HNT)	48
3.2 -Extração de DNA bacteriano	49
3.3-Reação em Cadeia da Polimerase (PCR)	50
3.3.1: Iniciadores utilizados para o estudo da variação de fase	51
3.4-Eletroforese em gel de agarose	52
3.5-Análise dos resultados da eletroforese em gel de agarose	52
3.6-Digestão enzimática	52
3.6.1:Preparo do DNA	53
3.6.2: Digestão enzimática	53
3.6.3: Eletroforese em gel de agarose	53
3.7- Identificação dos padrões de tipos dos antígenos H	54
3.8- Sequenciamento	54

3.8.1: Sequenciamento pela técnica de “cromossomo walking ou primer walking”	54
3.9: Clonagem do gene <i>fliC</i>	55
3.9.1: Inserção do gene <i>fliC</i> no vetor de clonagem.....	55
3.9.2: Obtenção de células competentes.....	55
3.9.3: Transformação por heat-choque.....	56
3.9.4: Extração plasmidial.....	56
3.9.5: Sequenciamento dos clones.....	57
3.9.6: Montagem dos “contigs” e análises das sequências	59
3.10- Preparo e obtenção do antígeno H das cepas de <i>E. coli</i> HNT	61
3.10.1: Ativação da motilidade.....	61
3.10.2: Obtenção do antígeno H.....	61
3.10.3: Obtenção de anti-soros anti-antígenos HNT.....	61
3.10.4: Aglutinação em tubo.....	62
3.10.5: Soroabsorção do anti-soro do antígeno H das amostras HNT.....	62
3.11 - Análise da estrutura secundária da proteína flagelina	63
IV- Resultados	64
4.1: Reação em cadeia da polimerase (PCR)	65
4.2: Análise da digestão enzimática	69
4.3: PCR dos genes de variação de fase	79
4.4: Sorologia	79
4.5: Sequenciamento	83
4.5.1: Amostra 1C.....	84
4.5.2: Amostra 5C.....	90
4.5.3: Amostra 8C.....	95
4.5.4: Amostra 9C.....	101
4.5.5: Amostra 11A.....	106
5: Análise da estrutura secundária	115

Índice Geral

V- Discussão	116
VI- Conclusões	134
VII- Referências Bibliográficas	136
VIII- Anexos	146
8.1: Anexo1.....	147
8.2: Anexo2.....	149

Índice de Figuras

- Figura 1:** Componentes flagelares da *Salmonella enterica*. A figura mostra as estruturas presentes na parede celular e na membrana celular. Componentes citoplasmáticos solúveis como FliH, FliL, FliJ, complexo ATPase, auxiliam na liberação dos substratos e na ordem da secreção dos substratos. A proteína do filamento consiste em FliC ou FljB que são alternativamente transcritas. Fonte: Chevance & Hughes (2008).....27
- Figura 2:** Resultado da montagem final do contig da cepa 1C através do programa Phred, Phrap, Consed.....60
- Figura 3:** Eletroesferogramas da amostra 1C, visualizados pelo programa ChromasPro....60
- Figura 4:** Perfil eletroforético em gel de agarose a 1,5% das amostras de *E. coli* padrão para antígeno H, utilizando o iniciador FliC(M). PM1: Padrão de peso molecular de 100 pb (Fermentas); PM2: Padrão de peso molecular de 1 Kb (Fermentas); Linha 1: *E. coli* H1; Linha 2: *E. coli* H2; Linha 3: *E. coli* H3; Linha 4: *E. coli* H4; Linha 5: *E. coli* H5; Linha 6: *E. coli* H6; Linha 7: *E. coli* H7; Linha 8: *E. coli* H8; Linha 9: *E. coli* H9; Linha 10: *E. coli* H10.66
- Figura 5:** Perfil eletroforético em gel de agarose a 1,5% das amostras de *E. coli* padrão para antígeno H, utilizando o iniciador FliC(F). PM2: Padrão de peso molecular de 1 Kb; Linha 1: *E. coli* H1; Linha 2: *E. coli* H2; Linha 3: *E. coli* H3; Linha 4: *E. coli* H5; Linha 5: *E. coli*

Índice de Figuras

H6; Linha 6: *E. coli* H7; Linha 7: *E. coli* H8; Linha 8: *E. coli* H9; Linha 9: *E. coli* H10;
Linha 10: *E. coli* H11; Linha 11: *E. coli* H12.66

Figura 6: Perfil eletroforético em gel de agarose 1,5% das amostras de *E. coli* com antígeno H não-tipável, utilizando o iniciador FliC(M). PM2: Padrão de peso molecular de 1 Kb; Linha 1: *E. coli* VTH-115 (1C); Linha 2: *E. coli* VTH-110 (2C); Linha 3: *E. coli* VTH-118 (3C); Linha 4: *E. coli* 28011a (4C); Linha 5: *E. coli* 33141a (5C); Linha 6: *E. coli* 46103b (6C); Linha 7: *E. coli* 40478b (7C); Linha 8: *E. coli* 48629c(2)(9C).....68

Figura 7: Perfil eletroforético em gel de agarose 1,5% das amostras de *E. coli* com antígeno H não-tipável, utilizando os iniciadores FliC(F). PM2: Padrão de peso molecular de 1 Kb; Linha 1: *E. coli* VTH-115 (1C); Linha 2: *E. coli* VTH-110 (2C); Linha 3: *E. coli* VTH-118 (3C); Linha 4: *E. coli* 28011a (4C); Linha 5: *E. coli* 33141a (5C); Linha 6: *E. coli* 46103b (6C); Linha 7: *E. coli* 40478b (7C); Linha 8: *E. coli* 48629c(2)(9C).....68

Figura 8: Perfil eletroforético da digestão dos amplicons obtidos com os iniciadores fliC(M) com a enzima *HhaI*, nas amostras de *E. coli* H padrão, gerado pelo programa GELCOMPAR II. PM50 pb – padrão de peso molecular de 50 pb (Invitrogen); dig – digerido.....70

Figura 9: Perfil eletroforético da digestão dos amplicons obtidos com os iniciadores FliC(F) com a enzima *RsaI*, nas amostras de *E. coli* H padrão, gerado pelo programa GELCOMPAR II. PM 100 pb - padrão de peso molecular de 100 pb; dig – digerido.....72

Índice de Figuras

- Figura 10:** Perfil eletroforético do produto da digestão de *HhaI*, dos amplicons obtidos com os iniciadores fliC(M), nas amostras de *E. coli* HNT, gerado pelo programa GELCOMPAR II. PM50 pb – padrão de peso molecular de 50 pb (Invitrogen); dig – digerido.....74
- Figura 11:** Perfil eletroforético do produto da digestão de *RsaI*, dos amplicons obtidos com os iniciadores fliC(F), nas amostras de *E. coli* HNT, gerado pelo programa GELCOMPAR II. PM100 pb – padrão de peso molecular de 100 pb; dig – digerido.....75
- Figura 12:** Resultado da análise de similaridade entre os produtos da digestão dos amplicons das amostras de *E. coli* padrão e *E. coli* HNT, obtidos com os iniciadores FliC(M) e com a enzima *HhaI*, gerado pelo programa GEL COMPAR II. PM100 pb – padrão de peso molecular de 100 pb (Fermentas); dig – digerido.....77
- Figura 13:** Resultado da análise de similaridade entre os produtos da digestão dos amplicons das amostras de *E. coli* padrão e *E. coli* HNT, obtidos com os iniciadores FliC(F) e com a enzima *RsaI*, gerado pelo programa GEL COMPAR II. PM100 pb – padrão de peso molecular de 100 pb (Fermentas); dig – digerido..... 78
- Figura 14:** Vista frontal do modelo da estrutura secundária da amostra 1C.....115
- Figura 15:** Vista superior do modelo da estrutura secundária da amostra 1C.....115

Índice de Tabelas

Tabela 1: Amostras de <i>E. coli</i> produtoras de antígeno flagelar (H) padrão.....	48
Tabela 2: Amostras de <i>E. coli</i> H não-tipável (HNT) isoladas de amostras humanas.....	49
Tabela 3: Reagentes utilizados para as reações de PCR do gene <i>fliC</i>	50
Tabela 4: Iniciadores utilizados para o estudo de variação de fase.....	51
Tabela 5: Iniciadores utilizados na reação de sequenciamento do gene <i>fliC</i>	58
Tabela 6: Tamanhos dos produtos da PCR para os antígenos H de <i>E. coli</i> padrão (H1 a H56), utilizando os iniciadores Flic(M) e FliC(F).....	67
Tabela 7: Tamanho molecular do gene <i>fliC</i> em pares de base (pb) das amostras de <i>E. coli</i> HNT de origem humana, obtidos pela PCR.....	69
Tabela 8: Resultados do polimorfismo da PCR-RFLP, utilizando os iniciadores FliC(M) e Flic(F).....	73
Tabela 9: Resultados do polimorfismo da PCR-RFLP, utilizando os iniciadores FliC(M) e Flic(F) nas amostras de <i>E. coli</i> HNT.....	76
Tabela 10: Resultado da titulação dos anti-soros de <i>E. coli</i> HNT frente ao seu antígeno homólogo.....	80
Tabela 11: Resultados da PCR-RFLP e da sorologia obtidos entre as amostras de <i>E. coli</i> HNT e <i>E. coli</i> padrão.....	81
Tabela 12: Resultados da titulação dos anti-soros de <i>E. coli</i> HNT frente aos 20 antígenos das amostras de <i>E. coli</i> HNT.....	83
Tabela 13: Resultados das técnicas de PCR-RFLP, do sequenciamento e da sorologia utilizando as amostras de <i>E. coli</i> HNT.....	114

RESUMO

Escherichia coli tem sido isolada, com certa frequência, apresentando antígenos flagelares (H) que não são reconhecidos por nenhum dos anti-soros disponibilizado pelo mais importante centro de referência de *E. coli*, The International *Escherichia* and *Klebsiella* Centre (WHO) do Statens Serum Institut, Copenhague, Dinamarca. Atualmente são reconhecidos 53 antígenos “H” e, nos últimos 29 anos, nenhuma modificação ocorreu na lista dos antígenos flagelares associados à *Escherichia coli*. Isto posto, os objetivos deste trabalho foram identificar os antígenos flagelares das cepas de *E. coli* que expressam H não tipável (HNT) e que apresentam fatores de virulência associados à diferentes enteropatias. Esta identificação foi realizada inicialmente, pela reação em cadeia da polimerase (PCR) do gene *fliC*, responsável pela proteína flagelina, das 53 amostras padrões para os antígenos H e das 20 amostras HNT (H não-tipável). Em seguida, os amplicons foram digeridos por enzimas de restrição e daquelas amostras que apresentaram perfis de restrições distintos daqueles observados para as amostras padrões de antígeno H, foram produzidos soros em coelhos. Foram realizados testes de titulação frente aos 53 antígenos padrões, frente ao antígeno homólogo e frente aos antígenos das amostras HNT. As sequências gênicas das amostras HNT, obtidas na reação de sequenciamento, foram comparadas aos diferentes genes de *fliC* armazenados no banco de dados do “National Center for Biotechnology Information” (NCBI) através do sistema BLAST, e o programa ClustalW foi utilizado para alinhamento das sequências. Os resultados demonstraram que estas amostras apresentaram similaridade com antígenos padrões, entretanto, elas não possuem a mesma sequência nucleotídica e também não reagiram fenotipicamente com o anti-soro esperado. Os dados obtidos permitem concluir que no conjunto de amostras estudado, treze amostras apresentaram antígeno flagelar diferente daqueles já descritos na literatura, quando utilizado as técnicas de PCR e/ou sorologia.

ABSTRACT

Escherichia coli has been isolated frequently, showing flagellar antigens that are not recognized by any of the antisera, provided by the most important reference center of *E. coli*, The International *Escherichia* and *Klebsiella* Centre (WHO) of the Statens Serum Institute, Copenhagen, Denmark. Are currently recognized 53 H antigens and in the last 29 years, no change occurred in the list of flagellar antigens associated with *Escherichia coli*. The objectives of this study were to identify the flagellar antigens of *E. coli* that do not express non-typeable H antigens and presenting the virulence factors associated with different diseases. This identification was performed initially by gene amplification of the *fliC*, (flagellin protein) by the polymerase chain reaction (PCR) in all 53 standards *E. coli* strains for the H antigens and 20 non-typeable H-antigens *E. coli* strains, being then, the amplicons were digested by restriction enzymes. Anti-sera were produced in rabbits, those strains that showed different restriction profiles of these patterns observed for the non-typeable H antigens *E. coli* strains. Agglutination testes were carried out against the 53 antigens standards, against the homologous antigen and H antigens of the non-typeable strains. DNA sequences were compared to different *fliC* genes stored in the database of the National Center for Biotechnology Information (NCBI) through the BLAST, and ClustalW program was used to align the sequences. The results showed that although these strains have homology with a standard H-antigen, they do not have the same nucleotide sequence and did not phenotypically reacted with the antiserum expected. The data obtained showed that thirteen strains had a different H antigen those already described in the literature when used the techniques of PCR-RFLP and/or serology.

I-INTRODUÇÃO

1- *Escherichia coli*

Escherichia coli é um típico membro da família *Enterobacteriaceae*, que tem como principal habitat a microbiota intestinal de mamíferos. Caracteriza-se por ser um bacilo gram-negativo, anaeróbico facultativo, podendo causar infecções entéricas (diarréia, disenteria, colite hemorrágica, síndrome urêmica hemolítica e doença do Edema) (Nataro & Kaper, 1998; Kaper *et al.*, 2004) ou extra-intestinais (infecções do trato urinário, septicemias, meningites, peritonites e infecções pulmonares) (Ørskov & Ørskov, 1985; Féria *et al.*, 2001).

Este bacilo apresenta uma natureza tanto patogênica quanto comensal, além de possuir um alto grau de diversidade genética, ocasionada por mutações, recombinações e transferência de DNA horizontal (Baldy-Chudzik & Stosik, 2007). O nicho da *E. coli* comensal é a mucosa intestinal de mamíferos. Neste sítio a bactéria caracteriza-se por ser um competidor, fazendo parte da mais abundante microbiota anaeróbica facultativa (Sussman, 1997). Entretanto, há clones de *E. coli* altamente adaptados, que têm adquirido atributos específicos de virulência, que lhes conferem um aumento na habilidade de se adaptar a novos nichos e então causar um grande espectro de doenças. A necessidade de diferenciar entre *E. coli* comensal e patógenos entéricos levou ao desenvolvimento de testes bioquímicos, que se tornou a base da taxonomia bacteriana moderna (Sussman, 1997).

Uma das principais contribuições para a caracterização da *E. coli* foi o estabelecimento dos métodos sorológicos por Kauffmann (1947), que permitiu um grande conhecimento sobre epidemiologia, ecologia e patogênese deste microrganismo. Esse autor propôs que amostras de *E. coli* fossem identificadas tendo-se como base seus principais

antígenos de superfície. Estes foram denominados de O e H representando, respectivamente, o lipopolissacarídeo (LPS) termoestável da membrana externa da parede celular de bactérias gram-negativas e a proteína do flagelo (flagelina). A estes, somam-se os antígenos capsulares que foram denominados de antígenos K. Os patótipos de *E. coli* tendem a ser grupos clonais que definem os sorogrupos (somente antígeno O) ou sorotipos (onde o antígeno H deve ser pesquisado, O:H) (Blanco *et al.*, 1994).

Pela presença destes antígenos, a identificação sorológica tem permitido estabelecer a correlação de amostras de *E. coli* com algumas doenças, principalmente com as enteropatias (Ørskov & Ørskov, 1992).

Atualmente, há seis categorias de *E. coli* diarreiogênicas bem descritas (Kaper *et al.*, 2004):

- 1) *E. coli* Enteropatogênica (EPEC), que produz a histopatologia da lesão A/E (*attaching and effacing*), onde a bactéria adere intimamente às células do epitélio intestinal, ocasionando mudanças no citoesqueleto, incluindo a acumulação de actina polimerizada diretamente abaixo da adesão bacteriana e destruição das microvilosidades. Está associada a surtos de diarreia infantil em países em desenvolvimento.
- 2) *E. coli* Enterohemorrágica (EHEC), caracteriza-se pela produção de Shiga-Toxinas (Stx1 e Stx2) ou verocitotoxinas (VT). Em adição, as linhagens de EHEC também contêm a ilha de patogenicidade LEE (*Locus of enterocyt effacement*) que codifica um sistema de secreção do tipo III, homólogo ao sistema produzido pela EPEC. EHEC está associada a epidemias e graves síndromes clínicas como colite hemorrágica e síndrome urêmica hemolítica.

- 3) *E. coli* Enterotoxigênica (ETEC), expressa enterotoxinas termoestáveis (ST) e termolábeis (LT), assim como diferentes fatores de colonização do hospedeiro: CFA (*colonization factor antigen*), CS (*coli surface antigen*) ou PCF (*putative colonization factor*). Está associada a diarreia aquosa em crianças e diarreia em viajantes.
- 4) *E. coli* Enteroagregativa (EAEC), demonstra um padrão de adesão do tipo agregativo às células Hep-2 e HeLa, possuindo uma fímbria de aderência agregativa (AAF - *aggregative adherence fimbriae*) e secreta a enterotoxina EAST-1 (*enteroaggregative E. coli ST*). Ocasiona surtos de diarreia em crianças e adultos tanto nos países desenvolvidos quanto nos países em desenvolvimento.
- 5) *E. coli* Enteroinvasivas (EIEC), são proximamente relacionadas com *Shigella* tendo em comum características bioquímicas, genéticas e de patogenicidade. São imóveis, negativas para lisina descarboxilase e lactose. Possuem genes associados à invasão das células epiteliais intestinais. Estão associadas a casos de colite inflamatória invasiva, diarreia aquosa e ocasionalmente disenteria.
- 6) *E. coli* Aderência difusa (DAEC), é definida pela presença do padrão de aderência difusa em células HEp-2. Aproximadamente 75% das linhagens produzem a adesina fimbrial F1845. DAEC tem sido associada a casos de diarreia, particularmente em crianças com mais de 12 meses de idade.

Escherichia coli também pode ocasionar infecções extraintestinais como as infecções no trato urinário, bacteremias, septicemias e meningites que, frequentemente, ocasionam sérias sequelas e morte (Johnson & Russo, 2002). *E. coli* Extraintestinal Patogênica (ExPEC), possuem fatores de virulência como as adesinas (Fímbria do tipo1,

fímbria P, fímbria S e adesina afimbrial), toxinas (α -hemolisina, fator necrozante citotóxico do tipo 1), sistema de aquisição de ferro (aerobactina), mecanismos de resistência ao soro e à fagocitose (LPS e cápsula) que facilitam a colonização e/ou invasão no hospedeiro, ocasionando injúrias no tecido e estimulando respostas inflamatórias.

2- Estruturas Flagelares

Algumas linhagens de *E. coli* quando observadas em microscópio óptico de luz ou em microscópio eletrônico apresentam apêndices helicoidais finos. Estes apêndices são tradicionalmente chamados de flagelos ou filamentos flagelares (Macnab, 1987).

Embora o termo flagelo também seja usado para a organela propulsora de células eucariotas, o flagelo bacteriano é completamente distinto. O filamento externo do flagelo bacteriano não funciona quimiomecanicamente, mas é passivamente direcionado por um motor em sua base e a fonte de energia é um potencial de prótons transmembrana. A atividade do flagelo é controlada por um aparato sensorial através de um sistema de controle que promove o movimento da célula bacteriana junto a um ambiente atrativo e a repele de uma via não atrativa (Lino, 1977; Macnab, 1987; Macnab 1992).

O flagelo bacteriano é facilmente reconhecido morfológicamente. Sua estrutura é um complexo de anéis associado a uma haste central chamada corpo basal, que está ancorada na membrana interna e externa da célula. Em seguida, há um segmento curto, curvado, denominado gancho que está localizado externamente à célula e providencia a junção entre o corpo basal e o filamento. O filamento protéico externo é constituído por subunidades repetidas da proteína flagelina em um arranjo helicoidal e está associado à

antigenicidade e morfologia do flagelo (Macnab, 1987; Macnab, 1992; Winstanley & Morgan, 1997; Chevance & Hughes, 2008).

2.1- Filamento flagelar:

O filamento flagelar não é uma parte do aparato de transdução de energia quimiomecânica, mas é o componente propulsor do flagelo que trabalha hidrodinamicamente no meio (Macnab, 1987).

O filamento flagelar consiste de 11 protofilamentos com a mesma polaridade e possui tipicamente um comprimento de 5 a 15 μm , tendo diâmetro constante de 20 nm por todo seu comprimento. Os parâmetros helicoidais variam levemente entre as espécies de bactérias e até entre diferentes linhagens bacterianas. Os valores típicos são de 2,0 a 2,5 μm pela extensão de uma onda helicoidal e 0,4 a 0,6 μm pelo diâmetro helicoidal. Para qualquer filamento, estes valores são constantes por todo seu comprimento, sendo que a estrutura tem uma geometria conservada e pode ser observada como um cristal unidimensional (DePamphilis & Adler, 1971; Silverman & Simon, 1977; Macnab, 1987; Samatey *et al.*, 2001).

Morfologicamente, os flagelos podem ser diferenciados através dos seus antígenos flagelares e agrupados em seis morfotipos (A-F), como observado em microscópio eletrônico de transmissão (Lawn *et al.*, 1977; Whitfield *et al.*, 1988; Seah & Kwang, 2000). Um simples filamento consiste de uma montagem quaternária de milhares de cópias de uma simples proteína, a flagelina (gene *fliC*). O peso molecular das subunidades de flagelina em *E. coli* pode variar de 27KDa a 69KDa e está correlacionado com o diâmetro dos filamentos quando comparado com os diferentes morfotipos (Whitfield *et al.*, 1988).

Filamentos flagelares isolados podem ser rapidamente despolimerizados pelo calor, pH ácido ou outros tratamentos. O processo reverso de polimerização a filamentos ocorre espontaneamente sob condições adequadas, rendendo uma estrutura que é indistinguível da original. Normalmente este processo requer a iniciação por um curto fragmento de “origem” do filamento. *In vitro* a razão do alongamento é constante, com monômeros sendo adicionados no topo distal do filamento (DePamphilis & Adler, 1971; Silverman & Simon, 1977; Macnab, 1987).

Baseado particularmente no mapeamento genético e sequenciamento da flagelina (gene *fliC*), concluí-se que as porções amino e carboxi-terminal são altamente conservadas, sendo essencial para a polimerização e secreção das moléculas de flagelina (Seah & Kwang, 2000). Concluí-se também, que a região central da flagelina pode divergir sem interferir no processo de montagem, sendo que cada divergência pode ser responsável pela variação antigênica, oferecendo os epítomos específicos dos sorotipos (DePamphilis & Adler, 1971; Silverman & Simon, 1977; Macnab, 1987).

A flagelina é responsável pela atividade antigênica do filamento flagelar, tão bem como a sua morfologia. Isto está bem demonstrado pelo isolamento de uma grande variedade de mutantes que resultam em mudanças na especificidade antigênica e forma flagelar (DePamphilis & Adler, 1971; Silverman & Simon, 1977; Macnab, 1987, Chevance & Hughes, 2008). Em estudo realizado por Kuwajima (1986), foi demonstrado que a região variável central da flagelina de *E. coli* K-12 pode ser deletada sem efeitos deletérios na função da proteína afetando porém, a antigenicidade e a morfologia do filamento, uma vez que a região central é a superfície exposta do filamento flagelar. Uma simples mutação na sequência *fliC* pode, potencialmente, resultar na geração de novos antígenos H e este mecanismo pode ter contribuído para a origem dos sorotipos H1 e H12, que diferem

somente em poucos aminoácidos (Schoenhals & Whitfield, 1993), e para os antígenos não tipáveis existentes. Também tem sido proposto que a grande diversidade de antígenos H pode ser menos dependente do acúmulo de mutações na região variável central e que os novos sorotipos derivam da transferência gênica lateral e recombinação do DNA dentro do gene *fliC* (Smith & Selander, 1990; Joys & Schodel, 1991; Schoenhals & Whitfield, 1993).

Smith e Selander (1990) demonstraram que as sequências das regiões centrais do gene *fliC* de *Salmonella* com diferentes genótipos eram constantes. Estes autores sugeriram que a região central altamente variável pode também contribuir para o processo de pressão seletiva e que esta região não está sujeita a taxas de mutação muito altas.

Wei e Joys (1985) sugeriram que a conservação da região amino-terminal da proteína flagelina de enterobactérias é necessária e reflete seu papel na montagem e regulação. A região amino-terminal da flagelina tem aproximadamente 180 resíduos de aminoácidos (aa) e 100 resíduos de aa na região carboxi-terminal. Estes são bem conservados e tem um papel importante na polimerização, secreção e polimorfismo dos filamentos flagelares (Reid *et al.*, 1999). A remoção do segmento carboxi-terminal afeta principalmente a habilidade de polimerização, enquanto que a ausência da porção amino-terminal reduz significativamente a estabilidade do filamento. A região central, que produz a parte antigênica exposta, é altamente variável dentro e entre espécies. Estudos em *Salmonella enterica* sugerem duas hipóteses: a primeira é que esta região submeteu-se a uma evolução não compelida, através da rápida fixação de mutações neutras pela deriva genética; a segunda hipótese é consideravelmente suportada pela evidência de que a transferência gênica lateral e a recombinação de DNA são as principais fontes de variação para se criar novos genes *fliC* e gerar esta diversidade antigênica (Reid *et al.*, 1999). Em outros casos, o papel da seleção natural pode ser problemático, já que, em certas partes da

molécula, a seleção pode operar como uma força conservadora contra a mudança de aa, enquanto que em outras partes o processo de seleção pode também ser uma força promotora de diversidade, gerando polimorfismo e rapidez na evolução (Reid *et al.*, 1999).

A região terminal conservada da flagelina está localizada na porção interna do filamento intacto, enquanto a superfície externa é formada pela região variável central da molécula de flagelina (Schoenhals & Whitfield, 1993).

2.2- Gancho

O filamento é conectado na célula pelo gancho, que é estruturalmente similar ao filamento, mas contém subunidades distintas. O gancho tem um comprimento bem definido, tendo 900 nm de comprimento e um diâmetro um pouco maior do que o filamento flagelar. A proteína do gancho tem um peso molecular de 42KDa e, entre o gancho e o filamento, estão duas junções ou proteínas acessórias do gancho (HAP1 e HAP3). Uma das grandes dificuldades de estudar a subunidade do gancho tem sido a disponibilidade do material. O gancho representa somente 1% do total da proteína da organela do flagelo (DePamphilis & Adler, 1971; Silverman & Simon, 1977; Macnab, 1987).

As funções do gancho podem ser atuar como uma “união universal” na base do filamento e permitir a transmissão eficiente do movimento de rotação para o resto da estrutura. Sua principal característica é a flexibilidade, que auxilia o flagelo a se mover corretamente. O topo distal do gancho pode atuar como um ponto de iniciação para o crescimento do filamento, e a extremidade proximal da estrutura termina na roda inserida através da membrana celular para a base do corpo basal (Silverman & Simon, 1977).

O gancho é muito mais estável do que o filamento. Condições de pH baixo, altas concentrações de uréia, solventes orgânicos, ou outros agentes desnaturantes que dissociam

completamente os filamentos flagelares têm pouco efeito no gancho. Esta observação tem sido a base para separar o gancho do filamento. Anticorpos específicos contra ambas as subunidades da proteína do gancho e do filamento têm sido preparados, sendo que o anticorpo anti-proteína do gancho não reage cruzadamente com as proteínas do filamento, mas demonstra uma reação cruzada com outros antígenos do gancho (Silverman & Simon, 1977).

2.3- Corpo Basal:

O gancho é conectado a uma estrutura complexa conhecida como corpo basal, que está imerso no complexo de membranas e parede celulares. Esta estrutura compreende aproximadamente 1% da massa da estrutura flagelar, entretanto, é a parte mais complexa do flagelo. O corpo basal tem uma simetria cilíndrica, consistindo de uma haste unida ao gancho e quatro tipos de anéis em bactérias gram-negativas. Os dois anéis externos (chamados L e P) estão separados, mas parecem ligados por uma parede cilíndrica. Os dois anéis internos (M e S) parecem encontrar-se diretamente ao lado dos anéis L e P (Figura 1). Ambas as considerações geométricas e bioquímicas sugerem que na célula intacta o anel L está no plano da membrana externa (lipopolissacarídeo), o anel P está no plano da camada peptidoglicana e os anéis M e S estão na membrana celular, provavelmente juntos na face externa, sendo que o anel M está diretamente associado com a membrana e o anel S está sobre a membrana. Em bactérias gram-positivas, a estrutura é menos complexa, sendo encontrados somente dois anéis. O anel interno aparece em contato com a membrana plasmática da célula e o anel externo pode estar associado com os polímeros de ácido teicóico na superfície externa (DePamphilis & Adler, 1971; Silverman & Simon, 1977; Macnab, 1987).

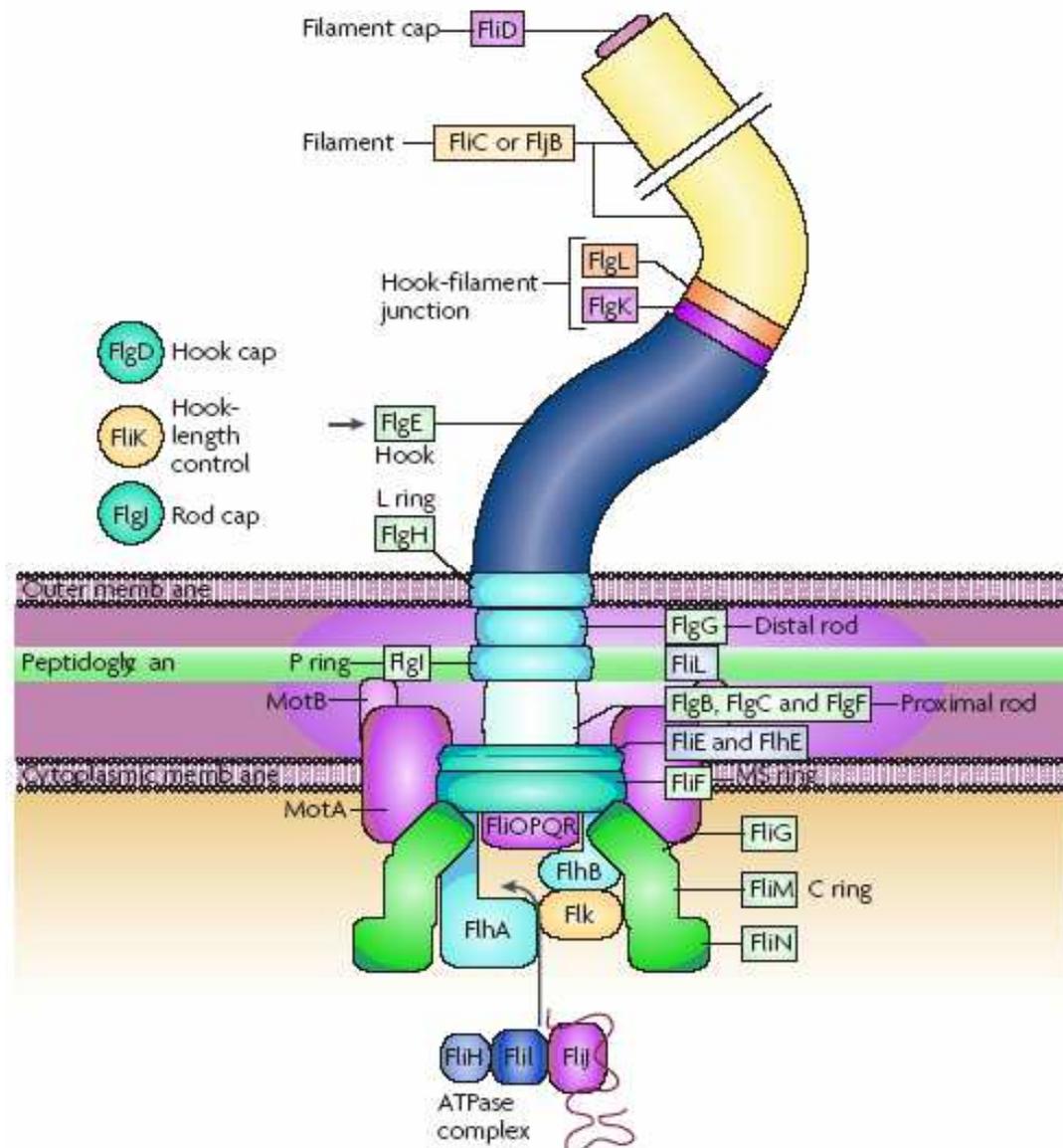


Figura 1: Componentes flagelares da *Salmonella enterica*. A figura mostra as estruturas presentes na parede celular e na membrana celular. Componentes citoplasmáticos solúveis como FliH, FliL, FliJ, complexo ATPase, auxiliam na liberação dos substratos e na ordem da secreção dos substratos. A proteína do filamento consiste em FliC ou FljB que são alternativamente transcritas. Fonte: Chevance & Hughes (2008).

Embora o corpo basal seja o principal componente do flagelo bacteriano, ele não contém as proteínas conhecidas por serem importantes para a função do motor. As

proteínas MotA e MotB são proteínas integrais da membrana celular e são necessárias para a rotação do motor, mas não para o movimento de troca de rota ou montagem do flagelo (Figura 1). Os corpos basais podem ser rotacionados pelas proteínas MotB recentemente sintetizadas. Estas observações sugerem que proteínas MotA e MotB podem ser arranjadas circunferencialmente ao redor dos anéis S e M (Macnab, 1987).

Pelo menos três proteínas não associadas ao corpo-basal (FlaBII_E/FlaAII.2_S, FlaAII_E/FlaQ_S, e MotD_E/FlaN_S) são necessárias para a montagem do flagelo e estão também envolvidas na rotação do motor e movimento de troca de rota. Quando super-produzidas elas aparecem no citoplasma e, no caso das proteínas FlaAII.2_S e FlaQ_S elas são retidas no citoplasma. Estas podem estar na periferia da membrana celular, possivelmente na face citoplasmática do anel M e podem formar um complexo que providencia a interface entre o mecanismo de transdução de energia e a informação sensorial do sistema de quimiotaxia (Macnab, 1987).

Um motor gerado pela força motora de prótons transmembrana dirige a rotação flagelar. O complexo motor-troca de rota (*switch*) é montado na face citoplasmática dos anéis M e S e compreende uma estrutura em forma de sino conhecida como o anel C. O complexo contém três proteínas (FliG, FliM e FliN) envolvidas na geração do movimento de retorcida e troca de rota de direção. Acredita-se que o movimento de rotação pertence aos anéis M e S, haste, gancho e filamento. A parte fixa do motor é feita por duas proteínas que cercam os anéis M e S. O domínio carboxi-terminal de MotB está provavelmente ancorado na parede celular e as quatro hélices hidrofóbicas de MotA interagem com a hélice de MotB atravessando a membrana para formar um canal condutor de prótons através do citoplasma da membrana. Acima de oito complexos de proteínas Mot estão

interagindo com o complexo motor-troca de rota para gerar rotação em resposta a uma corrente interna de prótons (Macnab, 1987).

3- O Sistema Gênico do Flagelo

Muitas espécies de bactérias movem-se utilizando os flagelos. Junto com os genes do aparato sensorial que controlam a operação de movimentação, o flagelo também possui um sistema complexo de genes estruturais. Os genes estruturais para flagelina foram denominadas *H1* e *H2* em *Salmonella*, *H* em *Bacillus subtilis* e *Pseudomonas aeruginosa* e *hag* (*fliC*) em *Escherichia coli*. Mutações em genes estruturais modificam qualquer um dos seguintes fenótipos: eficiência na formação do filamento, forma do filamento e especificidade do antígeno flagelar (Lino, 1977).

Os genes flagelares de *S. TYPHIMURIUM* e *E. coli* são encontrados em locações cromossômicas, onde formam *clusters* contínuos de considerável tamanho. Estes *clusters* foram determinados nas regiões flagelares I, II e III. Dentro de cada região, existem diversos *operons*, contendo de um a nove genes. Estes *operons* estão arranjados em uma hierarquia regulatória ou *regulon*. O *regulon* consiste de genes cujos produtos são componentes estruturais do flagelo e produtos que regulam a expressão destes genes estruturais ou a atividade completa do flagelo (Macnab, 1992).

Em 1988, foi proposta uma nova nomenclatura para os genes flagelares de *E. coli* e *S. TYPHIMURIUM* (Lino *et al.*, 1988). O símbolo base em todos os casos começava com *fl*. Os genes na região flagelar I foram designados pelo símbolo *flg* (flagelo), a região II foi chamada de *flh* e a região III (até então designada *hag*) foi renomeada para *fli*. Os genes da região da flagelina associados à mudança de fase da *S. TYPHIMURIUM* foi designada pelo símbolo *flj*.

No topo da hierarquia regulatória de bactérias entéricas está o operon classe I (*flhDC*). Este operon parece ser o sítio primário para o controle do regulon da motilidade. A informação é recebida de acordo com a nutrição celular, osmolaridade externa, temperatura e outras condições ambientais. Se as condições forem apropriadas para expandir quantidade significativa de energia necessária para motilidade, o operon é expresso e o complexo *flhDC* induz a transcrição dos genes da região II (Mytelka & Chamberlin, 1996).

Muitos dos genes para estruturas do corpo-basal e do gancho são encontrados na região I incluindo aqueles para a formação do filamento (*flgB*, *flgC*, *flgF*, *flgG*), dos anéis externos (*flgH*, *flgI*), do gancho (*flgE*) e da junção filamento-gancho (*flgK*, *flgL*). Também há outros genes como *flgD*, que controla o avanço da montagem do gancho para a montagem do filamento, o gene *flgA*, que permite a montagem de um dos anéis externos do corpo basal, e o gene *flgM*, que previne a expressão de genes tardios como o gene do filamento até que a montagem do complexo corpo basal-gancho esteja completa. Logo, a região I é dedicada para os componentes da estrutura flagelar externa próxima da célula (Lino, 1977; Silverman & Simon, 1977; Macnab, 1992).

Os genes de classe II incluem em sua maioria, genes estruturais para o complexo filamento – gancho – corpo basal, mais o fator σ *fliA*. O produto do gene *fliA*, o fator σ^F , é um membro da família de fatores σ^{28} e reconhece o promotor TAAAGTTX₁₁ τ GCCGATAA_C. Este fator dirige a transcrição dos genes de classe III, que codificam as proteínas associadas ao gancho, a proteína do filamento, as proteínas do motor e várias proteínas de quimiotaxia. A região II contém todos os genes associados ao processo de quimiotaxia (*cheA*, *cheB*, *cheR*, *cheW*, *cheY*, *cheZ*), dois genes receptores (*tar* e somente *tap* em *E. coli*), e dois genes para motilidade (*motA* e *motB*). Os produtos destes genes são

úteis somente quando a montagem flagelar está completa (Lino, 1977; Silverman & Simon, 1977; Macnab, 1992).

A região III contém muitos dos genes que estão envolvidos nos estágios iniciais da montagem do flagelo. Incluem-se os genes dos anéis MS (*fliF*), um componente inicial do corpo basal (*fliE*), a troca de direção (*fliG*, *fliM*, *fliN*) e parte do aparato de exportação (*fliH*, *fliI*). Outro gene pertencente a esta região é *fliK*, que é responsável pelo controle do comprimento do gancho. Pesquisas utilizando mutantes indicaram que todos os genes da região III são necessários no estágio inicial do processo de montagem (Lino, 1977; Silverman & Simon, 1977; Macnab, 1992).

Uma outra classe de gene III, *flgM*, codifica um regulador para σ^F associado a atividade negativa. Quando quaisquer dos genes estruturais de classe II são mutantes, FlgM liga-se a σ^F e bloqueia a transcrição de genes de classe III. Se todos os genes são funcionais, seus produtos são montados junto a um corpo basal completo que pode colocar FlgM fora da célula, permitindo a transcrição dos genes de classe III. A proteína FlgM provavelmente ajuda no controle da indução e repressão deste regulon durante o crescimento celular (Myltelka & Chamberlin, 1996; Chevance & Hughes, 2008).

Os genes flagelares são geralmente divididos em iniciais, intermediários e tardios (expressos pelos promotores I, II e III, respectivamente). Os genes iniciais (*flhDC*) codificam o ativador da transcrição para os promotores de classe II. Os genes intermediários codificam os anéis MS, C, P e L, o sistema de secreção do tipo 3, as proteínas do corpo-basal e do gancho, que juntos formam o complexo corpo basal-gancho na membrana interna. Em seguida, as proteínas distais da estrutura flagelar em crescimento são secretadas pelo sistema de secreção do tipo 3 (Chevance & Hughes, 2008).

A montagem completa da estrutura corpo basal-gancho desencadeia a mudança na especificidade dos substratos do sistema de secreção do tipo 3, das proteínas do gancho para os substratos de secreção tardia (ex: flagelina, proteína do capuz), que são requeridos nos estágios finais da montagem flagelar. Um segundo evento é a mudança da expressão dos genes intermediários para a expressão dos genes tardios, que são mediados pelas chaperonas do sistema de secreção do tipo 3 FliT, FliA e FlgN e está intimamente associado para a secreção de substratos. A exportação de proteínas usa o sistema de secreção do tipo 3, que é caracterizado por um peptídeo sinal amino-terminal de secreção (Chevance & Hughes, 2008).

FliT é a chaperona para a proteína do capuz FliD. Após a formação do complexo corpo basal-gancho, FliD é secretada no topo do gancho onde é facilitada a polimerização do filamento flagelar. A ausência de FliD no citoplasma permite que FliT se ligue a FlhC e, desta forma, iniba a transcrição dos genes intermediários cujos produtos não são requeridos para o processo de montagem (Brutinel & Yahr, 2008; Chevance & Hughes, 2008).

O gene intermediário *fliA* codifica o fator σ^F do flagelo, que é requerido para a ativação da transcrição dos genes tardios. Além disso, o fator σ^F também atua como uma chaperona, requerida para a secreção de FlgM. Antes de completar o complexo corpo basal-gancho, FlgM é retido no citoplasma onde é diretamente ligado a σ^F e inibe a expressão dos genes tardios. Após a formação completa do complexo corpo basal-gancho, a secreção de FlgM resulta na transcrição (dependente do fator σ^F) dos genes tardios, permitindo a montagem final do flagelo. Além disso, a redução citoplasmática de níveis de FlgM limita a duração da expressão dos genes tardios pela diminuição da meia vida do fator σ^F no citoplasma onde este é degradado por proteases (Brutinel & Yahr, 2008; Chevance & Hughes, 2008).

FlgK e FlgL estão entre as primeiras proteínas a serem secretadas, permitindo a mudança para a secreção dos substratos tardios. A chaperona para FlgK e FlgL é FlgN, e após a secreção de FlgK e FlgL, a FlgN aumenta a tradução de FlgM que é expressa pelo promotor de classe III, porém este mecanismo ainda não é bem compreendido (Brutinel & Yahr, 2008; Chevance & Hughes, 2008). Durante a montagem, FlgM está inviável para interagir com o fator σ^F até o término completo do filamento flagelar e, então, FlgM acaba acumulando-se no citoplasma. O aumento da tradução de FlgM é dependente de FlgN e isto permite que as células regulem os genes dependentes do fator σ^F permitindo a formação completa do filamento flagelar. Em adição, FlgM pode também regular expressão de *fliC* em níveis pós-transcricional (Chevance & Hughes, 2008).

4- Variação de Fase Flagelar em *E. coli*:

Por um longo tempo, acreditou-se que bactérias pertencentes a linhagens de *E. coli* possuíam apenas um simples gene *fliC* ou *hag* em seu cromossomo, que codificava o antígeno flagelar específico, expressando apenas um único tipo de antígeno flagelar (monofásico). Entretanto, têm sido encontradas na natureza, linhagens de *E. coli* que espontaneamente mudam de antígeno H, lembrando a presença de um sistema gênico semelhante ao sistema de variação de fase flagelar da *Salmonella sp* (Ratiner, 1982). A variação de fase do antígeno flagelar de *Salmonella sp*, é alternativamente expresso por dois diferentes genes para flagelina, *fljB* e *fliC* (Feng *et al.*, 2008). O fenômeno de variação de fase está relacionado à sobrevivência bacteriana na presença do sistema imune do hospedeiro, e desta forma está associado à virulência bacteriana.

Estudos realizados por Ratiner (1983) evidenciam que as linhagens de *E. coli* Bi7327-41 (antígeno H3) e P12b (antígeno H17) eram capazes de mudar seus antígenos

flagelares para H16 e H4, respectivamente. A mudança do antígeno H na linhagem Bi7327-41 seria controlada por um sistema gênico parecido com o sistema de *Salmonella*, que consistia de pelo menos dois grupos de genes conectados funcionalmente, porém ocupando *loci* diferentes (Ratiner, 1985). O gene que codifica para o tipo de antígeno H16 é um alelo do gene *hag* (localizado na posição 42 do mapa cromossômico da *E. coli* K-12) e foi denominado *hagA*, enquanto o outro grupo de genes que estava associado à função de mudança de fase e à manifestação do antígeno H3 foi denominado *hagB*. Um gene (*rhl* em *Salmonella*) localizado no operon *hagB*, e ativo quando este operon está sendo expresso, poderia determinar um repressor específico para muitos alelos *hagA* (Ratiner, 1983).

Estudo realizado por Ratiner (1987), demonstrou que os genes codificando os antígenos flagelares H36, H47 e H53 em *E. coli* são alelos do locus *hagB*. Em contraposição, estudos revelaram que os alelos *hagA* (H48) e *hagA* (H6) são insensíveis ao repressor-*hag* (Ratiner, 1983).

A habilidade de *hagA* (H42), *hagA* (H16) e *hagA* (H21) serem especificamente reprimidos quando *hagB* está ativo, sugere que a evolução esteja conectada com o caráter difásico e que, pelo menos uma parte das linhagens monofásicas com estes antígenos, representem variantes em que o alelo *hagB* tenha sido deletado, alterado ou estabilizado em um estado “off” (Ratiner, 1983). Além disso, este mecanismo difásico indica diversidade no arranjo dos genes responsáveis pela especificidade do antígeno flagelar (Ratiner, 1985).

Estudos associados com a variação do antígeno flagelar dependente da temperatura (Ratiner, 1999), demonstraram que em uma população de *E. coli*, podem existir linhagens que alteram sua especificidade ao antígeno flagelar de acordo com a temperatura de crescimento. Nas amostras investigadas, dois genes especificando a proteína flagelina foram encontrados: um alelo de *fliC*₄₀ sensível à repressão de fase, porém expresso em

amostras a 37°C mas não a 30°C e um alelo de *flkA*₅₃, expresso em ambas as temperaturas. Desta forma, dois antígenos flagelares (H40 e H53) foram simultaneamente expressos a 37°C. Como *fliC*₄₀ foi sensível à repressão de fase e sua expressão dependente de temperatura isto pode refletir na sensibilidade da temperatura do gene repressor dentro da região de *flkA*₅₃.

A manifestação simultânea de dois antígenos flagelares em *E. coli* tem sido reportada somente em células contendo um alelo *fliC* insensível e o locus *flkA* “on” ou *fljB* “on”. Entretanto, o gene *fliC*₄₀ descrito no artigo de Ratiner (1999), foi sensível para repressão e sua expressão foi dependente da temperatura, o que poderia então refletir na sensibilidade do gene repressor dentro da região *flkA*₅₃. Permanece, portanto a possibilidade permanece de que esta região possa conter ainda outro gene, cujo produto pode atuar a 37°C, mas não a 30°C e em baixa temperatura, para inativar (modificar) o repressor ou abolir a sensibilidade de *fliC* (Ratiner, 1999).

A dependência da temperatura (associada à morfogênese de estruturas de superfícies incluindo fimbria, atividade bioquímica, patogenicidade etc) tem sido reconhecida nas enterobactérias. A influência do ambiente, em particular na temperatura de crescimento e a presença de certos compostos químicos, foram bem estudados na flagelação e motilidade (Ratiner, 1999).

A variação de temperatura demonstra que em algumas amostras de *E. coli* a aglutinação com dois diferentes anti-soros pode não ser considerado errôneo. Além disso, a existência da dependência da temperatura na variação de fase do antígeno flagelar pode explicar algumas discrepâncias entre os resultados de tipagem de muitas *E. coli* em diferentes laboratórios (Ratiner, 1999).

Em isolados selvagens de *E. coli* foram encontrados dois genes para flagelina que eram alternativamente expressos: *hagA* e *hagB* (Ratiner, 1998). O primeiro é um alelo do gene *fliC* de *E. coli* K-12 enquanto o segundo está localizado a uma certa distância deste. Também foi demonstrado que a região de *hagB* é uma região distinta, integrada no cromossomo de *E. coli* K-12, mas fora da região conhecida por possuir as três regiões dos genes flagelares. Baseando-se na nova nomenclatura o gene *hagB* foi renomeado para *flkA* (Lino *et al.*, 1988). A alteração da atividade destes genes foi encontrada por ser usualmente unilateral (não-reversível): $fliC^{off} + flkA^{on} - fliC^{on} + flkA^{off}$. O estado “on” ou “off” está correlacionado com o estado de *flkA* especificamente, porque a região *flkA* produz um repressor similar aquele produto do gene *fljA* de *Salmonella*, atuando em muitos alelos *fliC* (sensíveis). Desta forma, a existência de um gene similar a *fljA* que é coexpresso com *flkA* foi postulado (Ratiner, 1982; Ratiner, 1983; Ratiner, 1998).

O mecanismo de mudança de fase é controlado pela orientação de um elemento invertido localizado a montante ("upstream") ao operon *fljBA*. Na orientação do segmento de DNA (segmento H +) o operon é expresso e o produto do gene *flja* (FljA) reprime a expressão do gene *fliC*, resultando na expressão da flagelina FljB, enquanto que na orientação H (-), o operon se torna inativo, resultando na expressão da flagelina FliC (Tominaga, 2004).

Estudo realizado por Ratiner (1998), revelou a presença de novos *loci* em *E. coli* *fliA* (alelos de *fliA₄₄* e *fliA₅₅*, determinando antígenos H44 e H55) e *flmA* (alelo *flmA₅₄* para o antígeno H54). Estas duas novas regiões flagelares diferem na capacidade de suprimir os alelos de *fliC* sensíveis à repressão específica de mudança de fase (Ratiner, 1998). O alelo *flmA₅₄* possui esta atividade, sendo então similar à região da flagelina de fase 2 de *Salmonella* e à região *flk* de *E. coli*. Em contraste, a região *fli* não exhibe esta característica

e, particularmente, pode ser esperada a existência de linhagens de *E. coli* que expressam dois genes para flagelina. Ratiner (1998), encontrou linhagens de *E. coli* do sorogrupo O18 exibindo dois antígenos flagelares simultaneamente (H14 e H55).

Wang *et al* (2003), demonstraram através da clonagem e expressão de *fliC* que 43 dos 53 antígenos H conhecidos são codificados pelo gene *fliC* e que os outros 10 estão localizados em outro *locus*. Neste trabalho, somente as linhagens H36 e H53 não foram amplificadas por nenhum par de iniciadores utilizados. Ambos os antígenos são conhecidos por serem codificados por um gene diferente de *fliC* (Ratiner, 1998) e presume-se que uma parte do gene *fliC* poderia estar parcialmente deletada. Logo, dos 53 tipos diferentes de antígeno H, 43 expressaram *fliC*, enquanto seis linhagens que tinham *fliC* expressaram um antígeno H alternativo quando clonado, mas não quando situado no cromossomo. As outras quatro linhagens são conhecidas por não expressarem *fliC*, mas por possuírem o antígeno H em um *locus* diferente de *fliC*.

Os genes *fliC* das linhagens de *E. coli* que expressam H3, H40, H47 e H55, expressaram os antígenos H16, 8, 21 e 38 respectivamente, enquanto as linhagens H17 e H44 expressaram ambas o antígeno H4. As linhagens H3, H36, H47 e H53 têm sido identificadas no *locus flkA*. As linhagens H44 e H55 no *locus flIA*, e a linhagem H54 no *locus flmA*. Entretanto, muitas destas linhagens expressando muitos destes antígenos H são conhecidas por também possuírem *fliC* (Wang *et al.*, 2003).

Em um estudo recente, Feng *et al.*, (2008) demonstraram que 44 linhagens de *E. coli* realmente codificavam a flagelina no *locus fliC*, enquanto que em algumas linhagens ela é codificada pelo *locus flk* (H3, H35, H36, H47 e H53), *flI* (H44 e H55) e *flm* (H54). A região *flkA* e o gene repressor *flkB* foram localizados entre os genes cromossomais *rnpB* e *yhaC* nas amostras de *E. coli* que expressaram H3 e H53, e o produto de *flkB* reprimiu a

expressão de *fliC* em *E. coli*. Desta forma, o gene *flkB* tem uma função similar ao *fljA* de *Salmonella enterica*.

Feng *et al.* (2008), demonstraram o mecanismo envolvido na variação de fase de amostras de *E. coli* positivas para *flk*. As regiões entre *yhaC* e *rnpB* nas amostras H3, H35, H36, H47 e H53, foram sequenciadas e o locus *flk* foi encontrado contendo um possível gene para integrase, sendo flanqueado por sequências repetidas diretas em ambas extremidades do gene *flkA* e do repressor *flkB*, indicando que pertencem a uma típica ilha de patogenicidade, e que foi denominada ilha de patogenicidade (IP) *flk*. Quando a IP *flk* está presente no cromossomo, o operon *flkAB* é expresso permitindo a produção da flagelina FlkA e do repressor *flkB*, sendo então reprimida a expressão de *fliC*. Quando a IP *flk* é excisada do cromossomo, *flkAB* é irreversivelmente deletada e, desta forma, a repressão de *fliC* é liberada e a flagelina (FliC) é produzida. Das cinco linhagens de *E. coli* que possuem o gene *flk*, o gene *fliC* das linhagens H35, H36 e H53 são interrompidos ou defeituosos. Desta forma, variantes negativas para *flk* poderiam ser imóveis. Estudos adicionais demonstraram que o que se pensava ser um gene para integrase, foi caracterizado como *orf204* e seria um possível gene para invertase (Feng *et al.*, 2008). A recombinação sítio-específica mediada pela invertase requer a presença de sequências repetidas invertidas que estão ausentes na região *flk*. Desta forma, o gene *orf204* é não-funcional ou não está associado à região *flk*. Logo, a origem da IP *flk* ainda permanece desconhecida. Especula-se que esta ilha poderia ter se desenvolvido de um ancestral simples e que tenha sido integrada a outros hospedeiros por recombinação no sítio *attB*, uma vez que a presença deste sítio de 23pb em muitas espécies bacterianas pertencentes a família *Enterobacteriaceae* permite a possibilidade da presença de IP em outras espécies bacterianas, entretanto, o potencial de integração da IP *flk* ainda permanece por ser estudado (Feng *et al.*, 2008).

5- Sorologia:

Quando Kauffmann iniciou a sorotipagem da *E. coli*, doenças diarreio gênicas não eram consideradas (Kauffmann, 1947). O foco do estudo foi em doenças extraintestinais, como infecções do trato urinário, septicemia e apendicite. Os resultados sorológicos demonstraram que certos grupos O de *E. coli* estavam frequentemente associados à patogenicidade para o homem e animais. Observou-se ainda que a toxicidade estava associada a certos tipos de antígenos O:K:H (Kauffmann, 1947; Ørskov & Ørskov, 1992).

Atualmente são reconhecidos 53 antígenos “H” numerados de 1 a 56 em *E. coli*, sendo que os números 13, 22 e 50 não são utilizados. Em 1975, Ørskov *et al.*, descreveram o último antígeno incluindo, o antígeno H56, reconhecido pela comunidade científica internacional.

Em infecções intestinais ocasionadas por *E. coli*, muitos sorotipos podem estar envolvidos, mas somente um número limitado de sorotipos está predominantemente entre os sorotipos patogênicos. Além disso, os sorotipos das amostras diarreio gênicas não são geralmente encontrados na microbiota normal do intestino do hospedeiro, como ocorre na *E. coli* extraintestinal (Ørskov & Ørskov, 1992).

Os sorotipos das *E. coli* associados às doenças extraintestinais são encontrados tanto em animais domésticos quanto no homem, enquanto que os sorotipos associados a *E. coli* diarreio gênicas demonstraram mais especificidade e são limitados à doença. Na maioria dos casos, a sorotipagem completa de O:K:H, é um bom indicador das relações clonais entre as amostras. Porém, é importante ressaltar que dois clones não associados mas possuindo o mesmo sorotipo podem ser encontrados (Achtman *et al.*, 1986).

Nos estudos iniciais de sorotipagem foi demonstrado que somente a tipagem do sorogrupo O era insuficiente para traçar a epidemiologia da *E. coli* enteropatogênica

(EPEC). Logo foi recomendado pelos centros de referência de *E. coli*, o uso da tipagem dos antígenos O e H para verificar quais eram os sorotipos associados com os surtos de diarreia infantil (Ewing *et al.*, 1986, Ørskov & Ørskov, 1992). Os sorotipos clássicos de EPEC são: O55:H6, O86:H34, O111:H2, O114:H2, O125ac:H6, O126:H2 , O127:H6, O128:H2, O142:H6 (Ørskov & Ørskov, 1992).

As amostras de *E. coli* enterotoxigênicas (ETEC) possuem um grande número de sorotipos O:K:H, porém um grupo prevalente e bem definido de sorotipos é encontrado em determinadas regiões geográficas, como o sorotipo O153:H45 que é comum no Chile, Argentina e Espanha (Ørskov & Ørskov, 1992). Já as amostras de *E. coli* enterohemorrágica (EHEC), apesar de produzirem verocitotoxina e esta ser encontrada em diferentes sorotipos, poucos clones são bem caracterizados como os sorotipos O157:H7 e O26:H11 (Ørskov & Ørskov, 1992).

As amostras de *E. coli* enteroinvasoras (EIEC) possuem um número restrito de sorotipos O:H, uma vez que a maioria das amostras são imóveis e não possuem antígenos K. Os sorotipos mais frequentes são: O28:H-, O124:H30, O136:H-, O144:H- e O173:H- que raramente são encontrados na microbiota do intestino ou associados com outras doenças por *E. coli* (Ørskov & Ørskov, 1992).

Reações sorológicas do antígeno H de *E. coli* são usadas para distinguir entre os diferentes sorotipos das linhagens. O teste de aglutinação é rotineiramente utilizado para a sorotipagem do antígeno H da *E. coli*. Entretanto, reações sorológicas cruzadas entre os antígenos H são frequentemente observadas e geralmente ocorrem entre sorotipos dentro do mesmo morfotipo (Whitfield *et al.*, 1988; Seah & Kwang, 2000). Os morfotipos flagelares são caracterizados pelas variações presentes nas estruturas de superfície dos diferentes flagelos de *E. coli* (Lawn, 1977) e são classificados de A-F pela análise da arquitetura da

superfície dos filamentos flagelares intactos, utilizando um microscópio eletrônico de transmissão (Schoenhals & Whitfield, 1993).

Uma vez produzido o anti-soro contra o antígeno H, este deve ser utilizado na diluição de 1:1000, porém, os resultados dos testes de titulação demonstraram que a técnica de soroabsorção é necessária para retirar as reações cruzadas e deixar o soro específico (Ewing, 1986). Na literatura encontramos que as reações cruzadas geralmente envolvem o anti-soro H1 com o antígeno H12 e vice-versa, o anti-soro H8 com o antígeno H40, o anti-soro H40 reage com os antígenos H8 e H11, o anti-soro H11 reage com os antígenos H21 e H40, já o anti-soro H21 reage com o antígeno H11, o anti-soro H37 reage com o antígeno H41 e o anti-soro H41 reage com os antígenos H37 e H39, e o anti-soro H49 reage com o antígeno H39 (Ewing, 1986).

6-PCR-RFLP:

Devido às regiões terminais 5' e 3' do gene *fliC* serem altamente conservadas, elas tornaram-se bons alvos para o método da Reação em Cadeia da Polimerase (PCR). Estas regiões conservadas permitem a amplificação de uma variedade de alelos com um único iniciador (Fields *et al.*, 1997; Machado *et al.*, 2000). A aplicação da técnica de PCR-RFLP permitiu a identificação dos alelos dos genes *fliC* característicos de *E. coli* e a avaliação do grau de sua diversidade (Baldy-Chudzik & Stosik, 2007). A alta diversidade dos perfis de restrição foi atribuída à variabilidade da região interna do gene (Moreno *et al.*, 2006). O método também tem sido adequado para a diferenciação do gene *fliC* em *E. coli* patogênica (Fields *et al.*, 1997, Moreno *et al.*, 2006) e também pode ser particularmente útil para a genética de população de numerosas linhagens de *E. coli* comensais, uma vez que o custo

com o sequenciamento de numerosos produtos de PCR e o acesso aos soros contra o antígeno H pode ser alto.

Não há dúvida que o ensaio de PCR-RFLP tem vantagens, como por exemplo, a capacidade de identificar todos os antígenos flagelares ou todos os antígenos H conhecidos em amostras imóveis ou em amostras pouco-flageladas (Prager *et al.*, 2003).

Em 1997, Fields *et al.*, desenvolveram um método de PCR-RFLP para identificar e caracterizar molecularmente o gene *fliC* de *E. coli* O157:H7 e O157:H- (imóvel). Neste trabalho foi observado que a PCR-RFLP é uma ferramenta útil para a caracterização de isolados imóveis e que os resultados obtidos podem permitir a detecção e inclusão de novos sorotipos de *E. coli* clinicamente importantes.

Machado *et al.*, (2000) propuseram a identificação dos diferentes tipos de antígenos flagelares de *E. coli* pela PCR-RFLP, com o propósito de construir um banco de dados contendo todos os padrões de restrição para subsequente dedução dos tipos de antígenos H.

Botelho *et al.*, (2003) propuseram a tipagem do antígeno H de EPEC, STEC, EAEC e ETEC que ocorrem no sorogrupo O da EPEC clássica, através da análise da PCR-RFLP. Os resultados obtidos demonstraram que a enzima de restrição *RsaI* era capaz de identificar que as amostras de *E. coli* eram pertencentes aos sorotipos H2, H6, H34, H40 e que estas amostras são a principal causa da diarreia infantil no Brasil.

Prager *et al* (2003), propuseram um estudo sobre a caracterização dos antígenos H de EPEC, comparando a metodologia de Fields *et al* (1997) e Machado *et al* (2000). Propuseram também, verificar a correlação dos resultados da PCR-RFLP com os resultados sorológicos. Foi demonstrado que amostras não-tipáveis sorologicamente poderiam ser classificadas de acordo com o polimorfismo de *fliC* e que a PCR-RFLP poderia ser usada para detectar novos antígenos H.

Amhaz *et al.*, (2004) realizaram um estudo onde a PCR-RFLP foi usada para a caracterização molecular do gene *fliC* em amostras de EIEC isoladas de pacientes com diarreia em diferentes países. As sequências nucleotídicas dos diferentes padrões de PCR-RFLP obtidos foram usadas para a construção de um dendrograma, e para a análise da diversidade genética entre os sorotipos de EIEC. Estes autores observaram que a maioria dos sorotipos estudados apresentaram baixa diversidade genética principalmente na região central.

Moreno *et al.*, (2006) formularam uma hipótese: poderia os métodos de sorotipagem clássicos serem trocados pela técnica de PCR-RFLP para a caracterização do antígeno H de amostras de ETEC? Neste trabalho, os autores concluíram que a técnica de PCR-RFLP foi capaz de direcionar a determinação do antígeno O após a caracterização do antígeno H e foi uma eficiente ferramenta na caracterização do gene *fliC* de ETEC, propondo então que a técnica de sorotipagem poderia ser trocada pela PCR-RFLP.

Baldy-Chudzik e Stosik, (2007) estudaram o gene *fliC* de *E. coli* isoladas de dez espécies de mamíferos de um zoológico. A pesquisa teve como objetivo identificar os alelos do gene *fliC* de amostras isoladas de várias espécies diferentes e provar a correlação entre a diversidade do gene *fliC* e a fonte de origem da amostra. Os resultados obtidos permitiram a identificação dos alelos *fliC* característicos das *E. coli* de cada espécie e avaliaram o grau da diversidade entre as espécies estudadas, demonstrando também que a pressão seletiva do ambiente do trato gastrointestinal contribuiu para a preservação de formas selecionadas de alelos dos genes *fliC* em *E. coli*.

Até o momento, apenas Prager *et al* (2003) utilizaram amostras de *E. coli* H não-típavel (HNT) e amostras fenotipicamente imóveis provenientes de vários casos clínicos. Nestas amostras, o fragmento de PCR obtido foi submetido aos testes de RFLP para a

caracterização do antígeno H. Entretanto, em algumas amostras de *E. coli*, o perfil de restrição encontrado não pertencia a nenhum dos 53 perfis de restrição obtidos das amostras padrão, sendo que estas amostras poderiam representar um novo tipo de antígeno H, o que poderia ser confirmado através da obtenção dos anti-soros e testes de aglutinação.

Neste trabalho, utilizamos a metodologia proposta por Fields *et al* (1997) e por Machado *et al* (2000) para a identificação do antígeno flagelar de vinte amostras de *E. coli* provenientes de casos clínicos como diarreia, síndrome urêmica hemolítica, colite hemorrágica e infecções extraintestinais, que frente aos 53 anti-soros disponíveis comercialmente, não tiveram os antígenos H tipados. Além do método da PCR-RFLP utilizamos as técnicas de sorologia e sequenciamento para a identificação de possíveis novos antígenos flagelares.

II-OBJETIVOS

OBJETIVOS

Este trabalho teve como objetivo a identificação de possíveis novos antígenos flagelares, em amostras de *E. coli* de origem humana. Para tal buscamos:

- Estabelecer um banco de dados dos padrões de perfis de restrição (P), do gene *fliC* das 53 amostras padrões para antígeno H, através da utilização da PCR-RFLP.
- Identificar o padrão de restrição (P) das amostras de *E. coli* HNT, através da utilização do banco de dados;
- Verificar a presença dos genes associados à variação de fase: *fljA*, *flkA* e *flmA*;
- Produzir anti-soro em coelhos para cada amostra de *E. coli* HNT;
- Verificar a especificidade dos anti-soros frente às amostras de *E. coli* padrão dos antígenos H e as amostras de *E. coli* HNT utilizando técnicas imunológicas;
- Sequenciar e comparar os nucleotídeos do amplicon do gene *fliC* obtido das amostras de *E. coli* HNT com àqueles já descritos.

III-MATERIAL E MÉTODOS

3.1 - Amostras de *Escherichia coli*:

3.1.1 - Amostras de *E. coli* padrões para o antígeno flagelar “H”:

Foram utilizadas 53 amostras de *E. coli* padrão para os antígenos H (Tabela 1). Estas amostras pertencem ao Laboratório de Antígenos Bacterianos II, do Departamento de Microbiologia e Imunologia, do Instituto de Biologia, da Universidade Estadual de Campinas, SP, Brasil.

Tabela 1: Amostras de *E. coli* produtoras de antígeno flagelar (H) padrão.

O6:H1	O9:H12	O86:H25	O86:H36	O156:H47
O3:H2	O18:H14	O38:H26	O42:H37	O16:H48
O53:H3	O23:H15	O58:H27	O69:H38	O6:H49
O50:H4	O46:H16	O148:H28	O110:H39	O8:H51
O4:H5	O15:H17	O138:H29	O41:H40	O11:H52
O120:H6	O17:H18	O86:H30	O137:H41	O148:H53
O1:H7	O32:H19	O73:H31	O70:H42	O161:H54
O105:H8	O126:H20	O114:H32	O140:H43	O4:H55
O30:H9	O146:H21	O60:H33	O3:H44	O139:H56
O108:H10	O158:H23	O160:H34	O125:H45	
O26:H11	O51:H24	O134:H35	O115:H46	

3.1.2 – Amostras de *E. coli* com antígeno H não - tipável (HNT):

Foram utilizadas nove amostras de *E. coli*, isoladas de fezes diarréicas de humanos (Tabela 2), gentilmente cedidas pelo Prof. Dr. Jorge Blanco, Diretor do Laboratório de Referência de *Escherichia coli*, da Faculdade de Veterinária da Universidade de Santiago de Compostela, Espanha, e onze amostras de *E. coli* provenientes de vários casos clínicos

de origem humana, gentilmente cedida pelo Prof. Dr. Helmut Tschäpe, do Instituto Robert Koch, Alemanha. A identificação dos antígenos flagelares foi realizada utilizando o conjunto completo de soros anti-H (53 soros) adquirido do Statens Serum Institut, Copenhague, Dinamarca. Todas as amostras foram re-ensaiadas para a confirmação dos resultados de HNT.

Tabela 2: Amostras de *E. coli* H não-tipável (HNT) isoladas de amostras humanas.

Nº	Código Original	Sorogrupo	Nº	Código Original	Sorogrupo
1C	VTH 15	O81	2A	01-03443	O55
2C	VTH 110	O84	3A	00-04915	O76
3C	VTH 118	O26	5A	00-04447	O91
4C	28011a	O84	7A	00-08242	O136
5C	33141a	ONT	8A	00-03034	O25
6C	46103B	ONT	9A	00-07153	O74
7C	40478B	ONT	10A	00-00848	O126
8C	48629c(1)	O86	11A	00-05951	Rugosa
9C	48629c(2)	ONT	13A	00-08712	O15
			14A	99-01406	O68
			15A	00-09775	O76

3.2 - Extração de DNA Bacteriano:

As amostras bacterianas foram cultivadas em 2,0 mL de caldo Luria-Bertani (LB) a 37°C por 24 horas. A extração do DNA das amostras foi realizada de acordo com o Kit “Wizard® Genomic DNA Purification System” (Promega/EUA). O DNA bacteriano foi armazenado a 4°C em água ultra pura estéril. Para quantificar o DNA, foi utilizado o

Espectrofotômetro Genequant II (Pharmacia Biotech/EUA) em comprimento de onda de 260-280nm.

Para verificar a integridade do DNA genômico nas amostras estudadas, este material foi aplicado em um gel de agarose 0,7%. A visualização da banda foi feita mediante incubação em solução de brometo de etídio (1,5 µg/mL) por 15 minutos e visualizada em transiluminador de luz UV (Amersham Pharmacia Biotech/EUA).

3.3 - Reação da Polimerase em Cadeia (PCR):

Para amplificação do gene *fliC* através da PCR utilizamos os protocolos descritos por Fields *et al* (1997) e Machado *et al.*, (2000). Os iniciadores foram designados FliC(F) e FliC(M) respectivamente. As sequências dos oligonucleotídeos são FliC(F)1: ATGGCACAAGTCATTAATACCCAAC e FliC(F)2: CTAACCCTGCAGCAGAGACA, FliC(M)1: CAAGTCATTAATAC(A/C)AACAGCC e FliC(M)2: GACAT(A/G)TT(A/G)G A(G/A/C)ACTTC(G/C)GT. Para a PCR utilizou-se os reagentes e concentrações indicadas na tabela 3.

Tabela 3: Reagentes utilizados para as reações de PCR do gene *fliC*:

Reagentes	Volume (µL)	Concentração final
10X "PCR Buffer" (Fermentas)	3	1X
25mM "dNTP mixture" (Fermentas)	0,24	0,2 mM cada
25 mM MgCl ₂ (Fermentas)	2,4	2,0mM
Iniciador 1	1,0	10pmol
Iniciador 2	1,0	10pmol
DNA extraído	1,0	50ng
" <i>Taq</i> -DNA Polimerase" (1U) (Fermentas)	0,3	1,5 U
Água ultra pura estéril (qsp)	30	

As soluções foram pré-aquecidas a 94°C/10' e submetidas a 30 ciclos de: 94°C/1', temperatura de anelamento variável de 55°C-60°C/1' para os iniciadores FliC(F) e FliC(M), temperatura de extensão de 72°C/2' e, em seguida, a 72°C/7'. O termociclador GeneAmp PCR System 9700 (Perkin-Elmer Applied Biosystems/EUA) foi utilizado para a execução das reações de PCR.

3.3.1: Iniciadores utilizados para o estudo da variação de fase:

Os iniciadores para os genes *fljA* (repressor), *flkA* e *flmA* (flagelina) foram desenhados por Cláudia de Moura (Aluna do curso de pós-graduação em Genética e Biologia molecular, IB/DMI/UNICAMP), e construídos de acordo com as sequências depositadas no GenBank por Tominaga (2004). As sequências dos oligonucleotídeos e os tamanhos dos produtos esperados estão na tabela 4.

Tabela 4: Iniciadores utilizados para o estudo de variação de fase:

Gene	Iniciadores	Sequências 5' - 3'	Produto (pb)
<i>fljA</i>	FljAF FljAR	TGCATATGGTTGCCATTAGG AGGAAGCTCGTTCGTATCCT	423
<i>flkA3</i>	FlkA3F FlkA3R	TGGCTTAGCAGTTGTCATGC AGCAACCTCCAACAACAAC	1072
<i>flkA53</i>	FlkA53F FlkA53R	GTCGCGTAGTCGGAGTCTTC AAGCAAGTCACGTGGAAGTC	933
<i>flmA54</i>	FlmA54F FlmA54R	TCGTAACGCCAATGATGGTA AGAAGACGAGCTTCGGCATA	1030

3.4 - Eletroforese em gel de agarose:

Para a leitura do resultado das amplificações, o produto da reação de PCR foi misturado ao tampão de amostra (0,25% de azul de bromofenol; 0,25% de xilenocianol; 25% de ficoll). Foram aplicados 10 µL da mistura no poço preparado em gel de agarose 1,5% (Promega/EUA), em Tampão TAE (Tris 2M, Ácido acético 0,04M, EDTA 0,01M pH 8,0). Foi utilizado um padrão de peso molecular de 1 Kb (Fermentas/EUA). A visualização das bandas foi feita mediante incubação do gel em solução de brometo de etídio (1,5µg/mL) por 15 minutos e visualizada em transiluminador de luz UV (Amersham Pharmacia Biotech/EUA). Para registro, os géis foram fotografados através do sistema ImageMaster VDS (Amersham Pharmacia Biotech/ EUA) e os arquivos salvos na extensão Tiff para posterior análise.

3.5 - Análise dos resultados da eletroforese em gel de agarose:

As imagens obtidas dos géis de agarose tiveram o peso molecular da banda amplificada determinados através do software ImageMaster TotalLab versão 2.01 (Amersham Pharmacia Biotec/EUA) e posteriormente pelo programa GEL COMPAR II versão 4.5 (Applied Maths, Bélgica).

3.6 - Digestão enzimática:

Para os produtos obtidos com o iniciador FliC(M), utilizou-se a enzima de restrição *HhaI* (*Haemophilus haemolyticus*/ Fermentas) e para os produtos da PCR, obtidos com o

iniciador FliC(F), foi utilizada a enzima de restrição *RsaI* (*Rhodopseudomonas sphaeroides*/ Invitrogen).

3.6.1: Preparo do DNA:

Os produtos obtidos através da amplificação do gene *fliC* foram concentrados com etanol, baseando-se no protocolo de Sambrook *et al.*, (1989). Para 100 µL de reação de PCR foram adicionados 4 µL de NaCl 5M. Em seguida, 2,5 vezes o volume de etanol absoluto foi adicionado e incubado a -20°C por 30 minutos. O produto da reação de PCR foi então centrifugado a 12000 rpm por 2 minutos, sendo que, em seguida, o pellet foi lavado e centrifugado três vezes com etanol 70% (estocado a -20°C). O *pellet* foi ressuspenso em 20µL de água ultra pura estéril durante 24 horas a 4°C e armazenado a -20°C.

3.6.2: Digestão enzimática:

O protocolo de digestão enzimática foi o mesmo para ambas as enzimas de restrição. O DNA foi digerido por 3 horas a 37°C ou overnight de acordo com o fabricante. Para a inativação da enzima *HhaI*, o material foi colocado em banho termostatizado por 20 minutos a 80°C e a 65°C por 10 minutos para a enzima *RsaI*.

3.6.3: Eletroforese em gel de agarose:

O produto da digestão foi separado por eletroforese em gel de Metaphor agarose 2% (Cambrex Bio Science/EUA). Padrões de peso molecular de 50 pb (Invitrogen); 100 pb (Fermentas) e de 1 Kb (Fermentas, Gibco) foram utilizados como padrão de tamanho molecular de fragmentos internos e externos. Os fragmentos de restrição foram visualizados mediante incubação em solução de brometo de etídio (1,5 µg/mL) e exposição à luz UV. A imagem obtida foi digitalizada e arquivada como descrito anteriormente.

3.7 - Identificação dos padrões de tipos dos antígenos H:

A identificação dos diferentes tipos de perfis de restrição dos antígenos H requer o conhecimento da margem de erro associada à determinação do tamanho do fragmento. O programa escolhido em geral minimiza este erro. Neste trabalho, as imagens foram submetidas ao programa Gel Compar II (Applied Maths/ Bélgica) para o estabelecimento dos diferentes perfis das amostras padrões. Uma tolerância entre os tamanhos dos fragmentos de 3,5% observada é aceitável, quando utilizado padrão de peso molecular interno. Quando o padrão de peso molecular externo é utilizado, há uma tolerância de 10%, devido ao erro associado com os fragmentos menores (Fields *et al.*, 1997; Machado *et al.*, 2000).

3.8 - Sequenciamento:

3.8.1 - Sequenciamento pela técnica de “cromossomo walking” ou “primer walking”:

Os fragmentos da PCR de cada amostra HNT foram primeiramente purificados pela enzima ExoSAP-IT (Promega/EUA). Para cada 5 µL de produto da PCR, foi adicionado 1 µL de ExoSAP-IT. Este material foi incubado a 37°C por 15 minutos e, posteriormente, a 80°C para inativação enzimática. Os produtos purificados foram estocados a -20°C até o momento de uso.

Para reação de sequenciamento foram utilizados 5,0 µL do produto da PCR purificado e 1,0 µL do iniciador (à montante ou à jusante) na mesma concentração utilizada na PCR. Além disso, foram adicionados 4,0 µL do pré-mix do kit DYEnamic™ ET Dye Terminator (Amersham Bioscience). O volume total da reação de sequenciamento foi de 10

μL. O programa da reação de sequenciamento foi de 30 ciclos (95°C/20seg, 50°C/15seg, 60°C 1min).

Após o término da reação de sequenciamento, o material foi precipitado conforme instruções do fabricante. A placa foi envolta em papel alumínio e colocada a -20°C *overnight* para ressuspender o material. As placas foram colocadas no MegaBace 1000 (Amersham Biosciences) em sistema capilar. O tempo e a voltagem da injeção das amostras foram de 60s e 750 V. Cada corrida foi realizada à temperatura de 44 °C.

Os resultados foram analisados de acordo com o item 3.9.6.

3.9. - Clonagem do gene *fliC*:

3.9.1: Inserção do gene *fliC* no Vetor de clonagem:

O protocolo de clonagem do gene *fliC* foi aprovado pela CiBio/IB/Unicamp – Parecer CIBio 01/2005 - Anexo 2) os fragmentos da PCR de cada amostra HNT, foram inseridos no vetor pGEM-T easy (Promega/EUA). Foram utilizados 5 μL do tampão (2X), 1 μL da enzima ligase, 1 μL do vetor pGEM-T easy, e 3 μL de produto da PCR. O mix foi mantido a 4°C *overnight* para a reação de ligação do fragmento de DNA no vetor, sendo em seguida estocado à -20°C.

3.9.2: Obtenção de células competentes:

As células competentes da linhagem DH5α foram utilizadas na transformação por “heat-shock”, baseando-se no protocolo de Sambrook *et al.*, (1989). No primeiro dia, uma colônia isolada da linhagem DH5α foi semeada em meio SOB (Tryptona, Extrato de levedura, NaCl, KCl, MgCl₂) e incubada a 37°C *overnight*. No segundo dia, 400 μL desta cultura foi adicionada a 40mL de meio SOB e colocado sob agitação até densidade óptica

(D.O) de 0,5nm. A cultura foi resfriada em banho com água e gelo, sendo em seguida centrifugada por 10min à 4000 rpm a 4°C. O sobrenadante foi descartado e o *pellet* foi ressuscitado em 20 mL de solução de CaCl₂ 50mM estéril e frio, e mantido no gelo por 3 horas. Em seguida, a suspensão foi centrifugada por 10 min a 4000 r.p.m a 4 °C, e ressuscitado em 4mL de CaCl₂ 50 mM contendo 15% de glicerol. As células competentes foram distribuídas em alíquotas de 100 µL em *ependorf* esterilizados e gelados, e congeladas imediatamente em nitrogênio líquido, sendo posteriormente armazenadas em biofreezer a -80°C.

3.9.3 - Transformação por “heat shock”:

As células competentes foram descongeladas em gelo. Foram adicionados 5µL do DNA clonado às células competentes e mantidos em gelo durante 30 min. Em seguida, este material foi aquecido a 42°C em banho termostatizado por 90 segundos e transferido para banho de gelo por 2 min. Em seguida, foi adicionado 1 mL de meio Luria-Bertani (LB) com glicose e incubado a 37°C sob agitação por 3 horas. Ao final, a suspensão foi centrifugada por 12000 rpm, por 1 minuto e plaqueada em meio LB-ágar contendo ampicilina (100 µg/mL de ampicilina, 20 mg/mL de X-Gal e 200 mg/mL de IPTG). As placas foram incubadas a 37°C “overnight”.

Foram selecionadas colônias brancas contendo o inserto e repicadas em meio LB contendo ampicilina.

3.9.4 - Extração plasmidial:

O protocolo para extração plasmidial foi realizado de acordo com o indicado por Ausubel *et al.* (1992). Uma colônia obtida no item anterior, foi inoculada em 5 mL de meio

LB contendo ampicilina, e deixada na estufa a 37°C *overnight*. Foi transferido 1,5 mL desta cultura para um tubo tipo eppendorf e centrifugado por 20 segundos a 12000 rpm. O sobrenadante foi descartado e as bactérias foram ressuspensas em 300µL de solução STET (8% sacarose, 5% Triton X-100, 50mM EDTA, 50m M Tris-HCl), e colocada no vórtex para homogeneização. Os *eppendorfs* foram colocados no gelo por 10 minutos e, em seguida, no banho maria a 100°C por 2 min. A suspensão foi centrifugada por 30 min a 12000 rpm, sendo que o sobrenadante foi transferido em um novo tubo tipo eppendorf, já contendo 200 µL de isopropanol gelado e deixado por 30min a -20°C. Foi feita uma nova centrifugação por 5 min a 12000 rpm. O sobrenadante foi removido até que todo o *pellet* tenha ficado seco, para ser ressuspendido em 50 µL de água ultra pura esterilizada.

A confirmação da presença do inserto no plasmídio foi feita aplicando-se 5 µl do produto da purificação de plasmídios em poços escavados em gel de agarose a 0,8% e submetido à eletroforese e visualizado após incubação em solução de brometo de etídeo sob luz UV. Em média, 50 a 100 colônias de cada placa tiveram seus plasmídios extraídos para verificar a presença do inserto completo.

Uma vez identificado os plasmídios contendo o inserto, foi feita a PCR utilizando os iniciadores propostos por Fields *et al* (1997), para verificar se o tamanho do inserto encontrado era do tamanho esperado. Para obter a sequência completa do gene *fliC*, os iniciadores: T7: TAATACGACTCACTATAGGG (à montante) e M13: AGCGGATAACAATTCACACAGG (à jusante) foram utilizados.

3.9.5 - Sequenciamento dos clones:

Os fragmentos da PCR de cada amostra HNT clonada foram primeiramente purificados pela enzima *ExoSAP-IT* (Promega), sendo em seguida utilizados na reação de sequenciamento (item 3.8.1).

Os iniciadores utilizados tanto para a reação de PCR, quanto para a reação de sequenciamento foram os iniciadores T7 e M13.

Após o sequenciamento, as sequências foram analisadas no programa Phred, Phrap, Consed (<http://www.washington.edu>) e no programa de alinhamento ClustalW. Cada reação de sequenciamento gerou uma sequência nucleotídica com boa qualidade de leitura de aproximadamente 500 pb e, para cada sequência obtida, voltamos aproximadamente 100 pb para a construção dos novos iniciadores internos (Tabela 5).

Tabela 5: Iniciadores utilizados nas reações de sequenciamento do gene *fliC* das amostras de *E. coli* HNT:

Amostras	Iniciadores
1C e 5C	FliC1: AACTAACGGTACTAACTCTGACA FliC1rev: CCACTACCGTCTCAGCTTT
2C	FliC(F)1: ATGGCACAAGTCATTAATACCCAAC FliC2cF: TTACAGGTTGGTGCAAACAG FliC2cR: CCAACCATGGCAACACCGTA FliC(F)2: CTAACCCTGCAGCAGAGACA
3C	FliC(F)1: ATGGCACAAGTCATTAATACCCAAC FliC1: AACTAACGGTACTAACTCTGACA FliC3cF: AATATCGATGGCGCGCAGAGA FliC(F)2: CTAACCCTGCAGCAGAGACA
4C	FliC1: AACTAACGGTACTAACTCTGACA FliC4Cf: TATGTTTCAGTTGGGTAC
6C	FliC3: CAGTTCAACGGCGTGAA FliC3rev: GTGGTGTGTTTCAGGTTGGT
7C	FliC7cF: CAGCTGATCGGCCTTTCGGT
8C, 9C, 11A	FliC2: AATTTGTCGATGGAGCTGAT FliC2rev: AGTTCAACGGCGTGAACGT
2A	FliC(F)1: ATGGCACAAGTCATTAATACCCAAC FliC(F)2: CTAACCCTGCAGCAGAGACA FliC1: AACTAACGGTACTAACTCTGACA
3A, 10A	FliC3: CAGTTCAACGGCGTGAA

Tabela 5. Continuação

Amostras	Iniciadores
5A	FliC3: CAGTTCAACGGCGTGAA FliC5A: GCTTCTTTGACTGAGGCTAG
7A	FliC3: CAGTTCAACGGCGTGAA FliC3rev: GTGGTGTTGTTTCAGGTTGGT FliC7af: GAACACTCGCGGAAGCGA
8A	FliC(F)1: ATGGCACAAGTCATTAATACCCAAC FliC8af: CCTGACAGCAGCGAAAATGG FliC8ar: CAAATCGACATCGGTGAAAT FliC(F)2: CTAACCCTGCAGCAGAGACA
9A	FliC1: AACTAACGGTACTAACTCTGACA FliC3rev: GTGGTGTTGTTTCAGGTTGGT
13A	FliC(F)1: ATGGCACAAGTCATTAATACCCAAC FliC13aF: GGCTTGATAACTTTAGGCGTA FliC(F)2: CTAACCCTGCAGCAGAGACA
14A	FliC14aR: GGTATAATGACTTGTGCCA
15A	FliC2rev: AGTTCAACGGCGTGAACGT

3.9.6 - Montagem dos *Contigs* e análise das sequências

Para a montagem dos *contigs* foi utilizado o programa Phred, Phrap, Consed. Cada eletroesferograma foi analisado para verificar a qualidade da leitura do sequenciamento. Cada base que apresentou um resultado discrepante foi conferida (Figuras 2 e 3). As sequências gênicas obtidas na reação de sequenciamento foram comparadas aos diferentes genes de *fliC* armazenados no banco de dados do “National Center for Biotechnology Information” (NCBI), através do sistema BLAST. O programa ClustalW foi utilizado para o alinhamento das sequências.

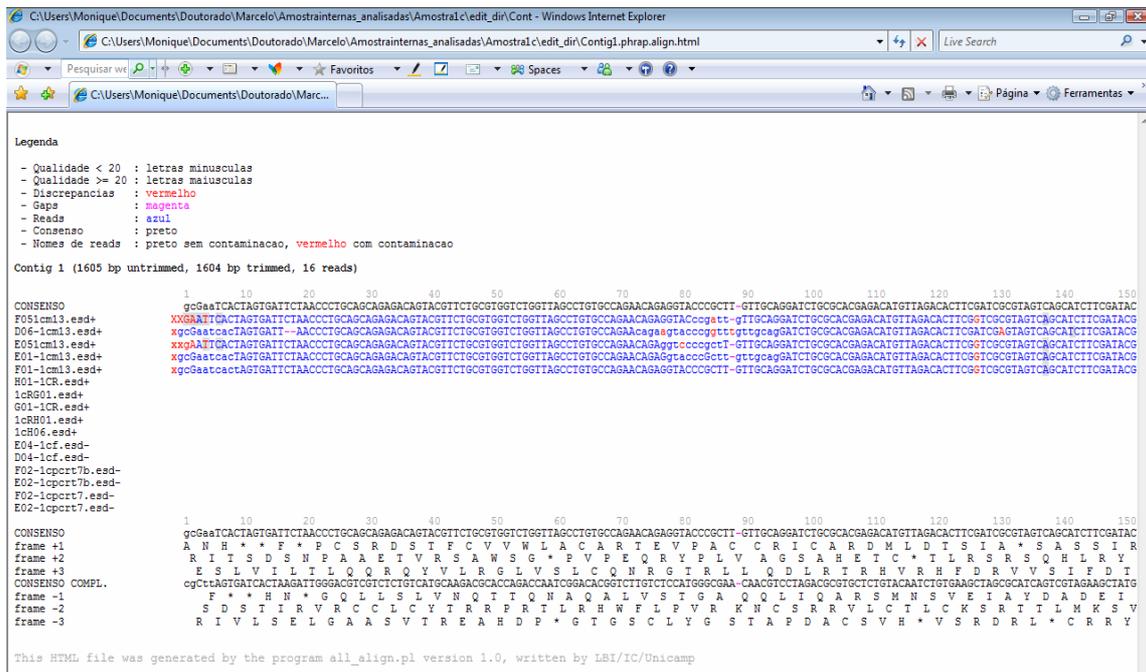


Figura 2: Resultado da montagem final do contig da amostra 1C através do programa Phred, Phrap, Consed.

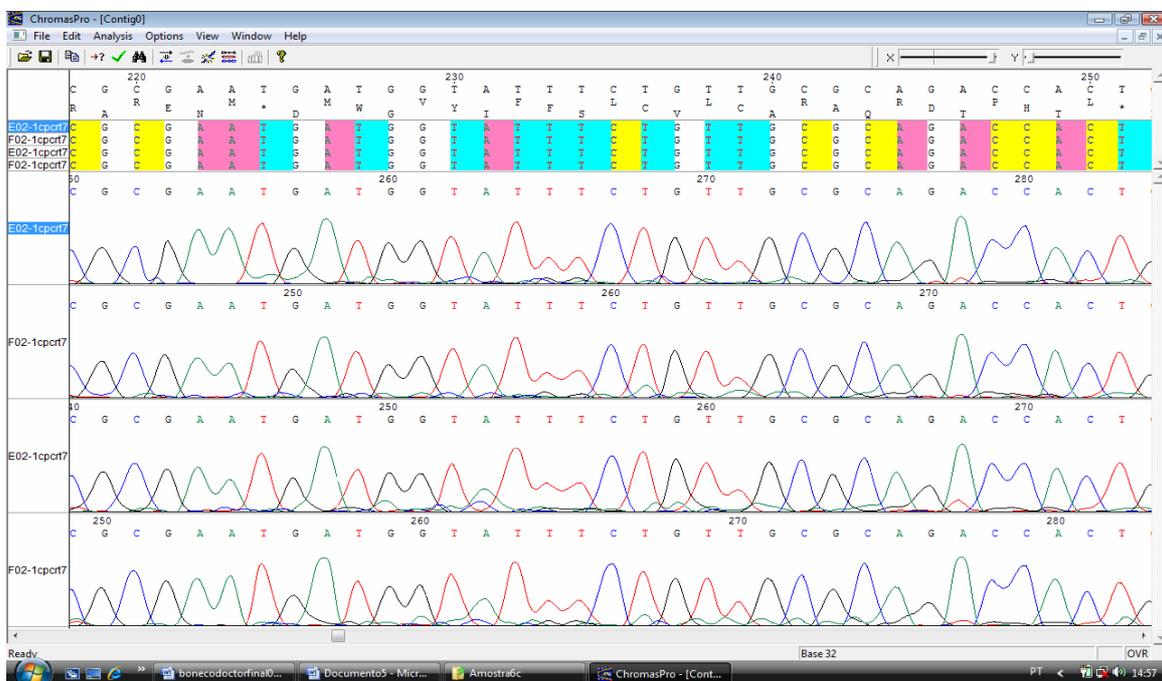


Figura 3: Eletroesferogramas da amostra 1C, visualizados pelo programa ChromasPro.

3.10 - Preparação e obtenção do antígeno “H” das amostras de *E. coli* HNT:

3.10.1- Ativação da motilidade:

As amostras de *E. coli* HNT foram semeadas em tubos U contendo meio de cultura semi-sólido MILi (Toledo *et al*; 1982) e incubadas por 24 horas a 37° C, ou até que o crescimento tenha alcançado a extremidade oposta à inicialmente semeada. Este procedimento foi repetido por cinco passagens consecutivas. Observando-se que o repique seguinte seja obtido na extremidade oposta daquela inicialmente semeada.

3.10.2- Obtenção do antígeno H:

A partir do último tubo U semeado as amostras bacterianas foram repicadas para tubos contendo 7 mL de caldo BHI e incubadas a 37°C em estufa com agitação orbital por 6-8 horas ou até que a turbidez fosse comparável ao tubo 2 da escala de McFarland. Em seguida, foram inativadas pela adição de igual volume de salina fisiológica contendo 0,5% de formalina e incubada *overnight* em temperatura ambiente. As suspensões foram mantidas sob refrigeração até o momento de seu uso.

3.10.3 - Obtenção de anti-soros anti-antígenos HNT:

Para cada amostra de *E. coli* HNT foram utilizados pelo menos dois coelhos albinos adultos, de doze semanas de idade (autorizado pelo Cômite de Ética Experimental – Protocolo 781-1- Anexo1). A suspensão bacteriana, obtida no item 3.9.2 foi utilizada como antígeno. Os coelhos foram inoculados com a suspensão bacteriana na veia marginal da orelha com volume de 0,5; 1,0; 2,0; 3,0 e 4,0 mL em intervalos de 4-5 dias (Ewing, 1986). De 6-8 dias após a última inoculação, os animais foram sangrados (sangria de prova) e os soros titulados. Uma vez determinado que altos títulos de aglutininas estavam presentes, os

animais foram anestesiados (ketamina/xilazina 35-50/5-10 mg/Kg) por via intramuscular e exsanguinados. Após a separação dos soros estes, foram clarificados por centrifugação, aliquotados e congelados a – 80°C.

3.10.4-Aglutinação em tubo:

Para a determinação do título dos anti-soros anti-H foram realizadas reações de aglutinação em tubos. Em baterias de tubos de vidro tipo Kahn, foram adicionados 200 µL de diferentes diluições dos anti-soros a serem titulados (Título inicial 1:100). A estas diluições, foram adicionados 200 µL dos antígenos H. A mistura foi incubada em banho termostático a 45°C por 3 horas.

Reação positiva foi aquela em que observamos a formação de uma película no fundo do tubo e o sobrenadante com aspecto transparente, enquanto que, na reação negativa o sobrenadante permaneceu turvo. Consideramos como título, a maior diluição do anti-soro que aglutinou com o antígeno homólogo.

3.10.5 - Soroabsorção do anti-soro do antígeno H das amostras HNT:

Para um volume do anti-soro diluído (1:50), igual volume de suspensão foi adicionado. A mistura foi incubada em banho termostático a 50°C por 1 hora e então centrifugada sob refrigeração. O anti-soro diluído foi então adicionado a um segundo volume da suspensão de antígeno H, reincubado por 1 hora em banho termostático e então centrifugado (Ewing, 1986). O anti-soro H absorvido foi usado em testes de aglutinação.

3.11: Análise da estrutura secundária:

A sequência de aminoácidos foi obtida pelo programa MEGA e foi submetida ao endereço eletrônico de bioinformática estrutural: <http://www.sbg.bio.ic.ac.uk/phyre> para verificar a similaridade com outras estruturas secundárias já descritas.

IV-RESULTADOS

4.1-Reação em Cadeia da Polimerase (PCR):

O gene *fliC* das 52 amostras de *E. coli* padrão para os antígenos H e das 20 amostras de *E. coli* HNT, foi amplificado pelo iniciador FliC(M) por PCR. Do mesmo modo, 50 amostras padrão para antígeno H e as 20 amostras de *E. coli* HNT foram amplificadas pelo iniciador FliC(F). Alguns resultados das ampliações das amostras padrões estão exemplificados nas Figuras 4 e 5. Na Tabela 6 estão apresentados os tamanhos dos produtos da PCR para as 53 amostras padrão.

As amostras com antígeno H padrão resultaram em produtos de PCR que variaram em tamanho molecular de 925pb a 2635pb utilizando o iniciador FliC(M), e de 1080 a 2850 pb utilizando o iniciador FliC(F).

O *fliC* do antígeno H53 não foi amplificado pelos iniciadores propostos por Machado *et al* (2000) e as amostras H17, H25 e H53 não foram amplificadas utilizando os iniciadores propostos por Fields *et al.*, 1997.

As amostras com antígeno HNT, quando utilizado o iniciador FliC(M), resultaram em produtos de PCR que variaram em tamanho molecular de 1180pb a 2605pb e de 1300 a 2690pb quando utilizado o iniciador FliC(F) (Figuras 6 e 7, Tabela 7).

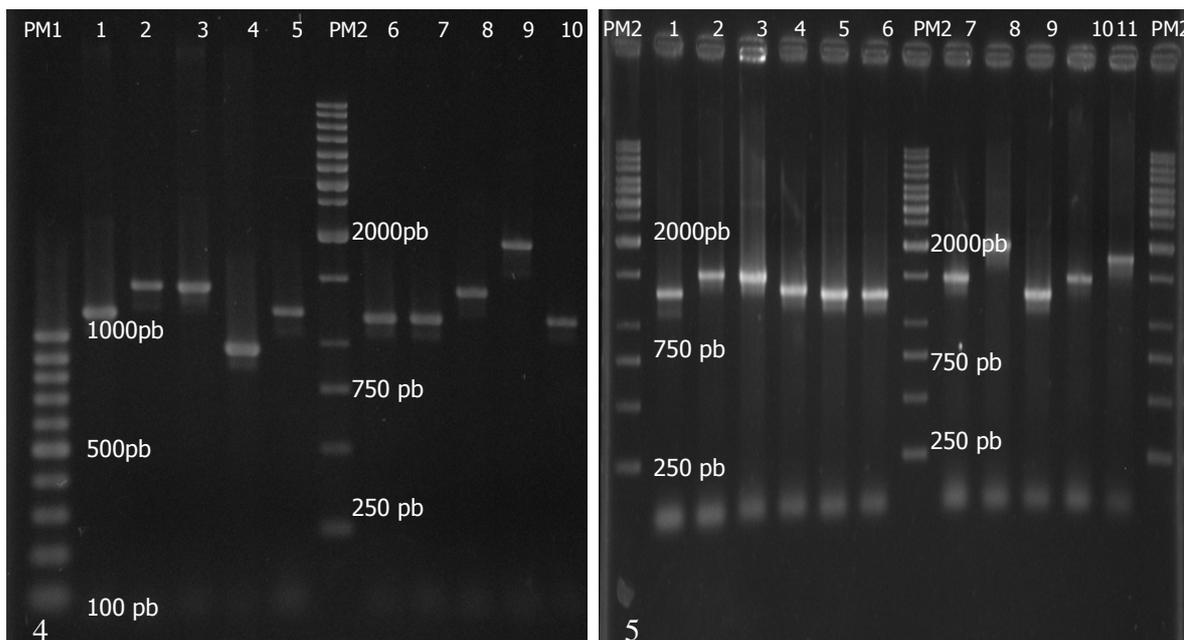


Figura 4: Perfil eletroforético em gel de agarose a 1,5% das amostras de *E. coli* padrão para antígeno H, utilizando o iniciador FliC(M). PM1: Padrão de peso molecular de 100 pb (Fermentas); PM2: Padrão de peso molecular de 1 Kb (Fermentas); Linha 1: *E. coli* H1; Linha 2: *E. coli* H2; Linha 3: *E. coli* H3; Linha 4: *E. coli* H4; Linha 5: *E. coli* H5; Linha 6: *E. coli* H6; Linha 7: *E. coli* H7; Linha 8: *E. coli* H8; Linha 9: *E. coli* H9; Linha 10: *E. coli* H10.

Figura 5: Perfil eletroforético em gel de agarose a 1,5% das amostras de *E. coli* padrão para antígeno H, utilizando o iniciador FliC(F). PM2: Padrão de peso molecular de 1 Kb; Linha 1: *E. coli* H1; Linha 2: *E. coli* H2; Linha 3: *E. coli* H3; Linha 4: *E. coli* H5; Linha 5: *E. coli* H6; Linha 6: *E. coli* H7; Linha 7: *E. coli* H8; Linha 8: *E. coli* H9; Linha 9: *E. coli* H10; Linha 10: *E. coli* H11; Linha11: *E. coli* H12.

Tabela 6: Tamanhos dos produtos da PCR para os antígenos H de *E. coli* padrão (H1 a H56), utilizando os iniciadores Flic(M) e FliC(F).

Antígeno H	FliC(M) (pb)	FliC(F) (pb)
H1	1400	1260
H2	1650	1450
H3	1540	1500
H4	980	1180
H5	1345	1290
H6	1280	1300
H7	1290	1260
H8	1520	1480
H9	2030	1970
H10	1250	1260
H11	1545	1450
H12	1705	1745
H14	1280	2035
H15	1730	1630
H16	1590	1610
H17	950	_*
H18	1260	1630
H19	1245	1250
H20	1870	1800
H21	1325	1300
H23	1200	1330
H24	1800	1750
H25	1280	_*
H26	1820	1635
H27	1330	1230
H28	1250	1195
H29	1355	1240
H30	1800	1730
H31	1255	1250
H32	1760	1690
H33	1150	1290
H34	1650	1665
H35	1420	1425
H36	2635	2850
H37	1770	1650
H38	1170	1375
H39	1300	1225
H40	1595	1530
H41	1770	1670
H42	1245	1290
H43	1575	1510
H44	925	1080
H45	1750	1730
H46	1800	1710
H47	1540	1510
H48	1565	1470
H49	1640	1680
H51	1770	1790
H52	1200	1350
H53	_*	_*
H54	1900	1770
H55	1250	1300
H56	1290	1300

* -: não houve reação positiva.

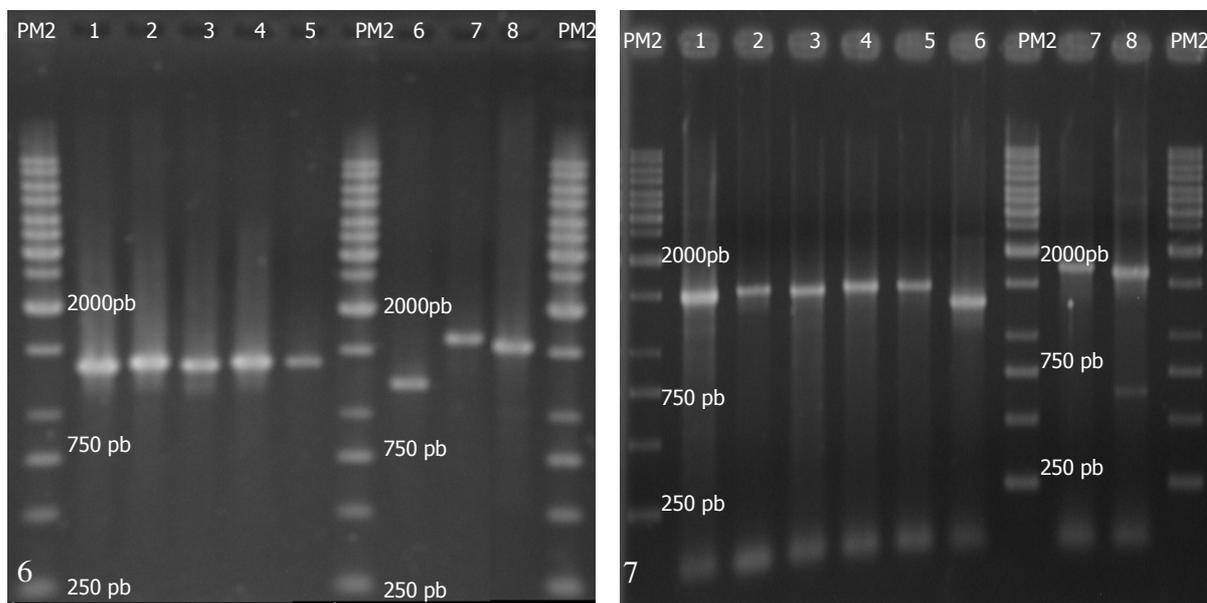


Figura 6: Perfil eletroforético em gel de agarose 1,5% das amostras de *E. coli* com antígeno H não-tipável, utilizando o iniciador FliC(M). PM2: Padrão de peso molecular de 1 Kb; Linha 1: *E. coli* VTH-115 (1C); Linha 2: *E. coli* VTH-110 (2C); Linha 3: *E. coli* VTH-118 (3C); Linha 4: *E. coli* 28011a (4C); Linha 5: *E. coli* 33141a (5C); Linha 6: *E. coli* 46103b (6C); Linha 7: *E. coli* 40478b (7C); Linha 8: *E. coli* 48629c(2)(9C).

Figura 7: Perfil eletroforético em gel de agarose 1,5% das amostras de *E. coli* com antígeno H não-tipável, utilizando os iniciadores FliC(F). PM2: Padrão de peso molecular de 1 Kb; Linha 1: *E. coli* VTH-115 (1C); Linha 2: *E. coli* VTH-110 (2C); Linha 3: *E. coli* VTH-118 (3C); Linha 4: *E. coli* 28011a (4C); Linha 5: *E. coli* 33141a (5C); Linha 6: *E. coli* 46103b (6C); Linha 7: *E. coli* 40478b (7C); Linha 8: *E. coli* 48629c(2)(9C).

Tabela 7 - Tamanho molecular do gene *fliC* em pares de base (pb) das amostras de *E. coli* HNT de origem humana, obtidos pela PCR:

Nº	Código Original	<i>fliC</i> (M) (pb)	<i>fliC</i> (F) (pb)
1C	VTH 15	1360	1420
2C	VTH 110	1390	1470
3C	VTH 118	1410	1430
4C	28011a	1395	1470
5C	33141a	1350	1445
6C	46103B	1180	1300
7C	40478B	1625	1790
8C	48629c(1)	1695	1740
9C	48629c(2)	1620	1670
2A	01-03443	1955	2050
3A	00-04915	1460	1495
5A	00-04447	1590	1660
7A	00-08242	1660	1720
8A	00-03034	1710	1785
9A	00-07153	1550	1775
10A	00-00848	1665	1725
11A	00-05951	1630	1500
13A	00-08712	1460	1555
14A	99-01406	2605	2690
15A	00-09775	1630	1720

4.2 - Análise da Digestão Enzimática:

Após a digestão enzimática do produto da amplificação do gene *fliC*, utilizando os iniciadores *FliC*(M) e a enzima *HhaI*, verificamos que os antígenos padrões H3 e H8; H6, H10, H19 e H27; H11 e H47; H23 e H43; H28 e H42; H55 e H56, apresentaram perfis de restrição semelhantes (Figura 8, Tabela 8).

Denominamos o perfil de restrição da amostra H1 de P1, H2 de P2 e assim sucessivamente. Aquela amostra que apresentou perfil de restrição semelhante fora denominada com a designação do perfil original, por exemplo: H3-P3, H8-P3, H6-P6, H10-P6, H19-P6.

RFLP

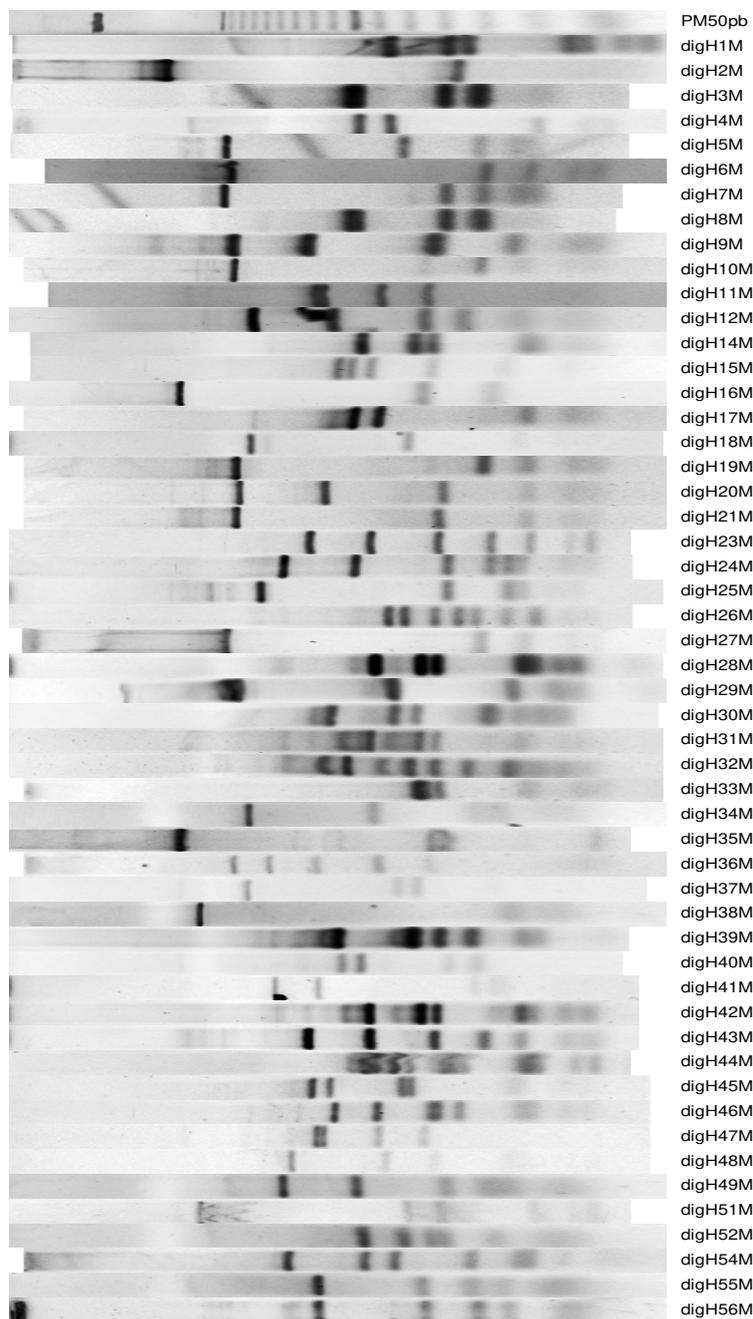


Figura 8: Perfil eletroforético da digestão dos amplicons obtidos com os iniciadores *fliC(M)* com a enzima *HhaI*, nas amostras de *E. coli* H padrão, gerado pelo programa GELCOMPAR II. PM50 pb – padrão de peso molecular de 50 pb (Invitrogen); dig – digerido.

Quando utilizados os iniciadores fliC(F) e a enzima *RsaI* foi observado que as amostras padrão H1, H28 e H31; H3 e H8; H2, H30 e H35; H9 e H14; H11 e H47; H7, H19 e H27; H55 e H56; quando analisadas através do programa Gel Compar II, apresentaram perfis de restrição similares, (cabe ressaltar que baseando-se em Machado *et al* (2000), dois fragmentos foram considerados idênticos quando seus tamanhos não diferiram mais que 3,5%). Os resultados estão apresentados na Figura 9 e Tabela 8.

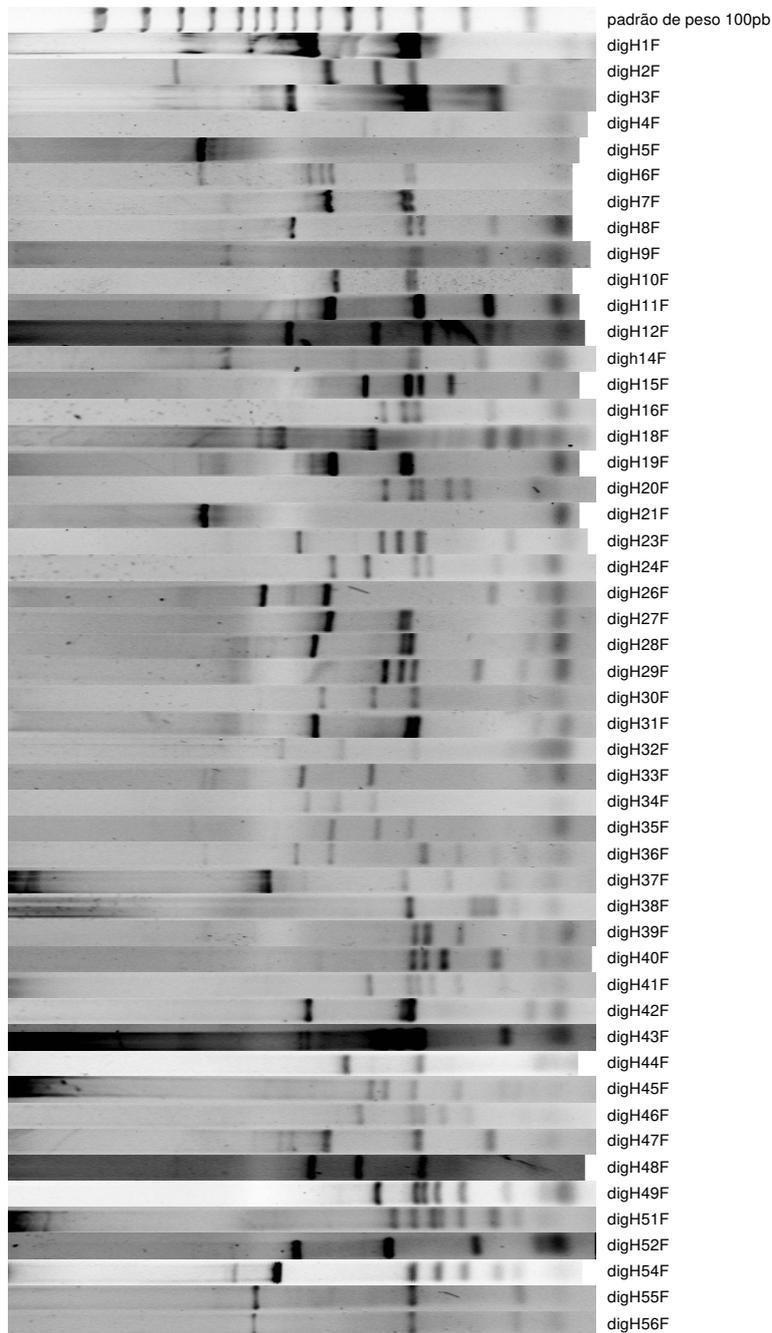


Figura 9: Perfil eletroforético da digestão dos amplicons obtidos com os iniciadores fliC(F) com a enzima *RsaI*, nas amostras de *E. coli* H padrão, gerado pelo programa GELCOMPAR II. PM100 pb - padrão de peso molecular de 100 pb; dig – digerido.

Tabela 8: Resultados do polimorfismo da PCR-RFLP, utilizando os iniciadores FliC(M) e FliC(F) para as amostras de *E. coli* padrão.

Antígeno H padrão	PCR - RFLP (<i>HhaI</i>) em pb	Padrão Molecular (M)	PCR -RFLP (<i>RsaI</i>) em pb	Padrão Molecular (F)
H1	285, 195, 170, 70	P1	630, 330, 310	P1
H2	1370, 180	P2	570, 410, 320, 120	P2
H3	360, 350, 150, 110	P3	720, 320, 290, 150	P3
H4	340, 285, 100, 60, 50	P4	440, 255, 230	P4
H5	770, 260, 160, 120	P5	1290	P5
H6	750, 150, 110, 70, 50	P6	565, 335, 320	P6
H7	790, 200, 150, 120, 105	P7	570, 340, 330	P7
H8	360, 350, 150, 110	P3	710, 330, 295, 150	P3
H9	735, 470, 215, 120, 70	P8	1115, 315, 170	P8
H10	740, 160, 115, 70, 50	P6	540, 320, 310	P9
H11	445, 435, 300, 220	P9	560, 300, 160	P10
H12	655, 410, 230, 175, 120	P10	730, 410, 280, 160, 130	P11
H14	340, 245, 220, 110, 105, 60	P11	1115, 315, 170	P8
H15	390, 360, 320, 215, 130	P12	440, 325, 300, 230, 95	P12
H16	1220, 230, 140	P13	390, 330, 300, 150	P13
H17	355, 305, 110, 70	P14	-	- ^a
H18	660, 250	P15	760, 420, 150, 120, 95	P14
H19	750, 150, 110, 70, 50	P6	550, 335, 325	P7
H20	710, 420, 200, 110, 60	P16	385, 315, 300, 230, 200	P15
H21	720, 210, 110, 70, 55	P17	1275	P16
H23	460, 320, 210, 145, 105, 70	P18	680, 390, 350, 300, 130	P17
H24	540, 340, 195, 145, 135	P19	550, 440, 310, 275, 140	P18
H25	625, 195, 130, 125	P20	-	-
H26	290, 260, 210, 180, 160, 130, 100	P21	860, 570, 150	P19
H27	740, 155, 110, 70, 50	P6	560, 340, 330	P7
H28	315, 235, 210, 110, 100, 80, 70	P22	620, 335, 320	P1
H29	740, 280, 125, 80, 70	P23	380, 340, 310, 175, 110	P20
H30	410, 280, 240, 150, 115, 100, 85	P24	590, 420, 310, 120	P2
H31	380, 320, 285, 240, 215, 115, 65	P25	610, 320, 310	P1
H32	430, 370, 300, 250, 210, 170, 130, 80	P26	760, 525, 305	P21
H33	235, 230, 210, 105	P27	670, 420	P22
H34	670, 315, 160, 135	P28	640, 535, 415	P23
H35	1210, 220, 195	P29	570, 410, 310, 120	P2
H36	740, 595, 445, 305, 220	P30	690, 560, 290, 210, 150, 105	P24
H37	680, 270, 240	P31	840, 330, 230, 130	P25
H38	995, 130	P32	320, 180, 165, 150, 120	P26
H39	390, 250, 210, 170, 110, 105	P33	310, 280, 270, 210, 110, 90	P27
H40	380, 340, 195, 160	P34	315, 290, 250, 145, 85	P28
H41	570, 440, 160, 130	P35	430, 320, 300, 270, 215, 130	P29
H42	320, 235, 210, 115, 70, 60	P22	640, 320, 310, 95	P30
H43	465, 320, 215, 150, 115, 75	P18	390, 350, 300, 290, 130	P31
H44	335, 315, 275, 250, 190, 110, 70	P36	710, 610, 500, 300, 90	P32
H45	455, 410, 260, 250, 115	P37	430, 380, 315, 215, 140, 130, 110	P33
H46	400, 310, 215, 180, 110, 80	P38	460, 315, 300, 250, 200, 105	P34
H47	445, 430, 300, 230	P9	575, 300, 155	P10
H48	515, 290, 210, 125, 100	P39	630, 470, 290, 95	P35
H49	540, 350, 200, 150, 130	P40	410, 310, 290, 260, 210, 130, 70	P36
H51	1000, 250, 205, 105	P41	360, 310, 270, 210, 150, 115	P37
H52	335, 260, 220, 140	P42	695, 375, 180, 90	P38
H53	-	-	-	-
H54	525, 330, 275, 165, 115, 110	P43	780, 315, 255, 200, 145, 115	P39
H55	440, 235, 165, 130, 80, 60	P44	900, 305, 105	P40
H56	440, 230, 160, 125, 80, 60	P44	900, 305, 105	P40

F-Fields

M-Machado

Os resultados das digestões enzimáticas das vinte amostras HNT amplificadas com os iniciadores Flic(M) e Flic(F) pela PCR, e digeridos pelas enzimas de restrição *HhaI* e *RsaI*, estão demonstrados na Tabela 9 e nas Figuras 10 e 11.

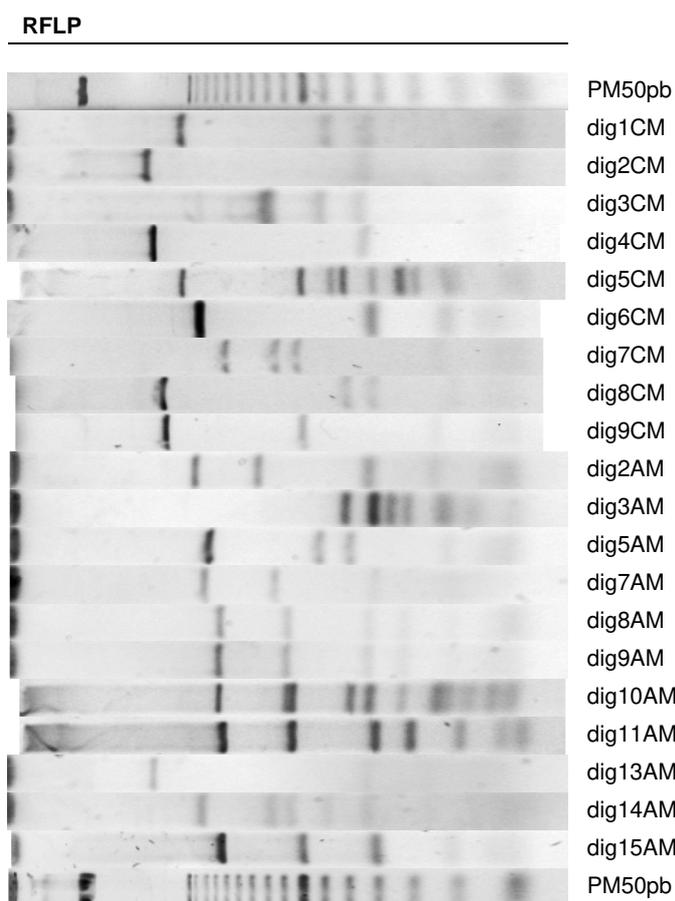


Figura 10: Perfil eletroforético do produto da digestão de *HhaI*, dos amplicons obtidos com os iniciadores fliC(M), nas amostras de *E. coli* HNT, gerado pelo programa GELCOMPAR II. PM50 pb – padrão de peso molecular de 50 pb (Invitrogen); dig – digerido.

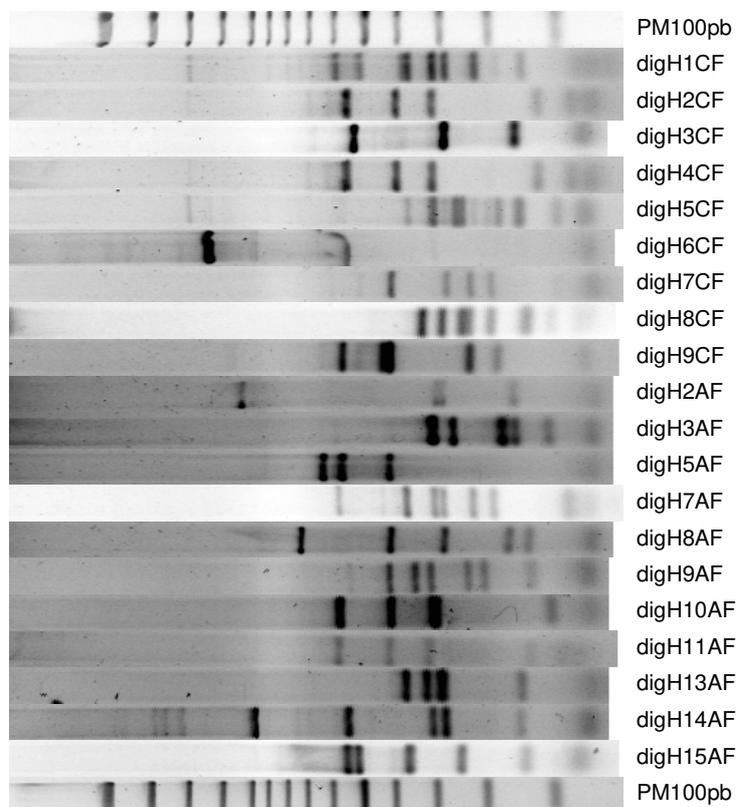


Figura 11: Perfil eletroforético do produto da digestão de *RsaI*, dos amplicons obtidos com os iniciadores *fliC(F)*, nas amostras de *E. coli* HNT, gerado pelo programa GELCOMPARE II. PM100 pb – padrão de peso molecular de 100 pb; dig – digerido.

Utilizando o programa GEL COMPARE II, os perfis de restrição obtidos na digestão enzimática com a endonuclease *HhaI* das amostras de *E. coli* HNT foram comparados com aqueles observados nas amostras de *E. coli* padrão. Sete amostras de *E. coli* HNT apresentaram perfis de restrição semelhantes aos observados nas amostras de *E. coli* padrão (Tabela 9; Figura 12). Já a digestão com a endonuclease *RsaI*, permitiu a identificação de oito perfis de restrição semelhantes (Tabela 9, Figura 13).

Tabela 9: Resultados do polimorfismo da PCR-RFLP, utilizando os iniciadores *FliC(M)* e *FliC(F)* nas amostras de *E. coli* HNT:

<i>E. coli</i> HNT	PCR-RFLP utilizando <i>FliC(M)</i>			PCR-RFLP utilizando <i>FliC(F)</i>		
	PCR-RFLP (<i>HhaI</i>) em pb	Padrão Molecular (M)	Tamanho molecular de <i>fliC</i> (pb)	PCR-RFLP (<i>RsaI</i>) em pb	Padrão Molecular (F)	Tamanho molecular de <i>fliC</i> em pb
1C	895, 295, 220	-	1360	595, 520, 375, 320, 285, 230, 140	-	1420
2C	1305, 220	P2	1390	560, 410, 320, 125	P2	1470
3C	445, 435, 315, 220	P9	1410	565, 290, 150	P10	1430
4C	1350, 180	P2	1395	565, 400, 315, 130	P2	1470
5C	875, 360, 285, 260, 210, 170, 150, 110	-	1350	310, 260, 185, 150, 105	-	1445
6C	735, 210, 115	-	1180	1320	-	1300
7C	615, 430, 370, 120	-	1625	415, 280, 230, 190, 95	-	1790
8C	1005, 260, 210, 100	P41	1695	335, 290, 250, 240, 190, 130, 105	-	1740
9C	1040, 350, 120	-	1620	570, 440, 420, 235, 180	-	1670
2A	760, 480, 215, 130	-	1955	1100, 320, 170,	P8	2050
3A	260, 210, 180, 160, 115, 85	-	1460	315, 270, 170, 150, 105	-	1495
5A	690, 310, 250, 110	-	1590	645, 570, 420	-	1660
7A	710, 425, 210	-	1660	580, 375, 310, 290, 225, 185	-	1720
8A	645, 400, 225, 165, 120	P10	1710	715, 430, 300, 175, 145	P11	1785
9A	635, 400, 215, 165, 115	P10	1550	420, 355, 320, 240, 205, 130	-	1775
10A	650, 380, 250, 215, 165, 120, 85, 65, 50	-	1665	560, 420, 320, 125	P2	1725
11A	615, 375, 205, 155, 95, 60, 50	-	1630	560, 420, 330, 135	P2	1500
13A	1350, 180	P13	1460	375, 325, 295, 140	P13	1555
14A	720, 435, 380, 290, 220	-	2065	1005, 550, 310 280, 140	-	2690
15A	625, 350, 205, 105	-	1630	555, 520, 370, 250, 135	-	1720

-: sem semelhança

RFLP

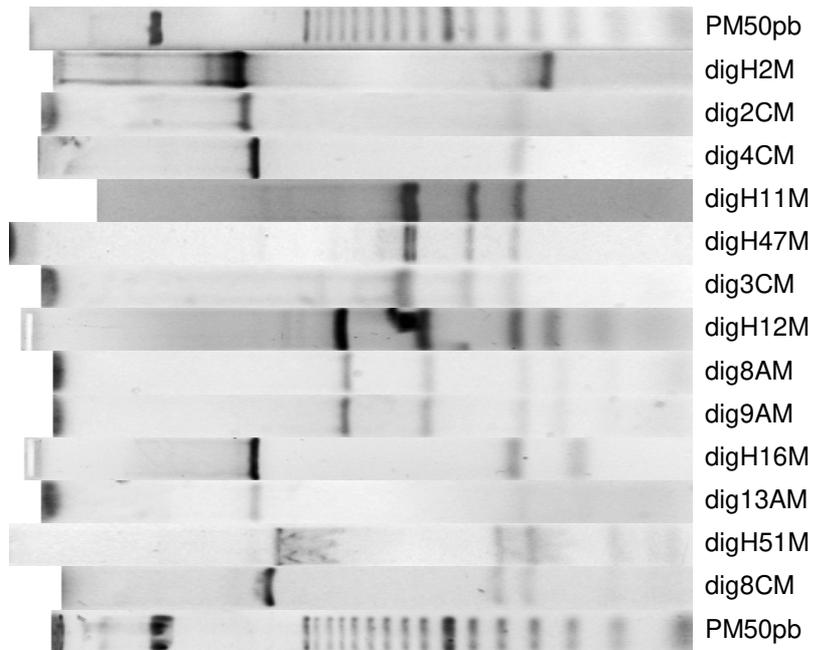


Figura 12: Resultado da análise de similaridade entre os produtos da digestão dos amplicons das amostras de *E. coli* padrão e *E. coli* HNT, obtidos com os iniciadores FliC(M) e com a enzima *HhaI*, gerado pelo programa GEL COMPAR II. PM100 pb – padrão de peso molecular de 100 pb (Fermentas); dig – digerido.

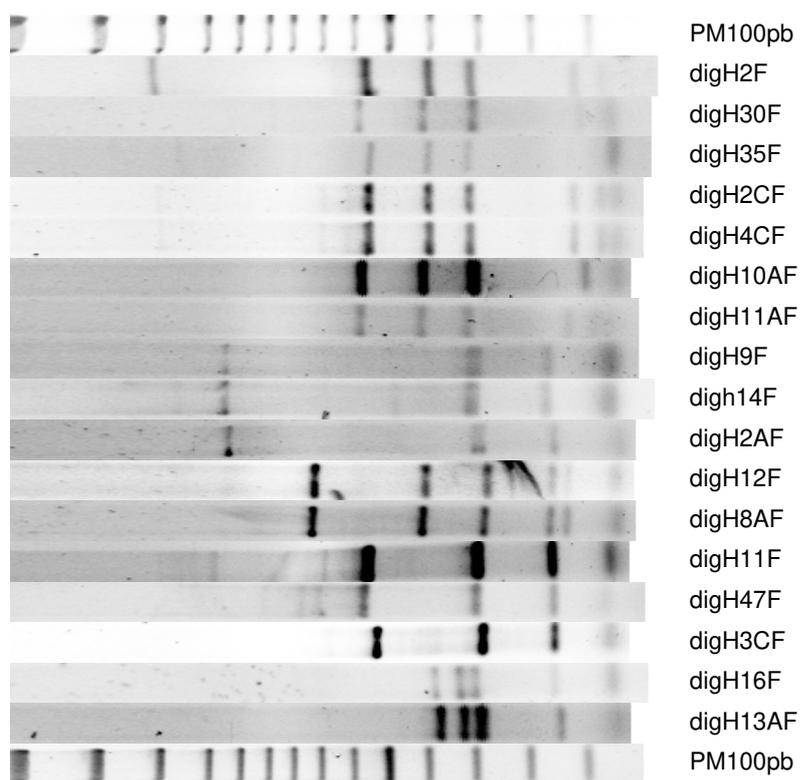


Figura 13: Resultado da análise de similaridade entre os produtos da digestão dos amplicons das amostras de *E. coli* padrão e *E. coli* HNT, obtidos com os iniciadores FliC(F) e com a enzima *RsaI*, gerado pelo programa GEL COMPAR II. PM100 pb – padrão de peso molecular de 100 pb (Fermentas); dig – digerido.

4.3 - PCR dos genes de variação de fase:

Após a validação dos testes de PCR, onde foram utilizadas as amostras padrões H3, H53 e H54 para padronização dos genes de variação de fase: *fljA* (repressor), *flkA* e *flmA* (flagelina) respectivamente, foi realizada a PCR de todas as vinte amostras HNT. Neste estudo todas as amostras HNT foram negativas para os genes associados à variação de fase.

4.4 - Sorologia:

Todos os antígenos das amostras HNT (2C, 3C, 4C, 8C, 2A, 8A, 9A, 10A, 11A, 13A), que apresentaram perfis de restrição semelhantes àqueles encontrados nas amostras H padrão, foram titulados frente os respectivos anti-soros padrões. Não foram observadas reações positivas em tais ensaios.

As amostras de *E. coli* HNT (1C, 5C, 6C, 7C, 9C, 3A, 5A, 7A, 14A, 15A) em que não foram encontrados perfis de restrição semelhantes aos perfis de restrição das amostras de *E. coli* padrão e, as amostras de *E. coli* HNT (2C, 3C, 4C, 8C, 2A, 8A, 9A, 10A, 11A, 13A) que apresentaram um perfil de restrição semelhante àqueles das amostras padrão, mas foram negativas nos testes sorológicos, foram selecionadas para a obtenção dos anti-soros.

Os anti-soros foram titulados frente a seus respectivos antígenos homólogos (Tabela 10).

Tabela 10: Resultado da titulação dos anti-soros de *E. coli* HNT frente ao seu antígeno homólogo:

N° da amostra	Título frente ao antígeno homólogo	N° da amostra	Título frente ao antígeno homólogo
1C	>12.800	2A	>12.800
2C	NR	3A	>12.800
3C	>12.800	5A	>12.800
4C	NO	7A	>3.200
5C	>12.800	8A	>12.800
6C	NO	9A	>12.800
7C	NR	10A	>12.800
8C	>12.800	11A	NO
9C	>12.800	13A	>12.800
		14A	NR
		15A	>12.800

NR: Não reagiu contra o antígeno homólogo

NO: Não obtido

Posteriormente, os anti-soros produzidos com as amostras de *E. coli* HNT foram titulados frente aos 53 antígenos flagelares padrão. Apesar dos antígenos das amostras HNT não terem reagido com os soros padrões, os anti-soros das *E. coli* HNT reagiram com alguns antígenos flagelares padrões. Os resultados desses ensaios são mostrados na tabela 11. O anti-soro da amostra 1C reagiu com o antígeno H11 (Título 1:6400), o anti-soro da amostra 2C reagiu com o antígeno H2 (Título 1: 12800), o anti-soro da amostra 3C reagiu com o antígeno padrão H11 (título 1:12.800), o anti-soro da amostra 2A reagiu com o antígeno padrão H9 (Título 1:12800), o anti-soro da amostra 8A reagiu com o antígeno padrão H12 (Título 1:12800), o anti-soro da amostra 10A reagiu com o antígeno padrão H30 (Título 1:12800), o anti-soro da amostra 13A reagiu com o antígeno padrão H16

(Título 1:12800).

A amostra 1C não apresentou similaridade com o antígeno H11 através da PCR-RFLP. Já as amostras 2C, 3C, 8A, apresentaram similaridade através da PCR-RFLP, apenas com estes antígenos que reagiram na sorologia. As amostras 3A, 10A, 13A, apresentaram similaridade com dois antígenos através da PCR-RFLP, porém na sorologia elas foram positivas para apenas um destes (Tabela 11).

Tabela 11: Resultados da PCR-RFLP e da sorologia obtidos entre as amostras de *E. coli* HNT e *E. coli* padrão:

Amostras	PCR-RFLP FliC(M)	PCR-RFLP FliC(F)	Título dos anti-soros das amostras HNT frente os antígenos padrões
1C	-	-	NR
2C	P2	P2	H2 (1:12800)
3C	P9	P10	H11 (1:12800)
5C	-	-	NR
7C	-	-	NR
8C	P41	-	NR
9C	-	-	NR
2A	-	P8	H9 (1:12800)
3A	-	-	NR
5A	-	-	NR
7A	-	-	NR
8A	P10	P11	H12 (1:12800)
9A	P10	-	NR
10A	-	P2	H30 (1:12800)
13A	P13	P13	H16 (1:12800)
14A	-	-	NR
15A	-	-	NR

- : Resultado negativo para PCR-RFLP (não houve similaridade com o antígeno padrão).

NR: Não houve reação do anti-soro das amostras HNT com nenhum antígeno padrão.

Posteriormente, estes anti-soros foram titulados frente aos vinte antígenos de *E. coli* HNT estudados. O anti-soro da amostra 1C reagiu com o antígeno 5C (Título 1:6400) e vice-versa. O anti-soro da amostra 2C reagiu com os antígenos 4C e 11A (Título 1:12800). Este resultado era esperado, uma vez que os antígenos 2C, 4C e 11A apresentaram o mesmo perfil de restrição. O anti-soro da amostra 8C reagiu com o antígeno 9C (Título 1:12800) e vice-versa. O anti-soro da amostra 9A reagiu com o antígeno da amostra 15A (Título 1:12800) e o anti-soro da amostra 10A reagiu com o antígeno da amostra 14A (Título 1:12800). Os resultados são apresentados na tabela 12.

Todos os anti-soros que apresentaram reações cruzadas foram absorvidos. Nos testes de titulação, realizados após o processo de absorção, os anti-soros já não apresentavam reações cruzadas, porém o título do anti-soro contra seu homólogo ficou abaixo de 1:100.

Tabela 12: Resultados da titulação dos anti-soros de *E. coli* HNT frente aos 20 antígenos das amostras de *E. coli* HNT:

Nº	Título frente antígeno HNT	Nº	Título frente antígeno HNT
1C	5C: >6400	2A	-
2C	4C/11A: >12800	3A	-
3C	-	5A	-
5C	1C: >12800	7A	-
7C	-	8A	-
8C	9C: >12800	9A	15A: >12800
9C	8C: >12800	10A	14A: >12.800
		13A	-
		14A	-
		15A	-

-:Reação negativa

4.5 - Sequenciamento:

Para a leitura das reações de sequenciamento foi utilizado o sequenciador MegaBace 1000 (GE Health Care) do Laboratório de Genômica e Biologia Molecular Bacteriana, sob supervisão do Prof. Dr. Marcelo Brocchi do Departamento de Genética, Evolução e Bioagentes, IB/Unicamp.

Todas as amostras de *E. coli* HNT que não apresentaram perfil de restrição semelhante aos das amostras padrão e as amostras HNT que apresentaram perfil de restrição semelhante àqueles das amostras padrão, mas foram negativas nos ensaios sorológicos, foram selecionadas para o sequenciamento de DNA.

Quando utilizamos a técnica de sequenciamento “primer walking” ou “cromossomo walking”, não foram obtidos resultados aproveitáveis na leitura dos fragmentos sequenciados. Em algumas amostras, o sequenciamento dos fragmentos de DNA não iniciava-se e, em outras reações, o número de bases nitrogenadas sequenciadas foi muito baixo. Das amostras estudadas, apenas as amostras 3A e 1C proporcionaram uma boa

leitura dos primeiros 500pb iniciais do gene *fliC* através desta técnica, porém, foram necessárias aproximadamente cinco repetições de cada amostra para a obtenção de resultados de boa qualidade.

Paralelamente, clonamos o gene *fliC* no vetor de clonagem pGEM-Teasy (Promega). Primeiramente, tentamos sequenciar diretamente os plasmídios utilizando tanto os iniciadores propostos por Fields *et al.* (1997), quanto os iniciadores T7 e M13, porém, não obtivemos uma boa qualidade nos resultados do sequenciamento. Em seguida, fizemos a PCR utilizando os plasmídios das *E. coli* HNT como DNA molde e os iniciadores T7 e M13. Quando este produto da PCR foi utilizado na reação de sequenciamento, obtivemos qualidade acima de 97% no sequenciamento.

Para a montagem dos *contigs* foi utilizado o programa Phred/ Phrap/Consed, sob orientação de Marcelo Falsarella Carazzolle, do Laboratório de Genômica e Expressão (DGE-IB-Unicamp).

As sequências gênicas obtidas na reação de sequenciamento foram comparadas aos diferentes genes de *fliC* armazenados no banco de dados do NCBI através do sistema BLAST. O programa ClustalW foi utilizado para alinhamento das sequências. O software BLASTx é uma excelente ferramenta para tentar a identificação de proteínas homólogas em uma região codante de nucleotídeos, além de combinar todos os seis *frames* de leitura. A sequência completa do gene *fliC* das amostras de *E. coli* HNT estudadas e os resultados obtidos pelo sistema BLASTx estão descritos abaixo.

4.5.1: Amostra 1C (VTH-15)

O sequenciamento da amostra 1C (VTH-15) resultou em uma sequência de 1541 pb. Os dois fragmentos em **negrito** demonstram o local de alinhamentos dos iniciadores FliC(F).

GCGAATCACTAGTGATTCT**AACCCTGCAGCAGAGACAGTACGTTCTGCGTGG**
TCTGGTTAGCCTGTGCCAGAACAGAGGTACCCGCTTGTTCAGGATCTGCGCAC
GAGACATGTTAGACACTTCGATCGCGTAGTCAGCATCTTCGATACGGCTACGG
GCAGAAGACAGGTTGTTTACGGTGTTCAGGTTGGTGTGATGGCAGAGTCGAA
ACGGTTTTGTACTGCACCGAGGTCAGAACGCAGATTGTCAACTTTAGCCAATGC
TTTGTGATAGTTTCGAGCGGGTTGGTGGTAGATTGCAACGATTTTGCTGCATC
TTCGTTTACCAGAATCGGGCTACCACCTTCTGATTTGCTCAGATACATGACTTT
ATTGTTACCAGAAGCAGTCTCCGTTATCGTTTTACCATCTGCACTAACATCGTA
AGTTGCACCGTTAACAACACTAACGTGCTTCCTGTTTTCTTGGCAGCGTTCAGATC
AAGATCAGATAGTGTGCTGCTTTATTTTCAACTTTTGTTCAGTCAATTGCCCT
GCACTATTTTTGTATAGGGCTGTCGTAGGGGCAACGGTTGCACTAGTACCTGAT
GTATCAGTACTCCCCGAAATAGTGAATTCAACAGCCTTACCATCAACATTGGCG
GTTAATTTACCATTACCTTTGGCATCTATAGCGACAGTGCCGGTATTTGTAAT
TCGATACCTTCGTAGACAATCTTATCCCTTGAGCTAAATCTTTAGCAGCCACT
GCAAGCTTAGTTGCATCAATCTTGAATTGAGTATCACTGCTGGTCGTAAGTGAA
CCATCCGCAGTACTCACATAAACTTGTTTGGCATCTTTATCCTGTACTACGCCAC
TATCTACATTAACAGTATAGTTATCAGTACCGTTAATTTGATAATTATCAGTAC
CTGTGCTTTAAATTTAGAAATCAGGTCAGTCCCGGTTGCTTTCTGCGCGCCAT
CGATATTAACCGTCCAGGCGGAGAGTTTTTCGCATCAATTTTTGCCAGATTGA
TAGTGATGGTTTACCATCATTAGCACCAACCTGAATTTTCATTTATTATTTTC
AGCAAGGACTTTCACGCCGTTAAACTGAGTTTGCTCAGATACACGGTCAATTTT
TTCCAGACGTTGAGTAATTTACGCCTGGATAGAAGAAAGATCGCTGTCAGAGT
TAGTACCGTTAGTTGCCTGAACAGAAAGTTACGAATACGCTGCAGGTTGTTGT
TAATTTCAATTCAGCGCACCTTCAGTGGTCTGCGCAACAGAAATACCATCATTTCG
CGTTACGGGAAGCCTGAGTCAGACCTTTAATATTTGCTGTAAAACGGTTAGCAA
TCGCCTGACCTGCTGCGTCATCTTTAGCACTGTTAATACGCAGGCCAGAAGAGA
GACGTTCAATGGCGGAGCTCAGAGAAGACTGAGATTTGTTTCAGGTTATTCTGG
GATCAACAGCGACAGGCTGTTGGGTATTAATGACTTGTGCCATAATCGATTC
CCGCGCCGACACGTGTGTGTGCCGTCGGCCCGCG

Através do BLASTn, foi observado que a sequência nucleotídica da amostra 1C apresentava 99% de similaridade com o antígeno H21, logo, foram escolhidas quatro sequências nucleotídicas do gene *fliC* do antígeno H21 depositadas com os nº de acessos: DQ459008, AY337484, AY250004, AF517663, para serem analisadas com a sequência nucleotídica obtida da amostra 1C através do programa ClustalW. Segue abaixo o resultado

da análise do ClustalW, sendo que as sequências de 1-4 foram numeradas de acordo com as sequências dos números de acessos descritos acima. A sequência de nº 5 é da amostra 1C.

```

seq1_ -----ATTA 4
seq4_ -----
seq3_ -----
seq5_ CGGCGGCCGACGGCACACACCGTGTGGCGCGGAATCGATTATGGCACAAAGTCATTAA 60
seq2_ -----ATGGCACAAAGTCATTAA 17

seq1_ ATACAAAACAGCCTGTCGCTGTTGA-CCCAGAATAACCTGAACAAATCTCAGTCTTCTCTG 63
seq4_ -----TCTCAGTCTTCTCTG 15
seq3_ -----AACAAATCTCAGTCTTCTCTG 21
seq5_ TACCCAACAGCCTGTCGCTGTTGATCCCAGAATAACCTGAACAAATCTCAGTCTTCTCTG 80
seq2_ TA-CCAACAGCCTGTCGCTGTTGA-CCCAGAATAACCTGAACAAATCTCAGTCTTCTCTT 75
*****

seq1_ AGCTCCGCCATTGAACGTCTCTTCTGGCCTGCGTATTAACAGTGCTAAAGATGACGCA 123
seq4_ AGCTCCGCCATTGAACGTCTCTTCTGGCCTGCGTATTAACAGTGCTAAAGATGACGCA 75
seq3_ AGCTCCGCCATTGAACGTCTCTTCTGGCCTGCGTATTAACAGTGCTAAAGATGACGCA 81
seq5_ AGCTCCGCCATTGAACGTCTCTTCTGGCCTGCGTATTAACAGTGCTAAAGATGACGCA 180
seq2_ AGCTCTGCTATTGAGCGTCTGTCTTCTGGTCTGCGTATTAACAGCGCAAAGACGATGCA 135
***** ** ***** ***** ***** ***** ***** ***** ***** ** ***** ** **

seq1_ GCAGGTCAGGCGATTGCTAACCGTTTTACAGCAAATATTAAGGTCTGACTCAGGCTTCC 183
seq4_ GCAGGTCAGGCGATTGCTAACCGTTTTACAGCAAATATTAAGGTCTGACTCAGGCTTCC 135
seq3_ GCAGGTCAGGCGATTGCTAACCGTTTTACAGCAAATATTAAGGTCTGACTCAGGCTTCC 141
seq5_ GCAGGTCAGGCGATTGCTAACCGTTTTACAGCAAATATTAAGGTCTGACTCAGGCTTCC 240
seq2_ GCAGGTCAGGCGATTGCTAACCGTTTTACAGCAAATATTAAGG-----CAGGCTTCC 188
*****

seq1_ CGTAACGCGAATGATGGTATTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATT 243
seq4_ CGTAACGCGAATGATGGTATTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATT 195
seq3_ CGTAACGCGAATGATGGTATTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATT 201
seq5_ CGTAACGCGAATGATGGTATTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATT 300
seq2_ CGTAACGCGAATGATGGTATTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATT 248
*****

seq1_ AACACAACCTGCAGCGTATTCTGTAACCTTCTGTTTCAGGCAACTAACGGTACTAECTCT 303
seq4_ AACACAACCTGCAGCGTATTCTGTAACCTTCTGTTTCAGGCAACTAACGGTACTAECTCT 255
seq3_ AACACAACCTGCAGCGTATTCTGTAACCTTCTGTTTCAGGCAACTAACGGTACTAECTCT 261
seq5_ AACACAACCTGCAGCGTATTCTGTAACCTTCTGTTTCAGGCAACTAACGGTACTAECTCT 360
seq2_ AACACAACCTGCAGCGTATTCTGTAACCTTCTGTTTCAGGCAACTAACGGTACTAECTCT 308
*****

seq1_ GACAGCGATCTTTCTTCTATCCAGGCTGAAATTAECTCAACGTCTGGAAGAAATTGACCGT 363
seq4_ GACAGCGATCTTTCTTCTATCCAGGCTGAAATTAECTCAACGTCTGGAAGAAATTGACCGT 315
seq3_ GACAGCGATCTTTCTTCTATCCAGGCTGAAATTAECTCAACGTCTGGAAGAAATTGACCGT 321
seq5_ GACAGCGATCTTTCTTCTATCCAGGCTGAAATTAECTCAACGTCTGGAAGAAATTGACCGT 420
seq2_ GACAGCGATCTTTCTTCTATCCAGGCTGAAATTAECTCAACGTCTGGAAGAAATTGACCGT 368
*****

seq1_ GTATCTGAGCAAACCTCAGTTTAAACGGCGTGAAAGTCCTTGCTGAAAATAATGAAATGAAA 423
seq4_ GTATCTGAGCAAACCTCAGTTTAAACGGCGTGAAAGTCCTTGCTGAAAATAATGAAATGAAA 375
seq3_ GTATCTGAGCAAACCTCAGTTTAAACGGCGTGAAAGTCCTTGCTGAAAATAATGAAATGAAA 381
seq5_ GTATCTGAGCAAACCTCAGTTTAAACGGCGTGAAAGTCCTTGCTGAAAATAATGAAATGAAA 480
seq2_ GTATCTGAGCAAACCTCAGTTTAAACGGCGTGAAAGTCCTTGCTGAAAATAATGAAATGAAA 428
*****

```

```

seq1_      ATTCAGGTTGGTGCTAATGATGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCG 483
seq4_      ATTCAGGTTGGTGCTAATGATGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCG 435
seq3_      ATTCAGGTTGGTGCTAATGATGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCG 441
seq5_      ATTCAGGTTGGTGCTAATGATGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCG 540
seq2_      ATTCAGGTTGGTGCTAATGATGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCG 488
*****

seq1_      AAAACTCTCGGCCCTGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGAC 543
seq4_      AAAACTCTCGGCCCTGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGAC 495
seq3_      AAAACTCTCGGCCCTGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGAC 501
seq5_      AAAACTCTCGGCCCTGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGAC 600
seq2_      AAAACTCTCGGCCCTGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGAC 548
*****

seq1_      CTGATTTCTAAATTTAAAGCGCAGGTACTGATAATTATCAAATTAACGGTACTGATAAC 603
seq4_      CTGATTTCTAAATTTAAAGCGCAGGTACTGATAATTATCAAATTAACGGTACTGATAAC 555
seq3_      CTGATTTCTAAATTTAAAGCGCAGGTACTGATAATTATCAAATTAACGGTACTGATAAC 561
seq5_      CTGATTTCTAAATTTAAAGCGCAGGTACTGATAATTATCAAATTAACGGTACTGATAAC 660
seq2_      CTGATTTCTAAATTTAAAGCGCAGGTACTGATAATTATCAAATTAACGGTACTGATAAC 608
*****

seq1_      TATACTGTTAATGTAGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTG 663
seq4_      TATACTGTTAATGTAGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTG 615
seq3_      TATACTGTTAATGTAGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTG 621
seq5_      TATACTGTTAATGTAGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTG 720
seq2_      TATACTGTTAATGTAGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTG 668
*****

seq1_      AGTACTGCGGATGGTTCAC TTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACT 723
seq4_      AGTACTGCGGATGGTTCAC TTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACT 675
seq3_      AGTACTGCGGATGGTTCAC TTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACT 681
seq5_      AGTACTGCGGATGGTTCAC TTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACT 780
seq2_      AGTACTGCGGATGGTTCAC TTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACT 728
***

seq1_      AAGCTTGCAGTGGCTGCTAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATC 783
seq4_      AAGCTTGCAGTGGCTGCTAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATC 735
seq3_      AAGCTTGCAGTGGCTGCTAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATC 741
seq5_      AAGCTTGCAGTGGCTGCTAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATC 840
seq2_      AAGCTTGCAGTGGCTGCTAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATC 788
*****

seq1_      GAATTTACAAATACCGGCAC TGTGCTATAGATGCCAAAGGTAATGGTAAATTAACCGCC 843
seq4_      GAATTTACAAATACCGGCAC TGTGCTATAGATGCCAAAGGTAATGGTAAATTAACCGCC 795
seq3_      GAATTTACAAATACCGGCAC TGTGCTATAGATGCCAAAGGTAATGGTAAATTAACCGCC 801
seq5_      GAATTTACAAATACCGGCAC TGTGCTATAGATGCCAAAGGTAATGGTAAATTAACCGCC 900
seq2_      GAATTTACAAAACCGGCAC TGGCGCTATAACTGCCACAGGTAATGGTAAATTAACCGCC 848
*****

seq1_      AATGTTGATGGTAAGGCTGTTGAATTCACTATTTTCGGGGAGTACTGATACATCAGGTACT 903
seq4_      AATGTTGATGGTAAGGCTGTTGAATTCACTATTTTCGGGGAGTACTGATACATCAGGTACT 855
seq3_      AATGTTGATGGTAAGGCTGTTGAATTCACTATTTTCGGGGAGTACTGATACATCAGGTACT 861
seq5_      AATGTTGATGGTAAGGCTGTTGAATTCACTATTTTCGGGGAGTACTGATACATCAGGTACT 960
seq2_      AATGTTGATGGTAAGGCTGTTGAATTCACTATTTTCGGGGAGTACTGATACATCAGGTACT 908
*****

seq1_      AGTGCAACCGTTGCCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCA 963
seq4_      AGTGCAACCGTTGCCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCA 915
seq3_      AGTGCAACCGTTGCCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCA 921
seq5_      AGTGCAACCGTTGCCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCA 1020
seq2_      AGTGCAACCGTTGCCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCA 968
*****

```

```

seq1_      ACAAAGTTGAAAATAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCCAAGAAA 1023
seq4_      ACAAAGTTGAAAATAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCCAAGAAA 975
seq3_      ACAAAGTTGAAAATAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCCAAGAAA 981
seq5_      ACAAAGTTGAAAATAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCCAAGAAA 1080
seq2_      ACAAAGTTGAAAATAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCCAAGAAA 1028
*****

seq1_      ACAGGAAGCACGTTAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACG 1083
seq4_      ACAGGAAGCACGTTAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACG 1035
seq3_      ACAGGAAGCACGTTAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACG 1041
seq5_      ACAGGAAGCACGTTAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACG 1140
seq2_      ACAGGAAGCACGTTAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACG 1088
*****

seq1_      ATAACGGAGACTGCTTCTGGTAAACAATAAAGTCATGTATCTGAGCAAATCAGAAGGTGGT 1143
seq4_      ATAACGGAGACTGCTTCTGGTAAACAATAAAGTCATGTATTTGAGCAAATCAGAAGGTGGT 1095
seq3_      ATAACGGAGACTGCTTCTGGTAAACAATAAAGTCATGTATCTGAGCAAATCAGAAGGTGGT 1101
seq5_      ATAACGGAGACTGCTTCTGGTAAACAATAAAGTCATGTATCTGAGCAAATCAGAAGGTGGT 1200
seq2_      ATAACGGAGACTGCTTCTGGTAAACAATAAAGTCATGTATCTGAGCAAATCAGAAGGTGGT 1148
*****

seq1_      AGCCCGATTCTGGTAAACGAAGATGCAGCAAATCGTTGCAATCTACCACCAACCCGCTC 1203
seq4_      AGCCCGATTTTGGTAAACGAAGATGCAGCAAATCGTTGCAATTTACCACCAACCCGCTC 1155
seq3_      AGCCCGATTCTGGTAAACGAAGATGCAGCAAATCGTTGCAATCTACCACCAACCCGCTC 1161
seq5_      AGCCCGATTCTGGTAAACGAAGATGCAGCAAATCGTTGCAATCTACCACCAACCCGCTC 1260
seq2_      AGCCCGATTCTGGTAAACGAAGATGCAGCAAATCGTTGCAATCTACCACCAACCCGCTC 1208
*****

seq1_      GAACTATCGACAAAGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTA 1263
seq4_      GAACTATCGACAAAGCATTGGCTAAAGTTGACAATTTGCGTTCTGACCTCGGTGCAGTA 1215
seq3_      GAACTATCGACAAAGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTA 1221
seq5_      GAACTATCGACAAAGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTA 1320
seq2_      GAACTATCGACAAAGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTA 1268
*****

seq1_      CAAAACCGTTTCGACTCTGCCATCACCAACCTTGGCAACACCGTAAACAACCTGTCTTCT 1323
seq4_      CAAAACCGTTTCGACTTTGCCATCACCAACCTTGGCAACACCGTAAACAACCTGTCTTCT 1275
seq3_      CAAAACCGTTTCGACTCTGCCATCACCAACCTTGGCAACACCGTAAACAACCTGTCTTCT 1281
seq5_      CAAAACCGTTTCGACTCTGCCATCACCAACCTTGGCAACACCGTAAACAACCTGTCTTCT 1380
seq2_      CAAAACCGTTTCGACTCTGCCATCACCAACCTTGGCAACACCGTAAACAACCTGTCTTCT 1328
*****

seq1_      GCCCGTAGCCGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCG 1383
seq4_      GCCCGTAGCCGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTTTCGTGCG 1335
seq3_      GCCCGTAGCCGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCG 1341
seq5_      GCCCGTAGCCGTATCGAAGATGCTGACTACGCGAATCGAAGTGTCTAACATGTCTCGTGCG 1440
seq2_      GCCCGTAGCCGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCG 1388
*****

seq1_      CAGATCCTGCAACAAGCGGGTACCTCTGTTCTGGCACAGGCTAACCAGACCACGCAGAAC 1443
seq4_      CAGATCCTGCAACAAGCGGGT----- 1356
seq3_      CAGATCCTGCAACAAGCGGGTACCTCTGTTCTGGCACAG----- 1380
seq5_      CAGATCCTGCAACAAGCGGGTACCTCTGTTCTGGCACAGGCTAACCAGACCACGCAGAAC 1500
seq2_      CAGATCCTGCAACAAGCGGGTACCTCTGTTCTGGCACAGGCTAACCAGACCACGCAGAAC 1448
*****

seq1_      GTACTGTCT----- 1452
seq4_      -----
seq3_      -----
seq5_      GTACTGTCTCTGCTGCAGGGTTAGAATCACTAGTGATTTCGC 1541
seq2_      GTACTGTCTCTGCTGCAGGGTTAA----- 1472

```

Quando comparamos a sequência nucleotídica da amostra 1C frente as demais sequências disponíveis, o programa de alinhamento ClustalW demonstrou que houve uma mudança de uma base de citosina por uma timina na região de 1415pb.

Quando analisada no BLASTx, a provável sequência de aminoácidos apresentou similaridade de 99% com a sequência do antígeno H21 (nº de acessos: ABI239661; ABE69176; AAP133121). As sequências de aminoácidos variaram em tamanho de 460 aa a 491 aa. Todas as sequências apresentaram uma mutação não sinônima, onde foi trocado o aminoácido I (Isoleucina) por T (Treonina) no aminoácido 457. A amostra 1C não apresentou o mesmo perfil de restrição que o antígeno H21 (pag.60), uma vez que, este antígeno padrão não possuiu sítios de restrição para a enzima *RsaI* e gerou fragmentos de diferentes tamanhos quando utilizamos a enzima *HhaI*.

gb|ABI23966.1| FliC [*Escherichia coli*]

Length=491

Score = 836 bits (2160), Expect = 0.0

Identities = 456/458 (99%), Positives = 457/458 (99%), Gaps = 0/458 (0%)

Frame = -1

Query	1397	SGLRINSAKDDAAGQAIANRFTANIKGLTQASRNANDGISVAQTTEGALNEINNNLQRIR	1218
		SGLRINSAKDDAAGQAIANRFTANIKGLTQASRNANDGISVAQTTEGALNEINNNLQRIR	
Sbjct	34	SGLRINSAKDDAAGQAIANRFTANIKGLTQASRNANDGISVAQTTEGALNEINNNLQRIR	93
Query	1217	ELSVQATNGTNSDSLSSIQAEITQRLEEIDRVSEQTQFNGVKVLAENNEMKIQVGANDG	1038
		ELSVQATNGTNSDSLSSIQAEITQRLEEIDRVSEQTQFNGVKVLAENNEMKIQVGANDG	
Sbjct	94	ELSVQATNGTNSDSLSSIQAEITQRLEEIDRVSEQTQFNGVKVLAENNEMKIQVGANDG	153
Query	1037	ETITINLAKIDAKTLGLDGFNIDGAQKATGSDLISKFKATGTDNYQINGTDNYTVNVDSG	858
		ETITINLAKIDAKTLGLDGFNIDGAQKATGSDLISKFKATGTDNYQINGTDNYTVNVDSG	
Sbjct	154	ETITINLAKIDAKTLGLDGFNIDGAQKATGSDLISKFKATGTDNYQINGTDNYTVNVDSG	213
Query	857	VVQDKDGKQVYVSTADGSLTTSSDTQFKIDATKLAVAADLAQGNKIVYEGIEFTNTGTV	678
		VVQDKDGKQVYVSTADGSLTTSSDTQFKIDATKLAVAADLAQGNKIVYEGIEFTNTGTV	
Sbjct	214	VVQDKDGKQVYVSTADGSLTTSSDTQFKIDATKLAVAADLAQGNKIVYEGIEFTNTGTV	273
Query	677	AIDAKNGKLTANVDGKAVEFTISGSTDTSGTSATVAPPTALYKNSAGQLTATKVENKAA	498
		AIDAKNGKLTANVDGKAVEFTISGSTDTSGTSATVAPPTALYKNSAGQLTATKVENKAA	
Sbjct	274	AIDAKNGKLTANVDGKAVEFTISGSTDTSGTSATVAPPTALYKNSAGQLTATKVENKAA	333

Query	497	TLSDLDLNAAKKTGSTLVVNGATYDVSADGKTI TETASGNNKVMYLSKSEGGSPILVNE	318
Sbjct	334	TLSDLDLNAAKKTGSTLVVNGATYDVSADGKTI TETASGNNKVMYLSKSEGGSPILVNE	393
Query	317	AAKSLQSTTNPLETIDKALAKVDNLRSDLGAVQNRFD SAI TN LGNTVNNLSSARSRIEDA	138
Sbjct	394	AAKSLQSTTNPLETIDKALAKVDNLRSDLGAVQNRFD SAI TN LGNTVNNLSSARSRIEDA	453
Query	137	DYATIEVSNMSRAQILQQAGTSVLAQANQTTQNVLSLLQ	24
Sbjct	454	DYAEVSNMSRAQILQQAGTSVLAQANQTTQNVLSLLR	491

4.5.2: Amostra 5C (33141a)

O sequenciamento da amostra 5C (33141a) obteve um fragmento sequenciado de 1522 pb. Os dois fragmentos de DNA abaixo em negrito demonstram o local de alinhamentos dos iniciadores FliC(F).

TTAGACGACATGGCGTAGCGGGAATCGATTAT**TGGCACACGTCATTAATACCC**
AACAGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCTTCTCTG
AGCTCCGCCATTGAACGTCTCTCTTCTGGCCCGCGTATTAACAGTGCTAAAGAT
GACGCAGCAGGTCAGGCGACTGCTAACCGTTTTACAGCAAATATTAAGGTCT
GACTCAGGCTTCCCGTAACGCGAATGATGGTATTTCTGTTGCGCAGACCACTGA
AGGTGCGCTGAATGAAATTAACAACAACCTGCAGCGTATTCGTGAACCTTCTGT
TCAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTCTTCTATCCAGGCTGA
AATTACTCAACGTCTGGAAGAAATTGACCGTGTATCTGAGCAAACCTCAGTTTAA
CGGCGTGAACCTTGCTGAAAATAATGAAATGAAAATTCAGGTTGGTGCTAATGA
TGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCC
TGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATT
TCTAAATTTAAAGCGACAGGTAAGTATCAAAATTAACGGTACTGATAAC
TATACTGTAAATGTAGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGT
TTATGTGAGTACTGCGGATGGTTCACTTACGACCAGCAGTGATACTCAATTCAA
GATTGATGCAACTAAGCTTGCAGTGGCTGCTAAAGATTTAGCTCAAGGGAATA
AGATTGTCTACGAAGGTATCGAATTTACAAATAGCGGCACTGTCGCTATAGAT
GCCAAAGGTAATGGTAAACTAACCGCCAATGTTGATGGTAAGGCTGTTGAATT
CACTATCTCGGGGAGTACTGATACATCAGGTAAGTACAGGTAAGTACAGGTAAGT
GACAGCCCTATACAAAATAGTGCAGGGCAATTGACTGCCACAAAGTTGAAAA
TAAAGCAGCGACACTATCTGATGTTGATCTGAACGCTGCCCAAGAAAACAGGA
AGCAGGTTAGTTGTTAACGGTGCACCTTACGATGTTAGTGCAGATGGTAAAC
GATAACGGGAGACTGCTTCTGGTAACAATAAAGTCATGTATCTGAGCAAATCAG
AAGGTGGTAGCCCGATTCTGGTAAACGAAGATGCAGCAAATCGTTGCAATCT
ACCACCAACCCGCTCGAAACTATCGACAAAGCATTGGCTAAAGTTGACAATCT
GCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCGACTCTGCCATCACCAACCT

TGGCAACACCGTAAACAACCTGTCTTCTGCCCGTAGCCGTATCGAAGATGCTGA
 CTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAACAATCGG
 GTACCTCTGTTCTGGCACAGGCTAACCAGACCACGCAGAACGTACTGTCTCTG
 CTGCAGGGTTAGAATCACTAGTGAATTC

Através do BLASTn, foi observado que a sequência nucleotídica da amostra 5C apresentava 98% de similaridade com o antígeno H21, logo, foram escolhidas quatro sequências do gene *fliC* do antígeno H21 depositadas com os nº de acessos: DQ459008, AY337484, AY250004, AF517663, para serem analisadas com a sequência nucleotídica obtida da amostra 5C através do programa de alinhamento ClustalW. Segue abaixo o resultado da análise do ClustalW, sendo que as sequências foram numeradas de acordo com as sequências dos números de acessos descritos acima. A sequência de nº 5 é a amostra 5C.

```

seq3_ -----
seq4_ -----
seq1_ -----ATTAATACAAACAGCCT 17
seq5_ TTAGACGACATGGCGTAGCGGAATCGATTATGGCACACGTCATTAATACCCAACAGCCT 60
seq2_ -----ATGGCACAAAGTCATTAATA-CCAACAGCCT 29

seq3_ -----AACAAATCTCAGTCTTCTCTGAGCTCCGCCATTGA 35
seq4_ -----TCTCAGTCTTCTCTGAGCTCCGCCATTGA 29
seq1_ GTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCTTCTCTGAGCTCCGCCATTGA 77
seq5_ GTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCTTCTCTGAGCTCCGCCATTGA 120
seq2_ GTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCTTCTCTTAGCTCTGCTATTGA 89
          ***** **

seq3_ ACGTCTCTCTTCTGGCCTGCGTATTAACAGTGCTAAAGATGACGCAGCAGGTCAGGCGAT 95
seq4_ ACGTCTCTCTTCTGGCCTGCGTATTAACAGTGCTAAAGATGACGCAGCAGGTCAGGCGAT 89
seq1_ ACGTCTCTCTTCTGGCCTGCGTATTAACAGTGCTAAAGATGACGCAGCAGGTCAGGCGAT 137
seq5_ ACGTCTCTCTTCTGGCCCGCGTATTAACAGTGCTAAAGATGACGCAGCAGGTCAGGCGAC 180
seq2_ GCGTCTGTCTTCTGGTCTGCGTATTAACAGCGCAAAGACGATGCAGCAGGTCAGGCGAT 149
          ***** *

seq3_ TGCTAACCGTTTTACAGCAAATATTAAGGTCTGACTCAGGCTTCCCCTAACGCGAATGA 155
seq4_ TGCTAACCGTTTTACAGCAAATATTAAGGTCTGACTCAGGCTTCCCCTAACGCGAATGA 149
seq1_ TGCTAACCGTTTTACAGCAAATATTAAGGTCTGACTCAGGCTTCCCCTAACGCGAATGA 197
seq5_ TGCTAACCGTTTTACAGCAAATATTAAGGTCTGACTCAGGCTTCCCCTAACGCGAATGA 240
seq2_ TGCTAACCGTTTTACGCAAATATTAAGG-----CAGGCTTCCCCTAACGCGAATGA 202
          *****

seq3_ TGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCA 215
seq4_ TGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCA 209
seq1_ TGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCA 257
seq5_ TGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCA 300
seq2_ TGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCA 262
          *****

```

```

seq3_      GCGTATTCGTGAACTTTCTGTTCAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTC 275
seq4_      GCGTATTCGTGAACTTTCTGTTCAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTC 269
seq1_      GCGTATTCGTGAACTTTCTGTTCAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTC 317
seq5_      GCGTATTCGTGAACTTTCTGTTCAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTC 360
seq2_      GCGTATTCGTGAACTTTCTGTTCAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTC 322
*****

seq3_      TTCTATCCAGGCTGAAATTACTCAACGCTCTGGAAGAAATTGACCGTGTATCTGAGCAAAC 335
seq4_      TTCTATCCAGGCTGAAATTACTCAACGCTCTGGAAGAAATTGACCGTGTATCTGAGCAAAC 329
seq1_      TTCTATCCAGGCTGAAATTACTCAACGCTCTGGAAGAAATTGACCGTGTATCTGAGCAAAC 377
seq5_      TTCTATCCAGGCTGAAATTACTCAACGCTCTGGAAGAAATTGACCGTGTATCTGAGCAAAC 420
seq2_      TTCTATCCAGGCTGAAATTACTCAACGCTCTGGAAGAAATTGACCGTGTATCTGAGCAAAC 382
*****

seq3_      TCAGTTTAAACGGCGTGAAAGTCTTGTCTGAAAATAATGAAATGAAATTCAGGTTGGTGC 395
seq4_      TCAGTTTAAACGGCGTGAAAGTCTTGTCTGAAAATAATGAAATGAAATTCAGGTTGGTGC 389
seq1_      TCAGTTTAAACGGCGTGAAAGTCTTGTCTGAAAATAATGAAATGAAATTCAGGTTGGTGC 437
seq5_      TCAGTTTAAACGGCGTGAAAGTCTTGTCTGAAAATAATGAAATGAAATTCAGGTTGGTGC 476
seq2_      TCAGTTTAAACGGCGTGAAAGTCTTGTCTGAAAATAATGAAATGAAATTCAGGTTGGTGC 442
*****

seq3_      TAATGATGGTGA AACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCCT 455
seq4_      TAATGATGGTGA AACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCCT 449
seq1_      TAATGATGGTGA AACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCCT 497
seq5_      TAATGATGGTGA AACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCCT 536
seq2_      TAATGATGGTGA AACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCCT 502
*****

seq3_      GGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATT 515
seq4_      GGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATT 509
seq1_      GGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATT 557
seq5_      GGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATT 596
seq2_      GGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATT 562
*****

seq3_      TAAAGCGACAGGTACTGATAATTATCAAATTAACGGTACTGATAACTATACTGTTAATGT 575
seq4_      TAAAGCGACAGGTACTGATAATTATCAAATTAACGGTACTGATAACTATACTGTTAATGT 569
seq1_      TAAAGCGACAGGTACTGATAATTATCAAATTAACGGTACTGATAACTATACTGTTAATGT 617
seq5_      TAAAGCGACAGGTACTGATAATTATCAAATTAACGGTACTGATAACTATACTGTTAATGT 656
seq2_      TAAAGCGACAGGTACTGATAATTATCAAATTAACGGTACTGATAACTATACTGTTAATGT 622
*****

seq3_      AGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTGAGTACTGCGGATGG 635
seq4_      AGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTGAGTACTGCGGATGG 629
seq1_      AGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTGAGTACTGCGGATGG 677
seq5_      AGATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTGAGTACTGCGGATGG 716
seq2_      AGATAGTGGAGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTGAGTACTGCGGATGG 682
*****

seq3_      TTCACTTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGCAGTGGC 695
seq4_      TTCACTTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGCAGTGGC 689
seq1_      TTCACTTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGCAGTGGC 737
seq5_      TTCACTTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGCAGTGGC 776
seq2_      TTCACTTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGCAGTGGC 742
*****

seq3_      TGCTAAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATCGAATTTACAAATAC 755
seq4_      TGCTAAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATCGAATTTACAAATAC 749
seq1_      TGCTAAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATCGAATTTACAAATAC 797
seq5_      TGCTAAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATCGAATTTACAAATAG 836
seq2_      TGCTAAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATCGAATTTACAAAAC 802
*****

```

```

seq3_      CGGCACTGTCGCTATAGATGCCAAAGGTAATGGTAAATTAACCGCCAATGTTGATGGTAA 815
seq4_      CGGCACTGTCGCTATAGATGCCAAAGGTAATGGTAAATTAACCGCCAATGTTGATGGTAA 809
seq1_      CGGCACTGTCGCTATAGATGCCAAAGGTAATGGTAAATTAACCGCCAATGTTGATGGTAA 857
seq5_      CGGCACTGTCGCTATAGATGCCAAAGGTAATGGTAAATTAACCGCCAATGTTGATGGTAA 896
seq2_      CGGCACTGGCGCTATAACTGCCACAGGTAATGGTAAATTAACCGCCAATGTTGATGGTAA 862
          *****

seq3_      GGCTGTTGAATTCACTATTTCCGGGGAGTACTGATACATCAGGTACTAGTGCAACCGTTGC 875
seq4_      GGCTGTTGAATTCACTATTTCCGGGGAGTACTGATACATCAGATACTAGTGCAACCGTTGC 869
seq1_      GGCTGTTGAATTCACTATTTCCGGGGAGTACTGATACATCAGGTACTAGTGCAACCGTTGC 917
seq5_      GGCTGTTGAATTCACTATTTCCGGGGAGTACTGATACATCAGGTACTAGTGCAACCGTTGC 956
seq2_      GGCTGTTGAATTCACTATTTCCGGGGAGTACTGATACATCAGGTACTAGTGCAACCGTTGC 922
          *****

seq3_      CCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCAACAAAAGTTGAAAA 935
seq4_      CCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCAACAAAAGTTGAAAA 929
seq1_      CCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCAACAAAAGTTGAAAA 977
seq5_      CCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCAACAAAAGTTGAAAA 1015
seq2_      CCCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCAACAAAAGTTGAAAA 982
          *****

seq3_      TAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCC-AAGAAAACAGGAAGCACGT 994
seq4_      TAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCC-AAGAAAACAGGAAGCACGT 988
seq1_      TAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCC-AAGAAAACAGGAAGCACGT 1036
seq5_      TAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCC-AAGAAAACAGGAAGCACGT 1075
seq2_      TAAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCC-AAGAAAACAGGAAGCACGT 1041
          *****

seq3_      TAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACGATAACGGGAGACTG 1054
seq4_      TAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACGATAACGGGAGACTG 1048
seq1_      TAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACGATAACGGGAGACTG 1096
seq5_      TAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACGATAACGGGAGACTG 1135
seq2_      TAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACGATAACGGGAGACTG 1101
          *****

seq3_      CTTCTGGTAACAATAAAGTCAATGATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGG 1114
seq4_      CTTCTGGTAACAATAAAGTCAATGATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGG 1108
seq1_      CTTCTGGTAACAATAAAGTCAATGATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGG 1156
seq5_      CTTCTGGTAACAATAAAGTCAATGATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGG 1195
seq2_      CTTCTGGTAACAATAAAGTCAATGATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGG 1161
          *****

seq3_      TAAACGAAGATGCAGCAAAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACA 1174
seq4_      TAAACGAAGATGCAGCAAAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACA 1168
seq1_      TAAACGAAGATGCAGCAAAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACA 1216
seq5_      TAAACGAAGATGCAGCAAAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACA 1255
seq2_      TAAACGAAGATGCAGCAAAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACA 1221
          *****

seq3_      AAGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCG 1234
seq4_      AAGCATTGGCTAAAGTTGACAATTTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCG 1228
seq1_      AAGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCG 1276
seq5_      AAGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCG 1315
seq2_      AAGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCG 1281
          *****

seq3_      ACTCTGCCATCACCAACCTTGGAACACCCGTAACAACCTGTCTTCTGCCCGTAGCCGTA 1294
seq4_      ACTTTGCCATCACCAACCTTGGAACACCCGTAACAACCTGTCTTCTGCCCGTAGCCGTA 1288
seq1_      ACTCTGCCATCACCAACCTTGGAACACCCGTAACAACCTGTCTTCTGCCCGTAGCCGTA 1336
seq5_      ACTCTGCCATCACCAACCTTGGAACACCCGTAACAACCTGTCTTCTGCCCGTAGCCGTA 1375
seq2_      ACTCTGCCATCACCAACCTTGGAACACCCGTAACAACCTGTCTTCTGCCCGTAGCCGTA 1341
          *****

```

```

seq3_      TCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAAC 1354
seq4_      TCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAAC 1348
seq1_      TCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAAC 1396
seq5_      TCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAAC 1435
seq2_      TCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAAC 1401
*****

seq3_      AAGCGGGTACCTCTGTTCTGGCACAG----- 1380
seq4_      AAGCGGGT----- 1356
seq1_      AAGCGGGTACCTCTGTTCTGGCACAGGCTAACCAGACCACGCAGAACGTACTGTCT---- 1452
seq5_      AATCGGGTACCTCTGTTCTGGCACAGGCTAACCAGACCACGCAGAACGTACTGTCTCTGC 1495
seq2_      AAGCGGGTACCTCTGTTCTGGCACAGGCTAACCAGACCACGCAGAACGTACTGTCTCTGC 1461
** *****

seq3_      -----
seq4_      -----
seq1_      -----
seq5_      TGCAGGGTTAGAATCACTAGTGAATTC 1522
seq2_      TGCAGGGTTAA----- 1472

```

Após a verificação das diferenças encontradas entre a sequência nucleotídica da amostra 5C e das demais amostras padrão H21 através do alinhamento pelo programa ClustalW, os eletroesferogramas da amostra 5C foram analisados, e verificamos que estes apresentaram uma alta qualidade na leitura do sequenciamento. Cada base que demonstrou um resultado diferente ao encontrado no alinhamento foi verificada e constatamos que nestas regiões, estas bases realmente estão presentes, e não são artefatos ou erros de leitura.

Quando analisada no BLASTx, a sequência de aminoácidos obteve 96% de similaridade com o antígeno H21 (nº de acessos: ABI239661; ABE691761; AAP133121; AAN771071), porém, foi encontrado um *stop códon* por volta do aminoácido 340, além de mutações sinônimas (+) e não-sinônimas (espaço em branco) nesta mesma região. Segue abaixo o resultado do BLASTx.

```

gb|ABI23966.1| FliC [Escherichia coli]
Length=491
Score = 794 bits (2051), Expect = 0.0
Identities = 438/457 (95%), Positives = 443/457 (96%), Gaps = 0/457 (0%)
Frame = +2

```

Query	134	GPRINSAKDDAAGQATANRFTANIKGLTQASRNANDGISVAQTTEGALNEINNNLQRIRE	313
Sbjct	35	GRINSAKDDAAGQAANRFTANIKGLTQASRNANDGISVAQTTEGALNEINNNLQRIRE	94
Query	314	LSVQATNGTNSDSLSSIQAEITQRLEEIDRVSEQTQFNGVKVLAENNEMKIQVGANDGE	493
Sbjct	95	LSVQATNGTNSDSLSSIQAEITQRLEEIDRVSEQTQFNGVKVLAENNEMKIQVGANDGE	154
Query	494	TITINLAKIDAKTLGLDGFNIDGAQKATGSDLISKFKATGTDNYQINGTDNYTVNVDSGV	673
Sbjct	155	TITINLAKIDAKTLGLDGFNIDGAQKATGSDLISKFKATGTDNYQINGTDNYTVNVDSGV	214
Query	674	VQDKDGKQVYVSTADGSLTSSDTQFKIDATKLAVAAKDLAQGNKIVYEGIEFTNSGTVA	853
Sbjct	215	VQDKDGKQVYVSTADGSLTSSDTQFKIDATKLAVAAKDLAQGNKIVYEGIEFTNTGTVA	274
Query	854	IDAKGNKLTANVDGKAVEFTISGSTDTSGTSATVAPTTALYKNSAGQLTATKLIKQRH	1033
Sbjct	275	IDAKGNKLTANVDGKAVEFTISGSTDTSGTSATVAPTTALYKNSAGQLTATKVENKAAT	334
Query	1034	YLMLI*TLPKKTGSTLVVNGATYDVSADGKTIETASGNNKVMYLSKSEGGSPILVNEDA	1213
Sbjct	335	LKKTGSTLVVNGATYDVSADGKTIETASGNNKVMYLSKSEGGSPILVNEDA	394
Query	1214	AKSLQSTTNPLETIDKALAKVDNLRSDLGAVQNRFDASAITNLGNTVNNLSSARSRIEDAD	1393
Sbjct	395	AKSLQSTTNPLETIDKALAKVDNLRSDLGAVQNRFDASAITNLGNTVNNLSSARSRIEDAD	454
Query	1394	YATEVSNMSRAQILQQSGTSLVLAQANQTTQNVLSLLO	1504
Sbjct	455	YATEVSNMSRAQILQQ+GTSVLAQANQTTQNVLSLLR	491

4.5.3: Amostra 8C (48629ca)

O sequenciamento da amostra 8C (48629ca) obteve um fragmento sequenciado de 1902 pb. Os dois fragmentos de DNA em negrito abaixo, demonstram o local de alinhamento dos iniciadores FliC(F).

GAATTCACTAGTGATTAT**TGGCACAAGTCATTAATACCCAACAGCCTCTCGCTG**
ATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGA
GCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCAGCGGGTC
AGGCGATTGCTAACCGTTTCACCTCTAACATTAAGGCCTGACTCAGGCTGCAC
GTAACGCTAACGATGGTATCTCTCTGGCGCAGACCACTGAAGGCGCACTGTCT
GAGATTAACAACAACCTTACAACGTGTGCGTGAGTTGACTGTACAGGCGACCAC
CGGTACTAACTCTGATTCTGACCTGGCTTCTATTCAGGACGAAATCAAATCCCCG
TTTGTCTGAAATTGACCGCGTATCCGGGCAGACCCAGTTCAACGGCGTGAACGT
ATTGTCTAAAGATGGCTCCCTGAAAATTCAGGTTGGCGCAAATGATGGTCAGA
CTATCTCTATCGACCTGAAGAAAATTGACTCTGATACTCTGGGTTTGAATGGTT

TCAACGTTAATGGTTCTGGTACCATTGCAAACAAAGCGGCCACAATCAGTGAC
 TTGACTGCTCAGAAAGCCGTTGACAACGGTAATGGTACTTATAAAGTTACAAC
 AGCGACGCTGCACTTACTGCATCTCAGGCATTAAGTAAGCTGAGTGATGGCGA
 TACTGTAGATATTGCAACCTATGCTGGTGGTACAAGTTCAACAGTTAGTTATAA
 ATACGACGCAGATCCAGGTAACCTCAGTTATAACAATACTGCAAACAAAACAA
 GTGCTGCGGCTGGAACCTCTGGCAGATACTCTTCTCCCGGCAGCTGGCCAGACTA
 AAACCGGTACTTACAAGGCTGCTACTGGTGATGTAACTTTAATGTTGACGCAA
 CTGGTAATCTGACAATTGGCGGACAGCAAGCCTACCTGACTACTGATGGTAAC
 CTTACAACAAACAACTCCGGTGGTGC GGCTACTGCAACTCTTAAAGAGCTGTTT
 ACTCTTGCTGGCGATGGTAAATCTCTGGGGAACGGCGGTACTGCTACCGTTACT
 CTGGATAATACTACGTATAATTTCAAAGCTGCTGCGAACGTTACTGATGGTGCT
 GGTGTCATCGCTGCTGCTGGTGTAACCTTATACAGCCACTGTTTCTAAAGATGTC
 ATTCTGGCACAACCTGCAATCTGCAAGTCAGGCAGCAGCAACCGCTACCGACGG
 TGATACTGTCGCAACGATCAACTATAAATCTGGTGTCATGATCGGTTCCGCTAC
 CTTTACCAATGGTAAAGGTAAGTACTGCCGATGGTATGACTTCTGGTACAACCTCCAGT
 CGTAGCTACAGGTGCTAAAGCTGTATATGTTGATGGCAACAATGAACTGACTTC
 CACTGCATCTTACGATACGACTTACTCTGTCAACGCAGATACAGGCGCAGTAA
 AAGTGGTATCAGGTACTGGTACTGGTAAATTTGAAGCTGTTGCTGGTGCGGATG
 CTTATGTAAGCAAAGATGGCAAATTAACGACAGAAACCACCAGTGCAGGACTG
 CAACCAAAGATCCTTTGGCTGCCCTGGATGCTGCTATCAGCTCCATCGACAAAT
 TCCGTTCCCTCCCTGGGTGCTATCCAGAACCGTCTGGATTCCGCAGTCACCAACC
 TGAACAACACCCTACTAACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCC
 GACTATGCGACCGAAGTGTCCAATATGTCGAAAGCGCAGATCATCCAGCAGGC
 CGGTAACCTCTGTGTGTTGGCAAAAGCTAACCAGGTACCGCAGCAGGTTCTGTC
TCTGCTGCAGGGTTAGAATCGTATTCCCGCGCTAACATGCGGCGTCTACCTCG
 TTTTTTCTGTTTCGTTTCGTCGTTCTCA

Através do BLASTn, foi observado que a sequência nucleotídica da amostra 8C apresentou 99% de similaridade com o antígeno H51, logo, foram escolhidas duas sequências do gene *fliC* do antígeno H51 depositadas com os nº de acessos: AY250027, AB028481 para serem analisadas com a sequência nucleotídica obtida da amostra 8C através do programa ClustalW. Segue abaixo o resultado da análise do ClustalW, sendo que as sequências foram numeradas de acordo com as sequências dos números de acessos descritos acima. A sequência de nº 3 é a amostra 8C.

```
seq1_ -----
seq2_ GCGATTCCTTTTATCTTTTCGACACGTAACGAATACCGGGGTTATCGGCCTGAATTGC 60
seq3_ -----
```

```

seq1_ -----
seq2_ GCAAAGTTTACGTTTAATTGTTTTTTTTTAATAGCGGGAATAAGGGGCAGAGAAAAGAGTA 120
seq3_ -----

seq1_ -----
seq2_ TTTCGGCGACTAACAAAAATGGCTGTTTGTGAAAAAATTTCTAAAGGTTGTTTTACGAC 180
seq3_ -----

seq1_ -----
seq2_ AGACGATAACAGGGTTGACGGCGATTGAGCCGACGGGTGGAAACCCAATACGTAATCAAC 240
seq3_ -----

seq1_ -----ATGGCACAAGTCATTAATACC--AACAGCCTCTCGC 34
seq2_ GACTTGCAATATAGGATAACGAATCATGGCACAAGTCATTAATACC--AACAGCCTCTCGC 299
seq3_ -----GAATTCAGTAGTGA--TTATGGCACAAGTCATTAATACC--AACAGCCTCTCGC 51
*****

seq1_ TGATCACTCAAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTC 94
seq2_ TGATCACTCAAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTC 359
seq3_ TGATCACTCAAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTC 111
*****

seq1_ TGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTA 154
seq2_ TGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTA 419
seq3_ TGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTA 171
*****

seq1_ ACCGTTTCACCTCTAACATTAAGGCCTGACTCAGGCTGCACGTAACGCTAACGATGGTA 214
seq2_ ACCGTTTCACCTCTAACATTAAGGCCTGACTCAGGCTGCACGTAACGCTAACGATGGTA 479
seq3_ ACCGTTTCACCTCTAACATTAAGGCCTGACTCAGGCTGCACGTAACGCTAACGATGGTA 231
*****

seq1_ TCTCTCTGGCGCAGACCCTGAAGGCGCACTGTCTGAGATTAACAACAACCTTACAACGTG 274
seq2_ TCTCTCTGGCGCAAACCCTGAAGGCGCACTGTCTGAGATTAACAACAACCTTACAACGTG 539
seq3_ TCTCTCTGGCGCAGACCCTGAAGGCGCACTGTCTGAGATTAACAACAACCTTACAACGTG 291
*****

seq1_ TGCGTGAGTTGACTGTACAGGCGACCACCGGTACTAACTCTGATTCTGACCTGGCTTCTA 334
seq2_ TGCGTGAGTTGACTGTACAGGCGACCACCGGTACTAACTCTGATTCTGACCTGGCTTCTA 599
seq3_ TGCGTGAGTTGACTGTACAGGCGACCACCGGTACTAACTCTGATTCTGACCTGGCTTCTA 351
*****

seq1_ TTCAGGACGAAATCAAATCCCGTTTGTCTGAAATTGACCGCGTATCCGGGCAGACCCAGT 394
seq2_ TTCAGGACGAAATCAAATCCCGTTTGTCTGAAATTGACCGCGTATCCGGGCAGACCCAGT 659
seq3_ TTCAGGACGAAATCAAATCCCGTTTGTCTGAAATTGACCGCGTATCCGGGCAGACCCAGT 411
*****

seq1_ TCAACGGCGTGAACGTATTGTCTAAAGATGGCTCCCTGAAAATTCAGGTTGGCGCAAATG 454
seq2_ TCAACGGCGTGAACGTATTGTCTAAAGATGGCTCCCTGAAAATTCAGGTTGGCGCAAATG 719
seq3_ TCAACGGCGTGAACGTATTGTCTAAAGATGGCTCCCTGAAAATTCAGGTTGGCGCAAATG 471
*****

seq1_ ATGGTCAGACTATCTCTATCGACCTGAAGAAAATTGACTCTGATACTCTGGGTTTGAATG 514
seq2_ ATGGTCAGACTATCTCTATCGACCTGAAGAAAATTGACTCTGATACTCTGGGTTTGAATG 779
seq3_ ATGGTCAGACTATCTCTATCGACCTGAAGAAAATTGACTCTGATACTCTGGGTTTGAATG 531
*****

```

```

seq1_      GTTTC AACGTTAATGGTTC TGGTACCATTGCAAACAAAGCGGCCACAATCAGTGACTTGA 574
seq2_      GTTTC AACGTTAATGGTTC TGGTACCATTGCAAACAAAGCGGCCACAATCAGTGACTTGA 839
seq3_      GTTTC AACGTTAATGGTTC TGGTACCATTGCAAACAAAGCGGCCACAATCAGTGACTTGA 591
*****

seq1_      CTGCTCAGAAAGCCGTTGACAACGGTAATGGTACTTATAAAAGTTACAAC TAGCAACGCTG 634
seq2_      CTGCTCAGAAAGCCGTTGACAACGGTAATGGTACTTATAAAAGTTACAAC TAGCAACGCTG 899
seq3_      CTGCTCAGAAAGCCGTTGACAACGGTAATGGTACTTATAAAAGTTACAAC TAGCAACGCTG 651
*****

seq1_      CACTTACTGCATCTCAGGCATTAAGTAAGCTGAGTGATGGCGATACTGTAGATATTGCAA 694
seq2_      CACTTACTGCATCTCAGGCATTAAGTAAGCTGAGTGATGGCGATACTGTAGATATTGCAA 959
seq3_      CACTTACTGCATCTCAGGCATTAAGTAAGCTGAGTGATGGCGATACTGTAGATATTGCAA 711
*****

seq1_      CCTATGCTGGTGGTACAAGTTCAACAGTTAGTTATAAAATACGACGCAGATGCAGGTAAC 754
seq2_      CCTATGCTGGTGGTACAAGTTCAACAGTTAGTTATAAAATACGACGCAGATGCAGGTAAC 1019
seq3_      CCTATGCTGGTGGTACAAGTTCAACAGTTAGTTATAAAATACGACGCAGATGCAGGTAAC 771
*****

seq1_      TCAGTTATAACAATACTGCAAACAAAACAAGTGCTGCGGCTGGAAC TCTGGCAGATACTC 814
seq2_      TCAGTTATAACAATACTGCAAACAAAACAAGTGCTGCGGCTGGAAC TCTGGCAGATACTC 1079
seq3_      TCAGTTATAACAATACTGCAAACAAAACAAGTGCTGCGGCTGGAAC TCTGGCAGATACTC 831
*****

seq1_      TTCTCCCGGCAGCTGGCCAGACTAAAACCGGTACTTACAAGGCTGCTACTGGTGATGTTA 874
seq2_      TTCTCCCGGCAGCTGGCCAGACTAAAACCGGTACTTACAAGGCTGCTACTGGTGATGTTA 1139
seq3_      TTCTCCCGGCAGCTGGCCAGACTAAAACCGGTACTTACAAGGCTGCTACTGGTGATGTTA 891
*****

seq1_      ACTTTAATGTTGACGCAACTGGTAATCTGACAATTGGCGGACAGCAAGCCTACCTGACTA 934
seq2_      ACTTTAATGTTGACGCAACTGGTAATCTGACAATTGGCGGACAGCAAGCCTACCTGACTA 1199
seq3_      ACTTTAATGTTGACGCAACTGGTAATCTGACAATTGGCGGACAGCAAGCCTACCTGACTA 951
*****

seq1_      CTGATGGTAACCTTACAACAAACAAC TCCGGTGGTGC GGCTACTGCAACTCTTAAAGAGC 994
seq2_      CTGATGGTAACCTTACAACAAACAAC TCCGGTGGTGC GGCTACTGCAACTCTTAAAGAGC 1259
seq3_      CTGATGGTAACCTTACAACAAACAAC TCCGGTGGTGC GGCTACTGCAACTCTTAAAGAGC 1011
*****

seq1_      TGTTTACTCTTGCTGGCGATGGTAAATCTCTGGGGAACGGCGGTACTGCTACCGTTACTC 1054
seq2_      TGTTTACTCTTGCTGGCGATGGTAAATCTCTGGGGAACGGCGGTACTGCTACCGTTACTC 1319
seq3_      TGTTTACTCTTGCTGGCGATGGTAAATCTCTGGGGAACGGCGGTACTGCTACCGTTACTC 1071
*****

seq1_      TGGATAATACTACGTATAATTTCAAAGCTGCTGCGAACGTTACTGATGGTGCTGGTGTC 1114
seq2_      TGGATAATACTACGTATAATTTCAAAGCTGCTGCGAACGTTACTGATGGTGCTGGTGTC 1379
seq3_      TGGATAATACTACGTATAATTTCAAAGCTGCTGCGAACGTTACTGATGGTGCTGGTGTC 1131
*****

seq1_      TCGCTGCTGCTGGTGTAAC TTATACAGCCACTGTTTCTAAAGATGTCATTCTGGCACAAC 1174
seq2_      TCGCTGCTGCTGGTGTAAC TTATACAGCCACTGTTTCTAAAGATGTCATTCTGGCACAAC 1439
seq3_      TCGCTGCTGCTGGTGTAAC TTATACAGCCACTGTTTCTAAAGATGTCATTCTGGCACAAC 1191
*****

seq1_      TGCAATCTGCAAGTCAGGCAGCAGCAACCGCTACCGACGGTGATACTGTCGCAACGATCA 1234
seq2_      TGCAATCTGCAAGTCAGGCAGCAGCAACCGCTACCGACGGTGATACTGTCGCAACGATCA 1499
seq3_      TGCAATCTGCAAGTCAGGCAGCAGCAACCGCTACCGACGGTGATACTGTCGCAACGATCA 1251
*****

```

```

seq1_ ACTATAAATCTGGTGTCTATGATCGGTTCCGCTACCTTTACCAATGGTAAAGGTTACTGCCG 1294
seq2_ ACTATAAATCTGGTGTCTATGATCGGTTCCGCTACCTTTACCAATGGTAAAGGTTACTGCCG 1559
seq3_ ACTATAAATCTGGTGTCTATGATCGGTTCCGCTACCTTTACCAATGGTAAAGGTTACTGCCG 1311
*****

seq1_ ATGGTATGACTTCTGGTACAACCTCCAGTCGTAGCTACAGGTGCTAAAGCTGTATATGTTG 1354
seq2_ ATGGTATGACTTCTGGTACAACCTCCAGTCGTAGCTACAGGTGCTAAAGCTGTATATGTTG 1619
seq3_ ATGGTATGACTTCTGGTACAACCTCCAGTCGTAGCTACAGGTGCTAAAGCTGTATATGTTG 1371
*****

seq1_ ATGGCAACAATGAACTGACTTCCACTGCATCTTACGATACGACTTACTCTGTCAACGCAG 1414
seq2_ ATGGCAACAATGAACTGACTTCCACTGCATCTTACGATACGACTTACTCTGTCAACGCAG 1679
seq3_ ATGGCAACAATGAACTGACTTCCACTGCATCTTACGATACGACTTACTCTGTCAACGCAG 1431
*****

seq1_ ATACAGGCGCAGTAAAAGTGGTATCAGGTACTGGTACTGGTAAATTTGAAGCTGTTGCTG 1474
seq2_ ATACAGGCGCAGTAAAAGTGGTATCAGGTACTGGTACTGGTAAATTTGAAGCTGTTGCTG 1739
seq3_ ATACAGGCGCAGTAAAAGTGGTATCAGGTACTGGTACTGGTAAATTTGAAGCTGTTGCTG 1491
*****

seq1_ GTGCGGATGCTTATGTAAGCAAAGATGGCAAATTAACGACAGAAACCACAGTGCAGGCA 1534
seq2_ GTGCGGATGCTTATGTAAGCAAAGATGGCAAATTAACGACAGAAACCACAGTGCAGGCA 1799
seq3_ GTGCGGATGCTTATGTAAGCAAAGATGGCAAATTAACGACAGAAACCACAGTGCAGGCA 1550
*****

seq1_ CTGCAACCAAAGATCCTTTGGCTGCCCTGGATGCTGCTATCAGCTCCATCGACAAATTCC 1594
seq2_ CTGCAACCAAAGATCCTTTGGCTGCCCTGGATGCTGCTATCAGCTCCATCGACAAATTCC 1859
seq3_ CTGCAACCAAAGATCCTTTGGCTGCCCTGGATGCTGCTATCAGCTCCATCGACAAATTCC 1610
*****

seq1_ GTTCCTCCCTGGGTGCTATCCAGAACCCTGTTGATTCCGCAGTCACCAACCTGAACAACA 1654
seq2_ GTTCCTCCCTGGGTGCTATCCAGAACCCTGTTGATTCCGCAGTCACCAACCTGAACAACA 1919
seq3_ GTTCCTCCCTGGGTGCTATCCAGAACCCTGTTGATTCCGCAGTCACCAACCTGAACAACA 1670
*****

seq1_ CCACTACTAACCTGTCTGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAG 1714
seq2_ CCACTACTAACCTGTCTGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAG 1979
seq3_ CCACTACTAACCTGTCTGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAG 1730
*****

seq1_ TGTCCAATATGTCGAAAGCGCAGATCATCCAGCAGGCCGGTAACTCTGTGT--TGGCAAA 1772
seq2_ TGTCCAATATGTCGAAAGCGCAGATCATCCAGCAGGCCGGTAACTCTGTGT--TGGCAAA 2037
seq3_ TGTCCAATATGTCGAAAGCGCAGATCATCCAGCAGGCCGGTAACTCTGTGTGTTGGCAAA 1790
*****

seq1_ AGCTAACCAGGTACCGCAGCAGTTCTGTCTCTGCTGCAGGGTTAA----- 1818
seq2_ AGCTAACCAGGTACCGCAGCAGTTCTGTCTCTGCTGCAGGGTTAATCGTTGTAACCTGA 2097
seq3_ AGCTAACCAGGTACCGCAGCAGTTCTGTCTCTGCTGCAGGGTTAGA-ATCGTATCCCG 1849
*****

seq1_ -----
seq2_ TTAAGTGAATTGCA 2156
seq3_ C--GCTAACATGCGGCGTCTACTCTGTTTTTTTC---TGTTCGTTTCGTCGTTCTCA-- 1901

seq1_ -----
seq2_ AGTTGAATTGCA 2168
seq3_ -----

```

Após a verificação das diferenças encontradas entre a sequência nucleotídica da amostra 8C e das demais amostras padrão H51 através do alinhamento pelo programa ClustalW, analisamos os eletroesferogramas da amostra 8C, e verificamos que estes apresentaram uma alta qualidade na leitura do sequenciamento. Cada base que demonstrou um resultado diferente ao encontrado no alinhamento foi verificada e constatamos que nestas regiões, estas bases realmente estão presentes.

Quando analisada no BLASTx, esta sequência de aminoácidos obteve similaridade de 99% com o antígeno H51 e foi observada uma mudança de leitura de *frame* no aminoácido 588. Os eletroesferogramas foram novamente analisados e observamos que a qualidade da leitura de sequenciamento estava acima de 98%. Segue abaixo o resultado do BLASTx.

```

dbj|BAA85090.1| flagellin [Escherichia coli]
gb|AAP13335.1| FliC [Escherichia coli]
Length=605
Score = 955 bits (2468), Expect(2) = 0.0
Identities = 579/581 (99%), Positives = 580/581 (99%), Gaps = 0/581 (0%)
Frame = +3

Query 39  NSLSLITQNNINKNQSAALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARN 218
Sbjct 8    NSLSLITQNNINKNQSAALSSSIERLSSGLRINSAKDDAAGQAIANRFTSNIKGLTQAARN 67

Query 219 ANDGISLAQTTEGALSEINNNLQRVRELTQVATTGTNSDSDLASIQDEIKSRLSEIDRVS 398
Sbjct 68  ANDGISLAQTTEGALSEINNNLQRVRELTQVATTGTNSDSDLASIQDEIKSRLSEIDRVS 127

Query 399  GQTQFNGVNVLSKDGSLKIQVGANDGQTISIDLKKIDSDTLGLNGFNVNGSGTIANKAAT 578
Sbjct 128  GQTQFNGVNVLSKDGSLKIQVGANDGQTISIDLKKIDSDTLGLNGFNVNGSGTIANKAAT 187

Query 579  ISDLTAQKAVDNGNGTYKVTTSDAALTASQALSCLSDGDTVDIATYAGGTSSTVSYKYDA 758
Sbjct 188  ISDLTAQKAVDNGNGTYKVTTSDAALTASQALSCLSDGDTVDIATYAGGTSSTVSYKYDA 247

Query 759  DPGNFSYNNTANKTSAAAGTLADTLLPAAGQTKTGTYYKAATGDVNFNVDATGNLTIGGQQ 938
Sbjct 248  DPGNFSYNNTANKTSAAAGTLADTLLPAAGQTKTGTYYKAATGDVNFNVDATGNLTIGGQQ 307

```

Query	939	AYLTTDGNLTTNNSGGAATATLKELFTLAGDGKSLGNGGTATVTLDNNTYNFKAAANVTD	1118
Sbjct	308	AYLTTDGNLTTNNSGGAATATLKELFTLAGDGKSLGNGGTATVTLDNNTYNFKAAANVTD	367
Query	1119	GAGVIAAAGVTTYTATVSKDVILAQLQSASQAAATATDGDTVATINYKSGVMIGSATFTNG	1298
Sbjct	368	GAGVIAAAGVTTYTATVSKDVILAQLQSASQAAATATDGDTVATINYKSGVMIGSATFTNG	427
Query	1299	KGTADGMTSGTTPVVATGAKAVYVDGNNELTSTASYDTTYSVNADTGAVKVVSGTGTGKF	1478
Sbjct	428	KGTADGMTSGTTPVVATGAKAVYVDGNNELTSTASYDTTYSVNADTGAVKVVSGTGTGKF	487
Query	1479	EAVAGADAYVSKDGKLTETTSAGTATKDPLAALDAAISSIDKFRSSLGAIQNRDLSAVT	1658
Sbjct	488	EAVAGADAYVSKDGKLTETTSAGTATKDPLAALDAAISSIDKFRSSLGAIQNRDLSAVT	547
Query	1659	NLNNTTTLNSEAQSRIQDADYATEVSNMSKAQIIQQAGNSV	1781
Sbjct	548	NLNNTTTLNSEAQSRIQDADYATEVSNMSKAQIIQQAGNSV	588

Score = 37.0 bits (84), Expect(2) = 0.0
 Identities = 18/18 (100%), Positives = 18/18 (100%), Gaps = 0/18 (0%)
 Frame = +2

Query	1781	VLAKANQVPQQVLSLLQG	1834
		VLAKANQVPQQVLSLLQG	
Sbjct	588	VLAKANQVPQQVLSLLQG	605

4.5.4: Amostra 9C (48629cb):

O sequenciamento da amostra 9C (48629cb) obteve um fragmento sequenciado de 1734 pb. Os dois fragmentos de DNA abaixo em **negrito** demonstram o local de alinhamentos dos iniciadores FliC(F).

GCGAATCACTAGTGATTAT**GGCACAAAGTCATTAATACCCAACAGCCTCTCGC**
 TGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCG
 AGCGTCTGTCTTCTGGCTTGCATTAACAGCGCGAAGGATGACGCCGCAGGTC
 AGGCGATTGCTAACCGTTTTACTTCTAATATTAAGGCCTGACTCAGGCTGCAC
 GTAACGCCAACGACGGTATTTCTGTTGCACAGACCACTGAAGGCGCGCTGTCC
 GAAATCAACAACAACCTTACAGCGTATTCGTGAGCTGACGGTTCAGGCGACGAC
 CGGAACACTCCACCTCTGACCTGGACTCCATTCAGGACGAAATCAAATCCC
 GTCTTGATGAAATTGACCGCGTATCCGGCCAGACCCAGTTCAACGGCGTGAAC
 GACTGGCAAAGACGGTTCGATGAAAATTCAGGTTGGTGCGAATGACGGCCA
 GACTATCACTATTGATCTGAAGAAAATTGACTCTGATACGCTGGGGCTGAGTG
 GGTTTAATGTGAATGGCAAAGGGGCTGTGGCTAATACTGCAGCGACTAAAGAT
 GATTTGGTCGCTGCATCAGTTTCAGCTGCGGTAGGTAATGAATACACTGTCTCT
 GCTGGCCTGTCAAATCAACTGCTGCTGATGTTATTGCTAGTCTCACGGATGGG
 GCGACAGTAACTGCGGCTGGTGTAAAGCAATGGTTTTGCTGCGGGGGCAACTGG

AAATGCTTATAAATTTAATCAAGCAAACAACACTTTTACTTACAATACCACCTC
AACAGCGGCAGAACTCCAATCTTACCTGACGCCTAAGGCGGGGGATAACCGCAA
CTTTCTCAGTTGAAATTGGTAGCACAAAGCAGGATGTTGTTCTGGCTAGTGATG
GCAAATCACAGCAAAGACGGGTCTAAACTTTATATTGACACCACTGGGAAT
TTAACCCAAAACGGTGGTGGTACTTTAGAAGAAGCTACCCTCAATGGTTTACGT
TTCAACCACTCTGGTCCAGCCGCTGCTGTACAATCTACTATTACTACTGCGGAT
GGAACCTCAATAGTTCTGGCAGGTTCCGGCGACTTTGGAACAACAAAACTGC
TGGGGCTATAAATGTCACAGGAGCAGTGATCAGTGCTGATGCACTTCTTTCCGC
TAGTAAAGCGACTGGGTTTACTTCTGGCGCTTATACCGTAGGTACAGATGGAGT
TGTTAAATCTGGCGGCAATGACGTTTATAACAAAGCTGACGGGACGGGATTAA
CTACTGACAATACCACAAAATATTATTTACAAGATGACGGGTCTGTAACTAATG
GTTCTGGTAAAGCTGTGTATGTTGATGCAACAGGAAAATACTACTGACGCT
GAACTAAAGCTGCAACCACCGCCGATCCCCTGAAAGCTCTGGACGAAGCCAT
CAGCTCCATCGACAAATTTTCGCTCCTCCTTGGGTGCGGTGCAAACCCTCTGGA
TTCTGCGGTGACCAACCTGAACAACACCACTACCAACCTGTCTGAAGCGCAGT
CCCGTATTCACGGACGCCGACTATGCGACCGAAGTGTCCAACATGTCGAAAGC
GCAGATTATCCAGCAGGCCGGTAACTCCGTGCTGGCAAAGCCAACCAGGTCC
CGCAGCAGGTTCTGTCTCTGCAGCAGGGGTAGATCGATTCCCAGCCTCACC
GGGTGCCGAATTTAAGGGGGGGGTCA

Através do BLASTn, foi observado que a sequência nucleotídica da amostra 9C apresentou 99% de similaridade com o antígeno H31, logo, foram escolhidas duas sequências do gene *fliC* do antígeno H31 depositadas com os nº de acessos: AY250013, CP000247, para serem analisadas com a sequência nucleotídica obtida da amostra 9C através do programa ClustalW. Segue abaixo o resultado da análise do ClustalW, sendo que as sequências foram numeradas de acordo com as sequências dos números de acessos descritos acima. A sequência de nº 3 é a amostra 9C.

```
seq2_      -----ATGGCACAAGTCATTAATACC--AACAGCCTCTCGCTGATCACT 42
seq3_      GCGAATCACTAGTGATTATGGCACAAGTCATTAATACC--AACAGCCTCTCGCTGATCACT 60
seq1_      -----ATGGCACAAGTCATTAATACC--AACAGCCTCTCGCTGATCACT 42
          *****

seq2_      CAAAATAATATCAACAAGAACCAGTCTGCGCTGTGCGAGTTCTATCGAGCGTCTGTCTTCT 102
seq3_      CAAAATAATATCAACAAGAACCAGTCTGCGCTGTGCGAGTTCTATCGAGCGTCTGTCTTCT 120
seq1_      CAAAATAATATCAACAAGAACCAGTCTGCGCTGTGCGAGTTCTATCGAGCGTCTGTCTTCT 102
          *****

seq2_      GGCTTGCGTATTAACAGCGCGAAGGATGACGCCGCGGGTCAGGCGATTGCTAACCGTTTT 162
seq3_      GGCTTGCGTATTAACAGCGCGAAGGATGACGCCGCGAGGTCAGGCGATTGCTAACCGTTTT 180
seq1_      GGCTTGCGTATTAACAGCGCTAAGGATGACGCCGCGGGTCAGGCGATTGCTAACCGTTTT 162
          *****
```

```

seq2_      ACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGATGGTATTTCTGTT 222
seq3_      ACTTCTAAATATTAAGGCCTGACTCAGGCTGCACGTAACGCCAACGCGGTATTTCTGTT 240
seq1_      ACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGCGGTATTTCTGTT 222
*****

seq2_      GCGCAGACCACCGAAGGCGCGCTGTCCGAAATTAACAACAACCTTACAGCGTATTCGTGAA 282
seq3_      GCACAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTATTCGTGAG 300
seq1_      GCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTATTCGTGAA 282
** *****

seq2_      CTGACGGTTCAGGCTTCTACCGGCACTAACTCTGATTCCGGATCTGGACTCCATTACAGGAC 342
seq3_      CTGACGGTTCAGGCGACGACCGGAACCTCCACCTCTGACTGGACTCCATTACAGGAC 360
seq1_      CTGACGGTTCAGGCTTCTACCGGGACTAACTCCGATTCGGATCTGGACTCCATTACAGGAC 342
*****

seq2_      GAAATCAAATCCCGTCTCGACGAAATTGACCCGCTATCCGGTCAGACCAGTTCAACGGC 402
seq3_      GAAATCAAATCCCGTCTTGATGAAATTGACCCGCTATCCGGCCAGACCAGTTCAACGGC 420
seq1_      GAAATCAAATCCCGTCTGGACGAAATTGACCCGCTATCTGGCCAGACCAGTTCAACGGC 402
*****

seq2_      GTGAACGTACTGGCAAAGACGGTTCGATGAAATTCAGGTTGGTGCGAATGACGGCCAG 462
seq3_      GTGAACGTACTGGCAAAGACGGTTCGATGAAATTCAGGTTGGTGCGAATGACGGCCAG 480
seq1_      GTGAACGTACTGGCAAAGACGGTTCATGAAATTCAGGTTGGTGCGAATGACGGCCAG 462
*****

seq2_      ACTATCACTATTGATCTGAAGAAAATTGACTCTGATACGCTGGGGCTGAGTGGGTTTAAAT 522
seq3_      ACTATCACTATTGATCTGAAGAAAATTGACTCTGATACGCTGGGGCTGAGTGGGTTTAAAT 540
seq1_      ACTATCACTATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAGTGGGTTTAAAT 522
*****

seq2_      GTGAATGGCAAAGGGGCTGTGGCTAATACTGCAGCGACTAAAGATGATTTGGTTCGCTGCA 582
seq3_      GTGAATGGCAAAGGGGCTGTGGCTAATACTGCAGCGACTAAAGATGATTTGGTTCGCTGCA 600
seq1_      GTGAATGGTGGCGGGGCTGTTGCTAATACTGCAGCGACTAAAGATGATTTGGTTCGCTGCA 582
*****

seq2_      TCAGTTTCAGCTGCGGTAGGTAATGAATACACTGTCTCTGCTGGCCTGTCGAAATCAACT 642
seq3_      TCAGTTTCAGCTGCGGTAGGTAATGAATACACTGTCTCTGCTGGCCTGTCGAAATCAACT 660
seq1_      TCAGTTTCAGCTGCGGTAGGTAATGAATACACTGTCTCTGCTGGCCTGTCGAAATCAACT 642
*****

seq2_      GCTGCTGATGTTATTGCTAGTCTCACGGATGGGGCGACAGTAACTGCGGCTGGTGTAAAGC 702
seq3_      GCTGCTGATGTTATTGCTAGTCTCACGGATGGGGCGACAGTAACTGCGGCTGGTGTAAAGC 720
seq1_      GCTGCTGATGTTATTGCTAGTCTCACAGATGGTGCAGTAACTGCGGCTGGTGTAAAGC 702
*****

seq2_      AATGGTTTTGCTGCGGGGCAACTGGAATGCTTATAAATTTAATCAAGCAAACAACACT 762
seq3_      AATGGTTTTGCTGCGGGGCAACTGGAATGCTTATAAATTTAATCAAGCAAACAACACT 780
seq1_      AATGGTTTTGCTGCGGGGCAACTGGAGATGCTTATAAATTTAATCAAGCAAACAACACT 762
*****

seq2_      TTTACTTACAATACCACCTCAACAGCGGCAGAACTCCAATCTTACCTGACGCCTAAGGCG 822
seq3_      TTTACTTACAATACCACCTCAACAGCGGCAGAACTCCAATCTTACCTGACGCCTAAGGCG 840
seq1_      TTTACTTACAATACCACCTCAACAGCGGCAGAACTCCAATCTTACCTGACGCCTAAGGCG 822
*****

seq2_      GGGGATACCGCAACTTTCTCAGTTGAAATTTGGTAGCACAAAGCAGGATGTTGTTCTGGCT 882
seq3_      GGGGATACCGCAACTTTCTCAGTTGAAATTTGGTAGCACAAAGCAGGATGTTGTTCTGGCT 900
seq1_      GGGGATACCGCAACTTTCTCCGTTGAAATTTGGTGGCACCACCAAGCAGGATGTTGTTCTGGCT 882
*****

```

```

seq2_      AGTGATGGCAAAATCACAGCAAAAGACGGGTCTAAACTTTATATTGACACCACTGGGAAT 942
seq3_      AGTGATGGCAAAATCACAGCAAAAGACGGGTCTAAACTTTATATTGACACCACTGGGAAT 960
seq1_      AGTGATGGCAAAATCACAGCAAAAGACGGGTCTAAACTTTATATTGACACCACTGGGAAT 942
*****

seq2_      TTAACCCAAAACGGTGGTGGTACTTTAGAAGAAGCTACCCTCAATGGTTTAGCTTTCAAC 1002
seq3_      TTAACCCAAAACGGTGGTGGTACTTTAGAAGAAGCTACCCTCAATGGTTTAGCTTTCAAC 1020
seq1_      TTAACCCAAAACGGTGGAGGTACTTTAGAAGAAGCTACCCTCAATGGTTTAGCTTTCAAC 1002
*****

seq2_      CACTCTGGTCCAGCCGCTGCTGTACAATCTACTATTACTACTGCGGATGGAACCTCAATA 1062
seq3_      CACTCTGGTCCAGCCGCTGCTGTACAATCTACTATTACTACTGCGGATGGAACCTCAATA 1080
seq1_      CACTCTGGTCCAGCCGCTGCTGTACAATCTACTATTACTACTGCGGATGGAACCTCAATA 1062
*****

seq2_      GTTCTGGCAGGTTCTGGCGACTTTGGAACAACAAAACTGCTGGGGCTATAAAATGTCACA 1122
seq3_      GTTCTGGCAGGTTCTGGCGACTTTGGAACAACAAAACTGCTGGGGCTATAAAATGTCACA 1140
seq1_      GTTCTAGCAGGTTCTGGCGACTTTGGAACAACAAAACTGCTGGGGCTATAAAATGTCACA 1122
*****

seq2_      GGAGCAGTGATCAGTGCTGATGCACTTCTTTCCGCTAGTAAAGCGACTGGGTTTACTTCT 1182
seq3_      GGAGCAGTGATCAGTGCTGATGCACTTCTTTCCGCTAGTAAAGCGACTGGGTTTACTTCT 1200
seq1_      GGAGCAGTGATCAGTGCTGATGCACTTCTTTCCGCCAGTAAAGCGACTGGGTTTACTTCT 1182
*****

seq2_      GCGCTTATAACCGTAGGTACAGATGGAGTTGTTAAATCTGGCGGCAATGACGTTTATAAC 1242
seq3_      GCGCTTATAACCGTAGGTACAGATGGAGTTGTTAAATCTGGCGGCAATGACGTTTATAAC 1260
seq1_      GGCCTTATAACCGTAGGTACAGATGGAGTTGTTAAATCTGGTGGCAATGACGTTTATAAC 1242
***

seq2_      AAAGCTGACGGGACGGGATTAACACTGACAATACCACAAAATATTATTTACAAGATGAC 1302
seq3_      AAAGCTGACGGGACGGGATTAACACTGACAATACCACAAAATATTATTTACAAGATGAC 1320
seq1_      AAAGCTGACGGGACGGGATTAACACTGACAATACCACAAAATATTATTTACAAGATGAC 1302
*****

seq2_      GGGTCTGTAACATAATGGTTCTGGTAAAGCTGTGTATGTTGATGCAACAGGAAAACCTA 1362
seq3_      GGGTCTGTAACATAATGGTTCTGGTAAAGCTGTGTATGTTGATGCAACAGGAAAACCTA 1380
seq1_      GGGTCTGTAACATAATGGTTCTGGTAAAGCTGTGTATGTTGATGCAACAGGAAAACCTA 1362
*****

seq2_      ACTGACGCTGAAACTAAAGCTGCAACCACCGCCGATCCCCTGAAAGCTCTGGACGAAGCC 1422
seq3_      ACTGACGCTGAAACTAAAGCTGCAACCACCGCCGATCCCCTGAAAGCTCTGGACGAAGCC 1440
seq1_      ACTGACGCTGAAACTAAAGCCGAAACCACCGCCGATCCCCTGAAAGCTCTGGACGAAGCC 1422
*****

seq2_      ATCAGCTCCATCGACAAAATTTCCGCTCCTCCTTGGGTGCGGTGCAAAACCGTCTGGATTCT 1482
seq3_      ATCAGCTCCATCGACAAAATTTCCGCTCCTCCTTGGGTGCGGTGCAAAACCGTCTGGATTCT 1500
seq1_      ATCAGCTCCATCGACAAAATTTCCGTTCTTCCCTCGGTGCGGTGCAAAACCGTCTGGATTCC 1482
*****

seq2_      GCGGTGACCAACCTGAACAACACCCTACCAACCTGTCTGAAGCGCAGTCCCCTATTCA- 1541
seq3_      GCGGTGACCAACCTGAACAACACCCTACCAACCTGTCTGAAGCGCAGTCCCCTATTCA- 1560
seq1_      GCGGTGACCAACCTGAACAACACCCTACCAACCTGTCTGAAGCGCAGTCCCCTATTCA- 1541
*****

seq2_      GGACGCCGACTATGCGACCGAAGTGTCCAACATGTGCGAAAGCGCAGATCATCCAGCAGGC 1601
seq3_      GGACGCCGACTATGCGACCGAAGTGTCCAACATGTGCGAAAGCGCAGATCATCCAGCAGGC 1620
seq1_      GGACGCCGACTATGCGACCGAAGTGTCCAACATGTGCGAAAGCGCAGATCATCCAGCAGGC 1601
*****

```

```

seq2_      CGGTAACTCCGTGCTGGCAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCA 1661
seq3_      CGGTAACTCCGTGCTGGCAAAGCQAACCAGGTCCGCAGCAGGTTCTGTCTCTGCA 1680
seq1_      CGGTAACTCCGTGCTGGCAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCA 1661
*****
seq2_      GGGTTAA----- 1668
seq3_      GGGGTAGATCGATTCCCGCGCCTCACCGGGTGCCGCAATTTAAGGGGGGGTCA 1734
seq1_      GGGTTAA----- 1668
*** **

```

Após a verificação das diferenças encontradas entre a sequência nucleotídica da amostra 9C e das demais amostras padrão H31 através do alinhamento pelo programa ClustalW, analisamos os eletroesferogramas da amostra 9C, e verificamos que estes apresentaram uma alta qualidade na leitura do sequenciamento. Cada base que demonstrou um resultado discrepante ao encontrado no alinhamento foi verificada e constatamos que nestas regiões, estas bases realmente estão presentes, e não são artefatos ou erros de leitura.

Quando analisadas no BLASTx, a sequência de aminoácidos obteve 99% de similaridade com o antígeno H31, porém, foi observada uma mudança de leitura de frame por volta do aminoácido 513. Os eletroesferogramas foram novamente analisados e observamos que a qualidade da leitura de sequenciamento estava acima de 98%, além disto, realizamos várias repetições diferentes de reações de sequenciamento da mesma amostra. Segue abaixo o resultado do BLASTx.

```

ref|YP_669759.1| flagellin [Escherichia coli 536]
ref|ZP_03034867.1| flagellin [Escherichia coli F11]
gb|ABG69858.1| flagellin [Escherichia coli 536]
gb|EDV66063.1| flagellin [Escherichia coli F11]
Length=555
Score = 824 bits (2129), Expect(2) = 0.0
Identities = 503/506 (99%), Positives = 504/506 (99%), Gaps = 0/506 (0%)
Frame = +1

Query 40  NSLSLITQNNINKNQSA LSSSIERLSSGLRINS AKDDAAGQAIANRFTSNIKGLTQAARN 219
          NSLSLITQNNINKNQSA LSSSIERLSSGLRINS AKDDAAGQAIANRFTSNIKGLTQAARN
Sbjct  8  NSLSLITQNNINKNQSA LSSSIERLSSGLRINS AKDDAAGQAIANRFTSNIKGLTQAARN 67

```

Query	220	ANDGISVAQTTEGALSEINNNLQRIRELTVQATTGTNSTSDLDSIQDEIKSRLDEIDRVS	399
Sbjct	68	ANDGISVAQTTEGALSEINNNLQRIRELTVQA+TGTNS SDLDSIQDEIKSRLDEIDRVS	127
Query	400	GQTQFNGVNVFLAKDGSMKIQVGANDGQTTITIDLKIDSDDLGLSGFNVNGKAVANTAAT	579
Sbjct	128	GQTQFNGVNVFLAKDGSMKIQVGANDGQTTITIDLKIDSDDLGLSGFNVNGKAVANTAAT	187
Query	580	KDDLVAASVSAAVGNEYTVSAGLSKSTAADVIASLTDGATVTAAGVSNNGFAAGATGNAYK	759
Sbjct	188	KDDLVAASVSAAVGNEYTVSAGLSKSTAADVIASLTDGATVTAAGVSNNGFAAGATGNAYK	247
Query	760	FNQANNTFTYNTTSTAELQSYLTPKAGDTATFSVEIGSTKQDVVLASDGKITAKDGSKL	939
Sbjct	248	FNQANNTFTYNTTSTAELQSYLTPKAGDTATFSVEIGSTKQDVVLASDGKITAKDGSKL	307
Query	940	YIDTTGNLTQNGGGTLEEATLNGLAFNHSGPAAAVQSTITTTADGTSIVLAGSGDFGTTKT	1119
Sbjct	308	YIDTTGNLTQNGGGTLEEATLNGLAFNHSGPAAAVQSTITTTADGTSIVLAGSGDFGTTKT	367
Query	1120	AGAINVTGAVISADALLSASKATGFTSGAYTVGTDGVVKS GGNDVYNKADGTGLTTDNTT	1299
Sbjct	368	AGAINVTGAVISADALLSASKATGFTSGAYTVGTDGVVKS GGNDVYNKADGTGLTTDNTT	427
Query	1300	KYYLQDDGSVTNGSGKAVYVDATGKLTDAETKAATTADPLKALDEAISSIDKFRSSLGA	1479
Sbjct	428	KYYLQDDGSVTNGSGKAVYVDATGKLTDAETKAATTADPLKALDEAISSIDKFRSSLGA	487
Query	1480	VQNPLDSAVTNLNNTTNLSEAQSRI	1557
Sbjct	488	VQN LDSAVTNLNNTTNLSEAQSRI	513

Score = 77.8 bits (190), Expect(2) = 0.0
Identities = 40/41 (97%), Positives = 40/41 (97%), Gaps = 0/41 (0%)
Frame = +2

Query	1562	DADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLQQG	1684
Sbjct	515	DADYATEVSNMSKAQIIQQAGNSVLAKANQVPQQVLSLQQG	555

4.5.5: Amostra 11A (00-05951)

O sequenciamento da amostra 11A (00-05951), obteve um fragmento sequenciado de 1869 pb. As duas sequências em **negrito** demonstram o local de alinhamento dos iniciadores FliC(F).

TGCAGCGGCGGAATCACTAGTGATTAT**TGGCACAAGTCATTAATACCCAACA**
GCCTCTCGCTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGA
GTTCTATCGAGCGTCTGTCTTCTGGCTTGCATTAACAGCGCGAAGGATGACG

CAGCGGGTCAGGCGATTGCTAACCGTTTCACCTCTAACATTAAGGCCTGACTC
AGGCGGCCCGTAACGCCAACGACGGTATCTCCGTTGCGCAGACCACCGAAGGC
GCGCTGTCCGAAATCAACAACAACCTTACAGCGTATCCGTGAACTGACGGTTCA
GGCTTCTACCGGGACTAACTCCGATTCGGATCTGGACTCCATTCAGGACGAAAT
CAAATCCCCGTCTGGACGAAATTGACCGCGTATCTGGCCAGACCCAGTTCAAC
GGCGTGAACGTACTGGCGAAAGACGGTTCAATGAAAATTCAGGTTGGTGCGAA
TGACGGCAGACTATCACGATTGATCTGAAGAAAATTGACTCAGATACGCTGGG
GCTGAATGGTTTAAACGTGAATGGTTCCGGTACGATAGCCAATAAAGCGGGCAG
CATTAGCGACCTGACAGCAGCGAAAATGGATGCTGCAACTAATACTATAACTA
CAACAATAATGCGCTGACTGCATCAAAGGCGCTTGATCAACTGAAAGATGGT
GACTGTACTATCAAAGCAGATGCTGCTCAAACCTGCCACGGTTTATACATAC
AATGCATCAGCTGGTAACTTCTCATTTCAGTAATGTATCGAATAATACTTCAGCA
AAAGCAGGTGATGTAGCAGCTAGCCTTCTCCCGCCGGCTGGGCAAACCTGCTAG
TGGTGTTTATAAAGCAGCAAGCGGTGAAGTGAACCTTTGATGTTGATGCGAATG
GTAATAATCACAATCGGAGGACAGAAAGCATATTTAACTAGTGATGGTAACTTA
ACTACAAACGATGCTGGTGGTGGCGACTGCGGCTACGCTTGATGGTTTATTCAAG
AAAGCTGGTGGTCAATCAATCGGGTTTAAAGAAGACTGCATCAGTCACGAT
GGGGGGAACAACCTTATAACTTTCAAACGGGTCTGATGCTGATGCTGCAACTG
CTAACGCAGGGTTATCGTTCACTGATACAGCTAGCAAAGATACCGTTTTAAATA
AAGTGGCTACAGCTTATACAAGGGCAAAGCAGTTGCAGCTGACCGGTGATACA
TCCGCAACAATGACCTATAAATCTGGCGTTCAGACGTATCAGGCTGTATTTGCC
GCAGGTGACGGTACTGCTAGCGCAAATATGCCGATAAAGCTGACGTTTCTAA
TGCAACAGCAACATACACTGATGCTGATGGTGAATGACTACAATTGGTTCAT
ACACCACGAAGTATTCAATCGATGCTAACAACGGCAAGGTAACCTGTTGATTCT
GGAACCTGGTACGGGTAAATATGCGCCGAAAGTAGGGGCTGAAGTATATGTTAG
TGCTAATGGTACTTTAACAACAGATGCAACTAGCGAAGGCACAGTAACAAAAG
ATCCACTGAAAGCTCTGGATGAAGCTATCAGCTCCATCGACAAATTCGGTCTT
CCCTGGGTGCTATCCAGAACCGTCTGGATTCCGCAGTCACCAACCTGAACAAC
ACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTCAGGACGCCGACTATGC
GACCGAAGTGTCCAACATGTTCGAAAGCGCAGATCATTTCAGCAGGCCGGTAACT
CCGTGCTGGCAAAGCCAACCAGGTACCGCAGCAGGTTCTGTCTCAGCTGCA
GGGTTAGAATCGAATTNCCCGCGCTGACCGTTGTGCGATTTCTCNTCCTTTGTC
CTCGT

Através do BLASTn, foi observado que a sequência nucleotídica da amostra 11A apresentou 99% de similaridade com o antígeno H12. Sequências do gene *fliC* do antígeno H12 depositadas com os nº de acessos: AY337474, AY337471, AY249997, L07389, foram utilizadas para análise da sequência nucleotídica da amostra 11A através do programa ClustalW. Segue abaixo o resultado da análise do ClustalW, sendo que as sequências foram

numeradas de acordo com as sequências dos números de acessos descritos acima. A sequência de nº 5 é a amostra 11A.

```

seq3_ -----ATGGCACAAAGTCATTAATACC--AACAGCCTCTCG 33
seq4_ -----ATGGCACAAAGTCATTAATACC--AACAGCCTCTCG 33
seq5_ TGCAGCGGCGGAATCACTAGTGATTATGGCACAAAGTCATTAATACC AACAGCCTCTCG 60
seq1_ -----ATGGCACAAAGTCATTAATACC--AACAGCCTCTCG 33
seq2_ -----ATGGCACAAAGTCATTAATACC--AACAGCCTCTCG 33
*****

seq3_ CTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCATATCGAGCGT 93
seq4_ CTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCATATCGAGCGT 93
seq5_ CTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCATATCGAGCGT 120
seq1_ CTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCATATCGAGCGT 93
seq2_ CTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCATATCGAGCGT 93
*****

seq3_ CTGTCTTCTGGCTTGCCTATTAAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCT 153
seq4_ CTGTCTTCTGGCTTGCCTATTAAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCT 153
seq5_ CTGTCTTCTGGCTTGCCTATTAAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCT 180
seq1_ CTGTCTTCTGGCTTGCCTATTAAACAGCGCGAAGGATGACGCCGAGGTCAGGCGATTGCT 153
seq2_ CTGTCTTCTGGCTTGCCTATTAAACAGCGCGAAGGATGACGCCGAGGTCAGGCGATTGCT 153
*****

seq3_ AACCGTTTCACCTCTAACATTTAAAGGCCTGACTCAGGCGGCCCGTAACGCCAACGACGGT 213
seq4_ AACCGTTTCACCTCTAACATTTAAAGGCCTGACTCAGGCGGCCCGTAACGCCAACGACGGT 213
seq5_ AACCGTTTCACCTCTAACATTTAAAGGCCTGACTCAGGCGGCCCGTAACGCCAACGACGGT 240
seq1_ AACCGTTTACTTCTAACATTTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGT 213
seq2_ AACCGTTTACTTCTAACATTTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGT 213
*****

seq3_ ATCTCCGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGT 273
seq4_ ATCTCCGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGT 273
seq5_ ATCTCCGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGT 300
seq1_ ATTTCCGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGT 273
seq2_ ATTTCCGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGT 273
** *****

seq3_ ATCCGTGAAC TGACGGTTCAGGCTTCTACCGGACTA ACTCCGATTCGGATCTGGACTCC 333
seq4_ ATCCGTGAAC TGACGGTTCAGGCTTCTACCGGACTA ACTCCGATTCGGATCTGGACTCC 333
seq5_ ATCCGTGAAC TGACGGTTCAGGCTTCTACCGGACTA ACTCCGATTCGGATCTGGACTCC 360
seq1_ ATCCGTGAAC TGACGGTTCAGGCTTCTACCGGACTA ACTCCGATTCAGATCTGGACTCC 333
seq2_ ATCCGTGAAC TGACGGTTCAGGCTTCTACCGGACTA ACTCCGATTCGGATCTGGACTCC 333
*****

seq3_ ATTCAGGACGAAATCAAATCCC-GTCTGGACGAAATTGACCGCGTATCTGGCCAGACCCA 392
seq4_ ATTCAGGACGAAATCAAATCCC-GTCTGGACGAAATTGACCGCGTATCTGGCCAGACCCA 392
seq5_ ATTCAGGACGAAATCAAATCCCGTCTGGACGAAATTGACCGCGTATCTGGCCAGACCCA 420
seq1_ ATTCAGGACGAAATCAAATCCC-GTCTGGACGAAATTGACCGCGTATCTGGCCAGACCCA 392
seq2_ ATTCAGGACGAAATCAAATCCC-GTCTGGACGAAATTGACCGCGTATCTGGCCAGACCCA 392
*****

seq3_ GTTCAACGGCGTGAACGTACTGGCGAAAGACGGTTCAATGAAAATTCAGGTTGGTGC GAA 452
seq4_ GTTCAACGGCGTGAACGTACTGGCGAAAGACGGTTCAATGAAAATTCAGGTTGGTGC GAA 452
seq5_ GTTCAACGGCGTGAACGTACTGGCGAAAGACGGTTCAATGAAAATTCAGGTTGGTGC GAA 480
seq1_ GTTCAACGGCGTGAACGTACTGGCGAAAGACGGTTCAATGAAAATTCAGGTTGGTGC GAA 452
seq2_ GTTCAACGGCGTGAACGTACTGGCGAAAGACGGTTCAATGAAAATTCAGGTTGGTGC GAA 452
*****

```

```

seq3_ TGACGGCCAGACTATCACGATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAA 512
seq4_ TGACGGCCAGACTATCACGATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAA 512
seq5_ TGACGGCAGACTATCACGATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAA 539
seq1_ TGACGGCCAGACTATCACGATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAA 512
seq2_ TGACGGCCAGACTATCACGATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAA 512
*****

seq3_ TGGTTTTAACGTGAATGGTTCGGTACGATAGCCAATAAAGCGGCGACCATTAGCGACCT 572
seq4_ TGGTTTTAACGTGAATGGTTCGGTACGATAGCCAATAAAGCGGCGACCATTAGCGACCT 572
seq5_ TGGTTTAAACGTGAATGGTTCGGTACGATAGCCAATAAAGCGGCGACCATTAGCGACCT 598
seq1_ TGGTTTTAACGTGAATGGTTCGGTACGATAGCCAATAAAGCGGCGACCATTAGCGACCT 572
seq2_ TGGTTTTAACGTGAATGGTTCGGTACGATAGCCAATAAAGCGGCGACCATTAGCGACCT 572
*****

seq3_ GACAGCAGCGAAAAATGGATGCTGCAACTAATACTATAACTACAACAAATAATGCGCTGAC 632
seq4_ GACAGCAGCGAAAAATGGATGCTGCAACTAATACTATAACTACAACAAATAATGCGCTGAC 632
seq5_ GACAGCAGCGAAAAATGGATGCTGCAACTAATACTATAACTACAACAAATAATGCGCTGAC 658
seq1_ GACAGCAGCGAAAAATGGATGCTGCAACTAATACTATAACTACAACAAATAATGCGCTGAC 632
seq2_ GACAGCAGCGAAAAATGGATGCTGCAACTAATACTATAACTACAACAAATAATGCGCTGAC 632
*****

seq3_ TGCATCAAAGGCGCTTGATCAACTGAAAGATGGTGACACTGTTACTATCAAAGCAGATGC 692
seq4_ TGCATCAAAGGCGCTTGATCAACTGAAAGATGGTGACACTGTTACTATCAAAGCAGATGC 692
seq5_ TGCATCAAAGGCGCTTGATCAACTGAAAGATGGTGACACTGTTACTATCAAAGCAGATGC 718
seq1_ TGCATCAAAGGCGCTTGATCAACTGAAAGATGGTGACACTGTTACTATCAAAGCAGATGC 692
seq2_ TGCATCAAAGGCGCTTGATCAACTGAAAGATGGTGACACTGTTACTATCAAAGCAGATGC 692
*****

seq3_ TGCTCAAACCTGCCACGGTTTATACATACAATGCATCAGCTGGTAACCTTCTCATTAGTAA 752
seq4_ TGCTCAAACCTGCCACGGTTTATACATACAATGCATCAGCTGGTAACCTTCTCATTAGTAA 752
seq5_ TGCTCAAACCTGCCACGGTTTATACATACAATGCATCAGCTGGTAACCTTCTCATTAGTAA 778
seq1_ TGCTCAAACCTGCCACGGTTTATACATACAATGCATCAGCTGGTAACCTTCTCATTAGTAA 752
seq2_ TGCTCAAACCTGCCACGGTTTATACATACAATGCATCAGCTGGTAACCTTCTCATTAGTAA 752
*****

seq3_ TGTATCGAATAATACTTCAGCAAAAAGCAGGTGATGTAGCAGCTAGCCTTCTCCCGCCGGC 812
seq4_ TGTATCGAATAATACTTCAGCAAAAAGCAGGTGATGTAGCAGCTAGCCTTCTCCCGCCGGC 812
seq5_ TGTATCGAATAATACTTCAGCAAAAAGCAGGTGATGTAGCAGCTAGCCTTCTCCCGCCGGC 838
seq1_ TGTATCGAATAATACTTCAGCAAAAAGCAGGTGATGTAGCAGCTAGCCTTCTCCCGCCGGC 812
seq2_ TGTATCGAATAATACTTCAGCAAAAAGCAGGTGATGTAGCAGCTAGCCTTCTCCCGCCGGC 812
*****

seq3_ TGGGCAAACCTGCTAGTGGTGTTTATAAAGCAGCAAGCGGTGAAGTGAACCTTTGATGTTGA 872
seq4_ TGGGCAAACCTGCTAGTGGTGTTTATAAAGCAGCAAGCGGTGAAGTGAACCTTTGATGTTGA 872
seq5_ TGGGCAAACCTGCTAGTGGTGTTTATAAAGCAGCAAGCGGTGAAGTGAACCTTTGATGTTGA 898
seq1_ TGGGCAAACCTGCTAGTGGTGTTTATAAAGCAGCAAGCGGTGAAGTGAACCTTTGATGTTGA 872
seq2_ TGGGCAAACCTGCTAGTGGTGTTTATAAAGCAGCAAGCGGTGAAGTGAACCTTTGATGTTGA 872
*****

seq3_ TGCGAATGGTAAAAATCACAATCGGAGGACAGAAAAGCATAATTTAACTAGTGATGGTAACTT 932
seq4_ TGCGAATGGTAAAAATCACAATCGGAGGACAGAAAAGCATAATTTAACTAGTGATGGTAACTT 932
seq5_ TGCGAATGGTAAAAATCACAATCGGAGGACAGAAAAGCATAATTTAACTAGTGATGGTAACTT 958
seq1_ TGCGAATGGTAAAAATCACAATCGGAGGACAGAAAAGCATAATTTAACTAGTGATGGTAACTT 932
seq2_ TGCGAATGGTAAAAATCACAATCGGAGGACAGAAAAGCATAATTTAACTAGTGATGGTAACTT 932
*****

seq3_ AACTACAAACGATGCTGGTGGTGC GACTGCGGCTACGCTTGATGGTTTATTCAAGAAAGC 992
seq4_ AACTACAAACGATGCTGGTGGTGC GACTGCGGCTACGCTTGATGGTTTATTCAAGAAAGC 992
seq5_ AACTACAAACGATGCTGGTGGTGC GACTGCGGCTACGCTTGATGGTTTATTCAAGAAAGC 1018
seq1_ AACTACAAACGATGCTGGTGGTGC GACTGCGGCTACGCTTGATGGTTTATTCAAGAAAGC 992
seq2_ AACTACAAACGATGCTGGTGGTGC GACTGCGGCTACGCTTGATGGTTTATTCAAGAAAGC 992
*****

```

```

seq3_ TGGTGATGGTCAATCAATCGGGTTTAAAGAAGACTGCATCAGTCACGATGGGGGGAACAAC 1052
seq4_ TGGTGATGGTCAATCAATCGGGTTTAAAGAAGACTGCATCAGTCACGATGGGGGGAACAAC 1052
seq5_ TGGTGATGGTCAATCAATCGGGTTTAAAGAAGACTGCATCAGTCACGATGGGGGGAACAAC 1078
seq1_ TGGTGATGGTCAATCAATCGGGTTTAAAGAAGACTGCATCAGTCACGATGGGGGGAACAAC 1052
seq2_ TGGTGATGGTCAATCAATCGGGTTTAAAGAAGACTGCATCAGTCACGATGGGGGGAACAAC 1052
*****

seq3_ TTATAACTTTAAAACGGGTGCTGATGCTGATGCTGCAACTGCTAACGCAGGGGTATCGTT 1112
seq4_ TTATAACTTTAAAACGGGTGCTGATGCTGATGCTGCAACTGCTAACGCAGGGGTATCGTT 1112
seq5_ TTATAACTTTCAAACGGGTCTGATGCTGATGCTGCAACTGCTAACGCAGGGTTATCGTT 1138
seq1_ TTATAACTTTAAAACGGGTGCTGATGCTGATGCTGCAACTGCTAACGCAGGGGTATCGTT 1112
seq2_ TTATAACTTTAAAACGGGTGCTGATGCTGATGCTGCAACTGCTAACGCAGGGGTATCGTT 1112
*****

seq3_ CACTGATACAGCTAGCAAAGAAACCGTTTTAAATAAAGTGGCTACAGCTAA-ACAAGG-C 1170
seq4_ CACTGATACAGCTAGCAAAGAAACCGTTTTAAATAAAGTGGCTACAGCTAA-ACAAGG-C 1170
seq5_ CACTGATACAGCTAGCAAAGAAACCGTTTTAAATAAAGTGGCTACAGCTTAAACAAGGGC 1198
seq1_ CACTGATACAGCTAGCAAAGAAACCGTTTTAAATAAAGTGGCTACAGCTAA-ACAAGG-C 1170
seq2_ CACTGATACAGCTAGCAAAGAAACCGTTTTAAATAAAGTGGCTACAGCTAA-ACAAGG-C 1170
*****

seq3_ AAAGCAGTTGCAGCTGAC-GGTGATACATCCGCAACAATTACCTATAAATCTGGCGTTCA 1229
seq4_ AAAGCAGTTGCAGCTGAC-GGTGATACATCCGCAACAATTACCTATAAATCTGGCGTTCA 1229
seq5_ AAAGCAGTTGCAGCTGACCGGTGATACATCCGCAACAATTACCTATAAATCTGGCGTTCA 1258
seq1_ AAAGCAGTTGCAGCTGAC-GGTGATACATCCGCAACAATTACCTATAAATCTGGCGTTCA 1229
seq2_ AAAGCAGTTGCAGCTGAC-GGTGATACATCCGCAACAATTACCTATAAATCTGGCGTTCA 1229
*****

seq3_ GACGTATCAGGCTGTATTTGCCGCAGGTGACGGTACTGCTAGCGCAAAATATGCCGATAA 1289
seq4_ GACGTATCAGGCTGTATTTGCCGCAGGTGACGGTACTGCTAGCGCAAAATATGCCGATAA 1289
seq5_ GACGTATCAGGCTGTATTTGCCGCAGGTGACGGTACTGCTAGCGCAAAATATGCCGATAA 1318
seq1_ GACGTATCAGGCTGTATTTGCCGCAGGTGACGGTACTGCTAGCGCAAAATATGCCGATAA 1289
seq2_ GACGTATCAGGCGGTATTTGCCGCAGGTGACGGTACTGCTAGCGCAAAATATGCCGATAA 1289
*****

seq3_ AGCTGACGTTTCTAATGCAACAGCAACATACACTGATGCTGATGGTGAATGACTACAAT 1349
seq4_ AGCTGACGTTTCTAATGCAACAGCAACATACACTGATGCTGATGGTGAATGACTACAAT 1349
seq5_ AGCTGACGTTTCTAATGCAACAGCAACATACACTGATGCTGATGGTGAATGACTACAAT 1378
seq1_ AGCTGACGTTTCTAATGCAACAGCAACATACACTGATGCTGATGGTGAATGACTACAAT 1349
seq2_ AGCTGACGTTTCTAATGCAACAGCAACATACACTGATGCTGATGGTGAATGACTACAAT 1349
*****

seq3_ TGGTTCATACACCACGAAGTATTCAATCGATGCTAACAAACGGCAAGGTAAGTGTGATTC 1409
seq4_ TGGTTCATACACCACGAAGTATTCAATCGATGCTAACAAACGGCAAGGTAAGTGTGATTC 1409
seq5_ TGGTTCATACACCACGAAGTATTCAATCGATGCTAACAAACGGCAAGGTAAGTGTGATTC 1438
seq1_ TGGTTCATACACCACGAAGTATTCAATCGATGCTAACAAACGGCAAGGTAAGTGTGATTC 1409
seq2_ TGGTTCATACACCACGAAGTATTCAATCGATGCTAACAAACGGCAAGGTAAGTGTGATTC 1409
*****

seq3_ TGGAAGTGGTACGGGTAAATATGCGCCGAAAGTAGGGGCTGAAGTATATGTTAGTGTCTAA 1469
seq4_ TGGAAGTGGTACGGGTAAATATGCGCCGAAAGTAGGGGCTGAAGTATATGTTAGTGTCTAA 1469
seq5_ TGGAAGTGGTACGGGTAAATATGCGCCGAAAGTAGGGGCTGAAGTATATGTTAGTGTCTAA 1498
seq1_ TGGAAGTGGTACGGGTAAATATGCGCCGAAAGTAGGGGCTGAAGTATATGTTAGTGTCTAA 1469
seq2_ TGGAAGTGGTACGGGTAAATATGCGCCGAAAGTAGGGGCTGAAGTATATGTTAGTGTCTAA 1469
*****

seq3_ TGGTACTTTAACAACAGATGCAACTAGCGAAGGCACAGTAACAAAAGATCCACTGAAAGC 1529
seq4_ TGGTACTTTAACAACAGATGCAACTAGCGAAGGCACAGTAACAAAAGATCCACTGAAAGC 1529
seq5_ TGGTACTTTAACAACAGATGCAACTAGCGAAGGCACAGTAACAAAAGATCCACTGAAAGC 1558
seq1_ TGGTACTTTAACAACAGATGCAACTAGCGAAGGCACAGTAACAAAAGATCCACTGAAAGC 1529
seq2_ TGGTACTTTAACAACAGATGCAACTAGCGAAGGCACAGTAACAAAAGATCCACTGAAAGC 1529
*****

```

```

seq3_      TCTGGATGAAGCTATCAGCTCCATCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAA 1589
seq4_      TCTGGATGAAGCTATCAGCTCCATCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAA 1589
seq5_      TCTGGATGAAGCTATCAGCTCCATCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAA 1618
seq1_      TCTGGATGAAGCTATCAGCTCCATCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAA 1589
seq2_      TCTGGATGAAGCTATCAGCTCCATCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAA 1589
*****

seq3_      CCGTCTGGATTCCGCAGTCACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCA 1649
seq4_      CCGTCTGGATTCCGCAGTCACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCA 1649
seq5_      CCGTCTGGATTCCGCAGTCACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCA 1678
seq1_      CCGTCTGGATTCCGCAGTCACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCA 1649
seq2_      CCGTCTGGATTCCGCAGTCACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCA 1649
*****

seq3_      GTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAACATGTCGAAAGCGCAGAT 1709
seq4_      GTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAACATGTCGAAAGCGCAGAT 1709
seq5_      GTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAACATGTCGAAAGCGCAGAT 1738
seq1_      GTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAACATGTCGAAAGCGCAGAT 1709
seq2_      GTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAACATGTCGAAAGCGCAGAT 1709
*****

seq3_      CATTACAGCAGGCCGGTAACCTCCGTGCTGGCAAAGCCAACCAGGTACCGCAGCAGGTTCT 1769
seq4_      CATTACAGCAGGCCGGTAACCTCCGTGCTGGCAAAGCCAACCAGGTACCGCAGCAGGTTCT 1769
seq5_      CATTACAGCAGGCCGGTAACCTCCGTGCTGGCAAAGCCAACCAGGTACCGCAGCAGGTTCT 1798
seq1_      CATTACAGCAGGCCGGTAACCTCCGTGCTGGCAAAGCCAACCAGGTACCGCAGCAGGTTCT 1769
seq2_      CATTACAGCAGGCCGGTAACCTCCGTGCTGGCAAAGCCAACCAGGTACCGCAGCAGGTTCT 1769
*****

seq3_      GTCTCTGCTGCAGGGTTAA----- 1788
seq4_      GTCTCTGCTGCAGGGTTAA----- 1788
seq5_      GTCTCTGCTGCAGGGTTAGAATCGAATTNCCCGCGCTGACCGTTGTGCGATTTCTCNTCC 1858
seq1_      GTCTCTGCTGCAGGGTTAG----- 1788
seq2_      GTCTCTACTGCAGGGTTAA----- 1788
*****

seq3_      -----
seq4_      -----
seq5_      TTTGTCCTCGT 1869
seq1_      -----
seq2_      -----

```

Quando analisadas no BLASTx, a sequência de aminoácidos obteve 97% de similaridade com o antígeno H12, estes dados não corroboram com os resultados encontrados pela PCR-RFLP, onde verificamos que esta amostra apresentou um padrão molecular P2, semelhante aos antígenos H2, H30 e H35. Além disto, também encontramos uma mudança de leitura de frame por volta do aminoácido 169. Segue abaixo o resultado obtido no BLASTx:

dbj|BAA85084.1| flagellin [*Escherichia coli*]

gb|AAP13305.1| FliC [*Escherichia coli*]

Length=595

Score = 639 bits (1647), Expect(2) = 0.0

Identities = 410/428 (95%), Positives = 416/428 (97%), Gaps = 1/428 (0%)

Frame = +3

```
Query 531 GAEFVFNVNGSGTIANKAATISDLTAAKMDAATNTITTTNNALTASKALDQLKDGDTVTIK 710
          G FNVNGSGTIANKAATISDLTAAKMDAATNTITTTNNALTASKALDQLKDGDTVTIK
Sbjct 169 GLNGFNVNGSGTIANKAATISDLTAAKMDAATNTITTTNNALTASKALDQLKDGDTVTIK 228

Query 711 ADAAQATATVYTYNASAGNFSFSNVSNNTSAKAGDVAASLLPPAGQTASGVYKAASGEVNF 890
          ADAAQATATVYTYNASAGNFSFSNVSNNTSAKAGDVAASLLPPAGQTASGVYKAASGEVNF
Sbjct 229 ADAAQATATVYTYNASAGNFSFSNVSNNTSAKAGDVAASLLPPAGQTASGVYKAASGEVNF 288

Query 891 DVDANGKITIGGQKAYLTSDGNLTTNDAGGATAATLDGLFKKAGDGQSIGFKKTASVTMG 1070
          DVDANGKITIGGQKAYLTSDGNLTTNDAGGATAATLDGLFKKAGDGQSIGFKKTASVTMG
Sbjct 289 DVDANGKITIGGQKAYLTSDGNLTTNDAGGATAATLDGLFKKAGDGQSIGFKKTASVTMG 348

Query 1071 GTTYNFQTGPDADAATANAGLSFTDTASKDVTLNKVATAYTRAKQLQLTGDTSATMTYKS 1250
          GTTYNF+TG DADAATANAG+SFTDTASK+TVLNKVATA + K + GDTSAT+TYKS
Sbjct 349 GTTYNFKTGADADAATANAGVSVFTDTASKETVLNKVATA-KQKAVAADGDTSATITYKS 407

Query 1251 GVQTYQAVFAAGDGTASAKYADKADVSNATATYTDADGEMTTIGSYTTKYSIDANNGKVT 1430
          GVQTYQAVFAAGDGTASAKYADKADVSNATATYTDADGEMTTIGSYTTKYSIDANNGKVT
Sbjct 408 GVQTYQAVFAAGDGTASAKYADKADVSNATATYTDADGEMTTIGSYTTKYSIDANNGKVT 467

Query 1431 VDSGTGTGKYAPKVGAEVYVSANGTLTTDATSEGTVTKDPLKALDEAISSIDKFRSSLGA 1610
          VDSGTGTGKYAPKVGAEVYVSANGTLTTDATSEGTVTKDPLKALDEAISSIDKFRSSLGA
Sbjct 468 VDSGTGTGKYAPKVGAEVYVSANGTLTTDATSEGTVTKDPLKALDEAISSIDKFRSSLGA 527

Query 1611 IQNRLDSAVTNLNNTTNLSEAQSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQ 1790
          IQNRLDSAVTNLNNTTNLSEAQSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQ
Sbjct 528 IQNRLDSAVTNLNNTTNLSEAQSRIQDADYATEVSNMSKAQIIQQAGNSVLAKANQVPQ 587

Query 1791 QVLSQLQG 1814
          QVLS LQG
Sbjct 588 QVLSLLQG 595
```

Score = 202 bits (513), Expect(2) = 0.0

Identities = 135/172 (78%), Positives = 142/172 (82%), Gaps = 14/172 (8%)

Frame = +1

```
Query 49 NSLSLITQNNINKNQALSSSIERLSSGLRINSKDDAAGQAIANRFTSNIKGLTQAARN 228
          NSLSLITQNNINKNQALSSSIERLSSGLRINSKDDAAGQAIANRFTSNIKGLTQAARN
Sbjct 8 NSLSLITQNNINKNQALSSSIERLSSGLRINSKDDAAGQAIANRFTSNIKGLTQAARN 67

Query 229 ANDGISVAQTTEGALSEINNNLQRIRELTVQASTGTNSDSDLDSIQDEIKS-----PS 387
          ANDGISVAQTTEGALSEINNNLQRIRELTVQASTGTNSDSDLDSIQDEIKS S
Sbjct 68 ANDGISVAQTTEGALSEINNNLQRIRELTVQASTGTNSDSDLDSIQDEIKSRLDEIDRVS 127

Query 388 GRN*PRIWDPVQRRERTGERRFNENSGWCE*RQTITIDLKIDSDTLGLNG 543
          G+ + V ++ + + + N G QTTITIDLKIDSDTLGLNG
Sbjct 128 GQT--QFNGVNVLAKDGJSMKIQVGANDG----QTITIDLKIDSDTLGLNG 172
```

Dentre as 20 amostras de *E. coli* HNT estudadas, cinco tiveram o gene *fliC* inteiramente sequenciado. Os iniciadores internos construídos não foram suficientes para permitir o sequenciamento de todo o fragmento de todas as amostras, uma vez que, para cada iniciador utilizado, uma sequência de aproximadamente 500pb era gerada e o tamanho molecular do gene *fliC* é variável.

Segue abaixo um resumo com todos os resultados obtidos entre todos os métodos utilizados para caracterizar o antígeno H (Tabela 13).

Tabela 13: Resultados das técnicas de PCR-RFLP, do sequenciamento e da sorologia utilizando as amostras de *E. coli* HNT.

Amostras	PCR-RFLP	Sequenciamento	Sorologia	Proposição
1C	-	H21	NR	Novo
2C	H2, H30, H35	-	H2 (1:12800)	H2
3C	H11 e H47	-	H11 (1:12800)	H11
4C	H2, H30, H35	-	NO	-
5C	-	H21	NR	Novo
6C	-	-	NO	Novo
7C	-	-	NR	Novo
8C	H51	H51	NR	Novo
9C	-	H31	NR	Novo
2A	H9 e H14	-	H9 (1:12800)	H9
3A	-	-	NR	Novo
5A	-	-	NR	Novo
7A	-	-	NR	Novo
8A	H12	-	H12 (1:12800)	H12
9A	H12	-	NR	Novo
10A	H2, H30, H35	-	H30 (1:12800)	H30
11A	H2, H30, H35	H12	NO	Novo
13A	H16	-	H16 (1:12800)	H16
14A	-	-	NR	Novo
15A	-	-	NR	Novo

-: Resultado negativo;

NO: Não foi obtido o anti-soro;

NR: Não houve reação com nenhum antígeno padrão;

HNT: Antígeno H não tipável

5- Análise da estrutura secundária da proteína flagelina:

Das cinco amostras sequenciadas, apenas a amostra 1C apresentou toda sequência do gene *fliC* sequenciado, sem a presença de mudanças de frame e/ou stop códons na região interna do gene. Logo, submetemos esta amostra à análise da estrutura secundária (Figuras 15 e 16). A amostra 1C apresentou 47% de similaridade com a estrutura secundária da proteína de *Salmonella* TYPHIMURIUM.

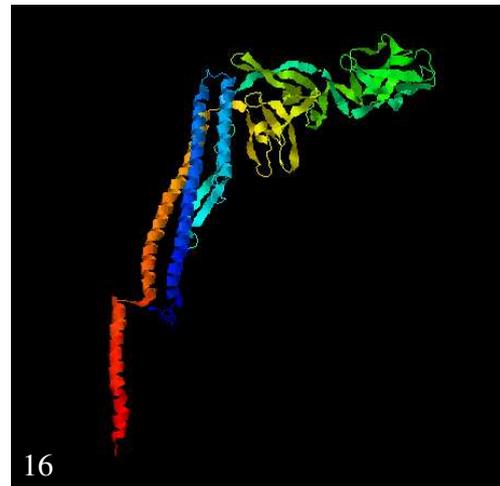
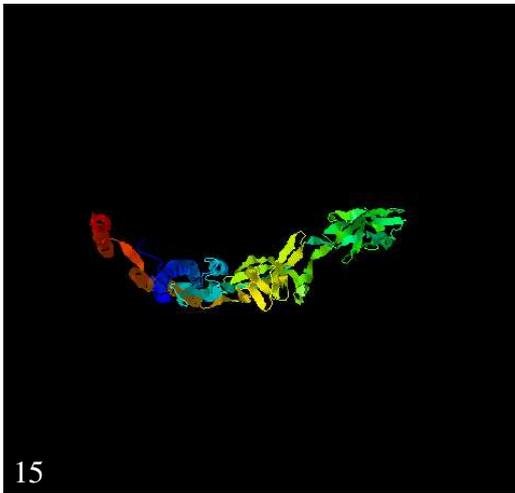


Figura 15: Vista frontal do modelo da estrutura secundária da amostra 1C.

Figura 16: Vista superior do modelo da estrutura secundária da amostra 1C.

V-DISSCUSSÃO

Em 1947, Kauffmann publicou um esquema antigênico, baseando-se nos principais antígenos de superfície da *E. coli*. Este esquema antigênico constituía em 25 grupos de antígenos O, 55 antígenos K e 20 antígenos H. Desde então, a sorologia permitiu um grande conhecimento sobre a epidemiologia, ecologia e patogênese deste organismo (Ewing, 1986; Ørskov & Ørskov, 1992).

A caracterização dos sorotipos é essencial para as investigações epidemiológicas, e na maioria dos casos, a sorotipagem completa de O:K:H, é um bom indicador das relações clonais entre as amostras. Inicialmente era realizado apenas a tipagem do grupo O, mas este não era suficiente para traçar os dados epidemiológicos. Logo, fez-se necessário a utilização da tipagem dos antígenos O e H (Ørskov & Ørskov, 1992).

A combinação específica dos antígenos O e H definem o sorotipo dos membros da família *Enterobacteriaceae*. O sorotipo específico de uma *E. coli* pode permitir a associação desta bactéria com determinadas síndromes clínicas, uma vez que antígenos de superfície podem estar correlacionados com sua virulência, como em EPEC ou EHEC (Amhaz *et al.*, 2004).

O flagelo é a organela responsável pela motilidade na maioria das espécies bacterianas e a unidade principal do filamento é a proteína flagelina codificada pelo gene *fliC*. A flagelina tem um domínio distinto, compreendendo regiões amino e carboxi terminais conservados e uma região central altamente variável, que possuem epítomos específicos (Amhaz *et al.*, 2004). Devido a conservação das extremidades, o gene *fliC* tornou-se um candidato ideal para a amplificação através da PCR, uma vez que estas

regiões conservadas permitem a amplificação de uma variedade de alelos com um único iniciador (Fields *et al.*, 1997), que posteriormente pode ser combinada com o polimorfismo em comprimento dos fragmentos de restrição (RFLP), já que o alvo será a análise da região central variável.

Devido a existência de dados epidemiológicos que se referem à presença dos diferentes grupos de *E. coli* associados à patogenias no homem e animais, observou-se que a sorotipagem é uma importante ferramenta epidemiológica, permitindo o acompanhamento da disseminação das linhagens, sua distribuição nos ambientes e nos hospedeiros, possibilitando a associação do antígeno em estudo com as diferentes doenças. Entretanto, algumas dificuldades foram relatadas, em particular, com a sorotipagem de *E. coli* quando esta é aplicada na rotina laboratorial: 1) a expressão de antígenos H pode ser muitas vezes demorada 2) depender de estímulos ambientais; 3) reações cruzadas podem ocorrer 4) kits diagnósticos para o antígeno H não estão comercialmente disponíveis.

A PCR-RFLP provou ser mais rápida do que a sorotipagem, levando dois dias completos contra sete ou mais dias necessários para a sorotipagem clássica (Prager, *et al.*, 2003; Moreno *et al.*, 2006; Beutin & Strauch, 2007). Por outro lado, estudos sobre a PCR-RFLP do gene *fliC* revelaram que este método poderia ainda ser utilizado para detectar e definir potenciais novas classes de antígenos H, assim como para confirmar resultados sorológicos (Fields *et al.*, 1997; Machado *et al.*, 2000; Prager *et al.*, 2003; Moreno *et al.*, 2006).

A técnica de PCR-RFLP foi utilizada para tipagem do antígeno H e análise filogenética em amostras de *E. coli* invasoras (Amhaz *et al.*, 2004). Em estudos realizados com ETEC, foram observados resultados discrepantes entre a sorologia e a PCR-RFLP, porém o antígeno H foi confirmado através da sorologia por expressar o antígeno esperado

pela PCR-RFLP (Moreno *et al.*, 2006). Prager *et al.* (2003), utilizando ambos os iniciadores (FliC(F) e FliC(M)), obtiveram bons resultados entre a correlação dos tipos de antígenos H identificados através do RFLP e de testes sorológicos, concluindo que ambos os iniciadores poderiam ser utilizados para a complementação dos resultados.

Por outro lado, a técnica de PCR-RFLP também pode ter suas limitações como: 1) o gene *fliC* pode não ser amplificado pela PCR devido a inadequada similaridade dos iniciadores, 2) é descrito na literatura a presença de alelos *fliC* desconhecidos, e os perfis de restrição destes alelos poderiam não coincidir com os perfis padrões já descritos (Moreno *et al.*, 2006).

A enzima *RsaI* foi utilizada por Fields *et al.*, (1997), uma vez que gerava uma variedade de tamanhos de bandas que diferenciavam os antígenos, gerando resultados discriminatórios. Por sua vez, Machado *et al.*, (2000), selecionaram a enzima *HhaI*, que gerava fragmentos que eram similares aos tamanhos dos fragmentos calculados de acordo com os mapas de restrição das sequências de DNA das flagelinas H1, H7 e H12.

Neste trabalho estabelecemos um banco de dados dos padrões de perfis de restrição (P), do gene *fliC* das amostras padrões para antígeno H, utilizando os iniciadores FliC(M) e FliC(F). Com os iniciadores FliC(M), 52 amostras de *E. coli* padrão foram amplificadas pela PCR, enquanto que, com os iniciadores FliC(F), foram amplificadas 50 amostras.

O gene do antígeno H53 não foi amplificado pelos iniciadores FliC(M) e os genes das amostras H17, H25 e H53 não foram amplificadas utilizando os iniciadores FliC(F). Estes resultados corroboram com os resultados encontrados na literatura, visto que, Prager *et al.*, (2003), utilizando os iniciadores FliC(F) não conseguiram obter o produto da PCR com as amostras padrão H17, H25 e H53. Do mesmo modo que Moreno *et al.*, (2006), utilizando o mesmo iniciador, não obtiveram sucesso com as amostras padrão H17, H53 e

H54. Machado *et al.* (2000), não obtiveram resultados satisfatórios com as amostras H51, H52, H53, H55 e H56 utilizando os iniciadores FliC(M).

A ausência da amplificação dos antígenos H17, H25 e H53 pela PCR, corroboram com os resultados presentes na literatura, onde alguns antígenos H são expressos por um sistema gênico diferente de *fliC* ou pela deleção parcial do gene *fliC* (Wang *et al.*, 2003). Por outro lado, as amostras padrões utilizadas por esses autores não são as mesmas, o que permite concluir que podem existir diferenças nos sistemas gênicos de amostras que apresentam o mesmo antígeno H. Em 1998, Ratiner descreveu novos *loci*, *fliA* (alelos de *fliA*₄₄ e *fliA*₅₅, determinando antígenos H44 e H55 em *E. coli*) e *flmA* (alelo *flmA*₅₄ para o antígeno H54). Wang *et al.*, (2003) demonstraram através da clonagem e expressão de *fliC*, que 43 dos 53 antígenos H conhecidos são codificados pelo gene *fliC*, e que os outros 10 estão localizados em outro *locus*. As linhagens H3, H35, H36, H47 e H53 tem sido identificadas no *locus flkA*. As linhagens H44 e H55 no *locus fliA*, e a linhagem H54 no *locus flmA*. Entretanto, muitas destas linhagens expressando muitos destes antígenos H, são conhecidas por também possuir o gene *fliC* (Wang *et al.*, 2003).

Utilizando os iniciadores FliC(M), as amostras padrão para o antígeno H resultaram em produtos de PCR que variaram em tamanho molecular de 925pb à 2635pb. Enquanto que, com os iniciadores FliC(F), foram obtidos produtos de PCR de tamanho molecular variando de 1080pb à 2850 pb.

Todas as amostras com antígeno HNT tiveram o gene *fliC* amplificado com ambos pares de iniciadores, resultando em produtos de PCR que variaram em tamanho molecular de 1180pb à 2605pb com FliC(M), e de 1300 à 2690 pb com FliC(F).

Em nosso trabalho, observamos que algumas amostras padrões quando amplificadas com FliC(M) apresentaram igual perfil de restrição com a enzima *RsaI*. Foram elas: H1,

H28 e H31; H3 e H8; H2, H30 e H35; H9 e H14; H11 e H47; H7, H19 e H27; H55 e H56. Por outro lado, quando o *fliC* das amostras padrões foram amplificadas com os iniciadores FliC(F) e a enzima *HhaI*, observamos perfis de restrições digeridos semelhantes com as amostras H3 e H8; H6, H10, H19 e H27; H11 e H47, H23 e H43; H28 e H42; H55 e H56. É interessante destacar que o gene *fliC* das amostras padrões H3 e H8; H11 e H47; H19 e H27; H55 e H56, apresentaram perfis de restrições semelhantes em ambos os experimentos, mas distintos quando comparados entre si.

De acordo com Prager *et al.*, (2003), o perfil de restrição não necessariamente está diretamente relacionado com antigenicidade expressada, pois diferentes amostras padrão podem apresentar o mesmo perfil de restrição pela técnica de PCR-RFLP, mas são fenotipicamente diferentes através da técnica de sorologia.

Em 2006, foi proposto por Moreno *et al.* que a técnica de PCR-RFLP poderia substituir os métodos clássicos de sorotipagem para a caracterização do antígeno H em amostras de ETEC. Um padrão comum foi observado no gene *fliC* das amostras H1 e H12, das amostras H3 e H16, H5 e H55, H11 e H27. Para diferenciar o padrão de digestão enzimática destas amostras, foi utilizado a endonuclease *HhaI*, e os genes H1, H11, H12 e H27 foram diferenciados nesta análise, enquanto que os perfis de restrição das amostras H3 e H16, H5 e H55 permaneceram indistinguíveis. É importante ressaltar que estes antígenos H podem apresentar reações cruzadas na sorologia, já que possuem epítomos comuns. Moreno *et al.*, (2006), propuseram que a análise de PCR-RFLP fosse utilizada apenas para determinar os perfis de restrição dos antígenos H das amostras de ETEC e compará-los com os perfis de restrição dos antígenos padrões, uma vez que a maioria dos antígenos H podem ser diferenciados pela PCR-RFLP.

Botelho *et al.*, (2003), identificaram que os sorotipos O128:H2 e O128:H35 (EPEC atípica) apresentaram o mesmo perfil de restrição através da análise da PCR-RFLP. Além disso, foi observado que os antígenos H2 e H35 diferiram somente pela substituição de sete aminoácidos ou 1,4% de toda a sequência de 498 aminoácidos. Isto sugere que as diferenças sorológicas entre os antígenos H decorrem de pequenas alterações na sequência protéica. Fields *et al.*, (1997) já tinham observado a ocorrência de um padrão comum do perfil de restrição entre os antígenos H2 e H35.

Machado *et al.*, (2000), demonstraram que as amostras H12 e H45 possuíam o mesmo perfil de restrição, entretanto estes antígenos são conhecidos por possuir reação cruzada na sorologia. Em nove tipos diferentes de amostra do mesmo antígeno H, mais de um padrão de digestão foi observado. Em nosso estudo, apenas uma amostra de cada antígeno H foi utilizada, logo, não foi possível verificar os diferentes perfis de restrição para o mesmo antígeno H. Prager *et al.*, (2003), também verificaram que para o mesmo antígeno H foram observados diferentes padrões de digestões, e vice-versa, ou seja, que o mesmo padrão de digestão enzimática foi observado em diferentes antígenos H. Contudo, a identificação dos padrões de digestão enzimática usando tanto os iniciadores FliC(F) quanto FliC(M), demonstraram uma boa correlação com os resultados sorológicos dos antígenos H.

Neste trabalho, 35 amostras padrão apresentaram o mesmo perfil de restrição daqueles encontrados por Moreno *et al* (2006) e em 30 amostras foram encontradas o mesmo perfil de restrição quando comparados com os resultados obtidos por Prager *et al* (2003). Destas 30 amostras, 25 amostras de *E. coli* apresentaram o mesmo resultado entre as duas metodologias empregadas, baseadas nos protocolos de Fields *et al* (1997) e Machado *et al* (2000).

Quando analisamos o resultado do padrão molecular (P) gerado pela técnica de PCR-RFLP nas amostras de *E. coli* HNT, observamos que das 20 amostras estudadas, 10 amostras (50%) apresentaram um padrão molecular semelhante àqueles encontrados no banco de dados das amostras de H padrão e em cinco amostras (2C, 3C, 4C, 8A e 13A) encontramos o mesmo padrão molecular quando utilizado tanto a técnica proposta por Fields *et al* (1997) quanto por Machado *et al* (2000). As amostras que apresentaram similaridade com dois ou mais antígenos, tiveram seus antígenos H confirmados posteriormente, utilizando técnicas de sorotipagem após a produção do anti-soro ou pela técnica de sequenciamento.

A técnica de PCR-RFLP se baseia na obtenção de um fragmento de DNA que é posteriormente submetido à endonucleases de restrição. As endonucleases são enzimas bacterianas que atuam reconhecendo sequências de pares de bases específicas em moléculas de DNA e cortando-as nesses pontos. Elas são altamente específicas, logo cada tipo de enzima reconhece e corta apenas uma determinada sequência de nucleotídeos, em geral constituída por 4 ou 6 pares de bases nitrogenadas. Portanto apesar de alguns antígenos H gerarem os mesmos perfis de restrição não necessariamente terão a mesma sequência nucleotídica.

Prager *et al* (2003), fizeram a tipagem molecular do antígeno H de 36 amostras imóveis de *E. coli* e sete amostras de *E. coli* HNT. Observaram que nove amostras de *E. coli* imóveis poderiam apresentar um novo antígeno H, e em duas amostras de *E. coli* HNT foi proposto representar um novo antígeno, já que o perfil de restrição obtido não foi correspondente aos perfis de restrição dos 53 antígenos padrões.

Em muitos casos, a PCR-RFLP pode ser incapaz de discriminar os antígenos H. Isto se deve a uma escolha errada das endonucleases, ou a falta de discriminação devido à

excessiva similaridade genética. A discriminação entre os antígenos, que é essencial na epidemiologia, poderia não ser executada pela análise do RFLP, uma vez que, o gene para flagelina pode diferir em poucos nucleotídeos e um único sítio de mutação pode potencialmente resultar na geração de novos antígenos H (Schoenhals & Whitfield, 1993; Dauga *et al.*, 1998). Assim, estes autores propuseram que estes resultados poderiam ser confirmados pela produção do anti-soro e testes de aglutinação.

Nos ensaios de PCR-RFLP utilizando FliC(F) e a enzima *RsaI*, as amostras 2C, 4C, 10A e 11A, apresentaram o padrão molecular (P2), padrão esse observado para os antígenos H2, H30 e H35. Nos ensaios de PCR-RFLP utilizando FliC(M) e a enzima *HhaI*, as amostras 2C e 4C apresentaram perfil P2, já as amostras 10A e 11A apresentaram distintos perfis de restrição.

Reações sorológicas do antígeno H de *E. coli* são utilizadas para distinguir entre os diferentes sorotipos (antígenos O e H) e associá-los às doenças. O teste de aglutinação ainda é rotineiramente utilizado em estudos epidemiológicos e na associação destes antígenos à patogenicidade para homens e animais.

Após a obtenção dos anti-soros contra o antígeno H das amostras HNT, estes foram titulados frente seus antígenos homólogos, frente aos 53 antígenos padrões e frente aos 20 antígenos HNT estudados. Dos anti-soros produzidos apenas o anti-soro da amostra 7A obteve título de 1:3200, os demais anti-soros apresentaram título acima de 1:12800.

As amostras 8A e 13A apresentaram perfis de restrição semelhantes aos dos antígenos H12 e H16 respectivamente, em ambas as técnicas de PCR-RFLP. Estes resultados foram confirmados após a produção dos anti-soros destas amostras e titulação frente aos antígenos H12 e H16. A única amostra que teve seu anti-soro positivo frente aos antígenos padrões (H11), mas não obteve nenhum resultado na PCR-RFLP foi a amostra

1C. É importante observar que os antígenos H podem apresentar reações cruzadas durante a sorotipagem clássica porque eles podem possuir epítomos comuns, sendo comum a reação cruzada entre os antígenos H11 e H21 (Ewing, 1986).

O antígeno H da amostra 10A que obteve o padrão molecular P2 (Fields), foi confirmado através da sorologia, na qual o antígeno H30 reagiu com o soro anti-10A em título acima de 1:12800.

A amostra 3C apresentou o mesmo perfil de restrição das amostras padrão H11 e H47 quando utilizado ambas as técnicas de PCR-RFLP, porém quando foi obtido seu anti-soro, este apresentou um resultado positivo frente ao antígeno H11 com título acima de 1:12800. O resultado da amostra 2A também foi confirmado após a produção do seu anti-soro e titulação frente ao antígeno H9 padrão. Através da PCR-RFLP foi obtido um perfil de restrição semelhante às amostras H9 e H14 quando utilizado os iniciadores FliC(F) e a enzima *RsaI*, sendo que não foi encontrado padrão semelhante quando utilizados FliC(M). Através da sorologia, o anti-soro da amostra 2A reagiu com o antígeno H9, com título acima de 1:12800.

Estes resultados indicam que apesar de determinados antígenos apresentarem perfis de restrição semelhantes, eles podem possuir uma sequência nucleotídica diferente, que pode ainda gerar uma proteína diferente.

Foi observado que podem existir diferenças fenotípicas e genotípicas nestas amostras HNT em relação aos resultados dos antígenos padrão, uma vez que, somente o anti-soro das amostras HNT reagiu com os antígenos padrão. A proteína flagelina das amostras HNT pode ser estruturalmente diferente do antígeno padrão, mas pode apresentar epítomos comuns com aquele do antígeno H padrão. É relatado na literatura que um mesmo

antígeno H pode ter sequências genótípicas diferentes, apesar de gerarem os mesmos resultados fenotípicos (Wang *et al.*, 2000; Johnson & Stell, 2001; Beutin & Strauch, 2007).

Para caracterizar o antígeno H das amostras HNT que apresentaram perfis de restrição diferentes aos das amostras padrão, e verificar as possíveis diferenças nas sequências nucleotídica, executamos o sequenciamento do gene *fliC* das amostras HNT. Foi observado que algumas amostras possuem o mesmo perfil de restrição porém não necessariamente possuem a mesma sequência nucleotídica.

Como o fragmento da PCR do gene *fliC* pode variar de 1.8 à 2.6 kbp, foi necessário a construção de iniciadores internos, uma vez que as sequências obtidas por cada um dos iniciadores, estando orientado no sentido *forward* ou *reverse*, normalmente fornecem de 500 a 600 pares de bases (pb) dependendo da qualidade do sequenciamento. Além disso, precisávamos obter a sequência completa do gene *fliC*, ou seja, também tínhamos que ter as extremidades sequenciadas. Logo, baseando-se na literatura, uma alternativa para o sequenciamento do gene *fliC* foi utilizada paralelamente. Os fragmentos da PCR das amostras de *E. coli* HNT, amplificados pelos iniciadores propostos por Fields *et al* (1997), foram inseridos no vetor de clonagem pGEM –T easy.

Os plasmídios foram sequenciados, utilizando inicialmente os iniciadores T7 e M13. Cada reação de sequenciamento foi realizada pelo menos em triplicata. Depois da obtenção de cada *contig* da sequência do gene *fliC*, através do programa Phred/Phrap/ Consed, fez-se o alinhamento de cada sequência com as outras depositadas no Genbank (BLASTn) e cada base duvidosa foi checada em todos os eletroesferogramas de cada réplica, garantindo assim a confiabilidade nos resultados obtidos.

As sequências nucleotídicas das amostras 1C e 5C apresentaram similaridade com o antígeno padrão H21, mas, não houve a identificação destas amostras com nenhum perfil de restrição dos antígenos padrões, através da técnica de PCR-RFLP. Porém, quando analisamos nossos resultados com os resultados já descritos na literatura, observamos que Prager *et al* (2003), utilizando os iniciadores fliC(M) e a endonuclease *HhaI*, obteve o mesmo perfil de restrição entre as amostras H21 e a amostra 1C. O resultado do sequenciamento confirmou que a sequência de aminoácidos da amostra 1C apresenta 99% de similaridade com o antígeno H21, sendo que um aminoácido isoleucina foi trocado por uma treonina.

Esta troca de aminoácidos é chamada de mutação não sinônima, sendo um tipo de mutação pontual onde ocorre a mudança de um único nucleotídeo, provocando a substituição de um aminoácido. Esta troca pode alterar a proteína ou torná-la não-funcional (Griffiths *et al.*, 2006). Podemos observar nos resultados obtidos pelo programa ClustalW que houve uma mudança de uma base de citosina por uma timina na região de -1415. Isto poderia explicar o porquê desta amostra não ser caracterizada pela sorologia, uma vez que seu antígeno pode ser diferente dos antígenos padrões já descritos, e a presença de mutações podem levar a expressão de um fenótipo diferente.

Já o resultado obtido na sorologia, onde o anti-soro da amostra 1C reagiu com o antígeno H11 com título acima de 1:6400 poderia ser explicado por reações cruzadas, uma vez que ambos os antígenos estão agrupados dentro do mesmo morfotipo D (Lawn *et al.*, 1977) e é constatado na literatura a necessidade da técnica de soroabsorção, uma vez que há reações cruzadas entre os antígenos H11 e H21 (Edwards & Ewing, 1986). O antígeno H desta amostra possivelmente possui um novo antígeno H, uma vez que, foi detectada uma

mutação e seu anti-soro não foi positivo para o teste de aglutinação com o antígeno padrão H21.

A sequência de aminoácidos da amostra 1C foi submetida ao endereço eletrônico de bioinformática estrutural: <http://www.sbg.bio.ic.ac.uk/phyre>. O resultado da amostra 1C apresentou 47% de similaridade com a estrutura secundária protéica da flagelina de fase 1 de *Salmonella thyphimurium*. Quando analisamos a sequência de aminoácidos do antígeno flagelar H11, verificou-se 43% de similaridade com a proteína de flagelina de fase 1 de *Salmonella thyphimurium* e 51% de similaridade do antígeno flagelar H21.

Já a amostra 5C, apesar de apresentar uma sequência de aminoácidos homóloga ao antígeno H21 (96%), apresentou um stop códon no aminoácido 341. A presença de um stop códon na região codante da proteína acaba resultando em uma proteína não funcional. Este stop códon pode ter sido gerado pela inserção de bases (C) ou troca de nucleotídeos (G) verificados na região do nucleotídeo -1038. Nesta mesma região foi verificada ainda a presença de mutações sinônimas, onde ocorre à mudança de uma base nucleotídica por outra, que acaba gerando o mesmo aminoácido, ou até por outros aminoácidos, que possuem as mesmas características bioquímicas e moleculares, que acabam não alterando a proteína. Também foram encontradas mutações não-sinônimas nesta mesma região.

Nos resultados sorológicos da amostra 5C verificamos que o anti-soro obtido reagiu com título acima de 1:12800 frente ao seu antígeno homólogo e frente ao antígeno da amostra 1C, e não reconheceu os antígenos H padrões, em especial o antígeno H21.

A reação cruzada do soro anti-5C e o antígeno 1C era esperado tendo em vista as semelhanças observadas nas sequências nucleotídicas desses antígenos. Assim, considerando o conjunto dos resultados propomos que o antígeno H da amostra 5C seja considerado um provável novo antígeno H.

O resultado da PCR-RFLP utilizando os iniciadores FliC(M) e a endonuclease *HhaI*, identificou que o gene *fliC* da amostra 8C possuía o mesmo perfil de restrição da amostra padrão H51. Este resultado foi confirmado pelo sequenciamento que obteve uma sequência de aminoácidos com 99% de similaridade com o antígeno H51. Através do BLASTx, observou-se que esta amostra possuía uma mudança de leitura de frame no aminoácido 588. A mudança de leitura de frame pode ser ocasionada pela inserção ou deleção de bases. Através do programa ClustalW, verificamos a presença da inserção de duas bases nucleotídicas (GT) na posição de -1784, que produziram a mudança na leitura de frame.

Reid *et al* (1999), estudando a sequência nucleotídica de amostras de *E. coli* O157:H-, verificaram a presença da inserção de duas bases na região conservada 5', quando comparada com a sequência nucleotídica da amostra de *E. coli* O157:H7. Estas inserções produziam uma mudança de frame que introduzia um stop códon prematuro. Além disso, estes autores também descreveram que os alelos *fliC* do antígeno H7 diferem de quatro a seis aminoácidos entre eles, o gene *fliC* da amostra O157:H7 difere da amostra O128:H7 em 11 nucleotídeos, os alelos do antígeno H6 diferem em dois aminoácidos entre eles, o antígeno H2 difere em dois aminoácidos do antígeno H21 e o antígeno H1 difere em seis aminoácidos em relação aos antígenos H2, H12 e H21.

O anti-soro 8C não foi positivo para o teste de aglutinação com o antígeno padrão H51, mas reconheceu o antígeno da amostra 9C e vice-versa. Pode haver alguma alteração na proteína final decorrente da alteração e/ou presença do stop códon e da mudança de frame, que expõe os epítomos que estariam mascarados no antígeno “selvagem” e explicariam as reações cruzadas entre as amostras HNT e as semelhanças das sequências nucleotídicas.

A amostra 9C não apresentou similaridade com nenhum perfil de restrição através da PCR-RFLP, porém os resultados do sequenciamento identificaram que esta amostra tinha 99% de similaridade com o antígeno H31. Uma mudança de frame no aminoácido 513 foi encontrada. Esta mudança de frame foi ocasionada pela inserção de uma base nucleotídica (C) na posição do nucleotídeo -1540. O anti-soro desta amostra obteve um título acima de 1:12800 frente ao seu antígeno homólogo, porém foi negativo para os antígenos padrões, em particular para o antígeno H31. Este anti-soro apresentou reações cruzadas frente ao antígeno da amostra 8C. Este resultado era esperado, uma vez que o anti-soro da amostra 8C havia apresentado reação cruzada com o antígeno 9C.

De acordo com Lawn *et al.*, (1977) os flagelos H51 e H31, apresentaram diferenças quanto a sua morfologia, o antígeno H51 faz parte do morfotipo E, e o antígeno H31 do morfotipo F. As reações cruzadas observadas entre os antígenos 8C e 9C com os seus respectivos soros nos permite inferir que as proteínas apresentam epítomos comuns na forma quaternária do flagelo, por outro lado, não há semelhança entre as sequências nucleotídicas de 8C e 9C.

O sequenciamento do gene *fliC* da amostra 11A, e posterior análise demonstrou a existência de 99% de similaridade com o antígeno padrão H12. Estes dados não corroboram com os resultados obtidos pela PCR-RFLP, visto que a amostra 11A apresentou padrão molecular P2 (H2, H30 e H35). Além da mudança de frame por volta do aminoácido 169, que foi gerada pela deleção de uma base nucleotídica (C) no nucleotídeo -488, foram encontradas várias mutações sinônimas e mutações não-sinônimas. Também foi verificada a presença de stop códon no aa 132 e 153 que foi gerado pela inserção de uma base nucleotídica (C) na posição -393. Até o momento, não consta na literatura semelhanças entre os perfis de restrição e/ou sequenciamento entre as amostras padrão H2 e H12.

Morfologicamente, a amostra H2 faz parte do morfotipo C enquanto a amostra padrão H12 está agrupada no morfotipo E (Lawn *et al.*, 1977).

A sequência de aminoácidos da flagelina entre as várias linhagens bacterianas demonstrou dois domínios extremamente conservados que correspondem aos 170 resíduos da porção amino-terminal e aproximadamente 90 resíduos da porção carboxi-terminal. Aproximadamente uma dúzia de sítios de mutações pontuais tem sido identificada na linhagem do tipo selvagem de *S. typhimurium* SJW1103 e estão associadas à morfologia do filamento, e todas, exceto uma, são encontradas nas regiões conservadas (Samatey *et al.*, 2001). Os mutantes polimórficos possuem filamentos flagelares de várias formas, como: filamento reto, *curly*, semi-enrolado e enrolado. A região carboxi-terminal tem uma alta concentração de sítios de mutações do que outra região. Isto foi observado pela forte correlação entre os tipos polimórficos dos filamentos e os domínios dos sítios de mutação.

Neste trabalho, utilizamos amostras de *E. coli* de origem humana provenientes de vários casos clínicos e, através dos métodos utilizados, foi observado que estas amostras apresentaram os antígenos H2, H9, H11, H12, H16, H21, H30, H31, H51.

Os antígenos H2, H9, H12, H16, H21 e H30 estão associados aos sorotipos de *E. coli* isolados de humanos com infecções intestinais. Os antígenos H2, H9, H12, H21 são alguns dos principais sorotipos associados à EPEC e ocasionam surtos epidêmicos e esporádicos de diarreia infantil no mundo todo (Edwards & Ewing, 1986; Campos *et al.*, 2004).

Os antígenos H9, H11, H12, H16, estão associados à ETEC que ocasionam diarreia em adultos e crianças, sendo mais conhecida como a diarreia do viajante. O antígeno H11 também é encontrado em EHEC do sorotipo O26:H11, que frequentemente ocasionam colite hemorrágica.

Os antígenos H2 e H30 estão associados à EIEC que ocasiona disenteria em crianças e adultos, uma vez que a bactéria penetra nas células epiteliais na superfície da lâmina própria e produz pequenas úlceras acompanhadas por uma reação inflamatória. Este grupo possui uma pequena lista de sorotipos associados O:H, já que a maioria das amostras são imóveis. Já o antígeno H31 está associado às infecções extraintestinais, sendo encontrado em bacteremias (Ørskov & Ørskov, 1992).

Das vinte amostras estudadas, seis amostras (5C, 9C, 3A, 5A, 7A, 15A), não tiveram o antígeno H caracterizado pela técnica de PCR-RFLP e os anti-soros destas amostras também não reagiram com nenhum dos 53 antígenos padrões. Apesar destas amostras terem sido sequenciadas, os resultados obtidos do sequenciamento não foram suficientes para caracterizar o antígeno H, visto que o sequenciamento destas amostras apresentou uma baixa qualidade na leitura de sequenciamento e, devido ao tamanho do gene *fliC* e da variabilidade genética da sequência de *fliC*, foram necessários a seleção de vários iniciadores internos para sequenciar todo o gene *fliC*. Devido ao fato de termos somente as regiões conservadas destas amostras, não houve como analisá-las nos programas utilizados neste trabalho, uma vez que, a ausência da região central, que confere variabilidade genética ao gene *fliC*, geraria falta de confiabilidade nos resultados.

Para as amostras 4C, 6C, e 11A, não conseguimos obter seus anti-soros e o soro das amostras 2C, 7C, 14A não reagiram frente aos seus antígenos homólogos, sugerindo que estudos complementares com tais amostras sejam realizados.

Em 1956, Ewing descreveu oito possíveis novos grupos O em *E. coli* isoladas de diarréia infantil, que foram designados de OX1 à OX8. Somente em 2004, Scheutz *et al.*, publicaram um estudo utilizando métodos de sorotipagem, estudos epidemiológicos e hibridização de genes de fatores de virulência associados à *E. coli* e propuseram que estes

antígenos fossem designados como O174 à O181. O último relato de identificação de um novo antígeno H, foi feito em 1975 por Ørskov *et al.*, que identificou o antígeno H56.

Finalmente, para comprovação, aceitação e reconhecimento dos resultados obtidos neste trabalho, será necessário submeter estas amostras ao Centro de Referência Internacional em *Escherichia* e *Klebsiella* (The International *Escherichia* and *Klebsiella* Centre), em Copenhague, Dinamarca.

CONCLUSÕES

- Os protocolos de PCR-RFLP, propostos por Fields e col (1997) e Machado e col, (2000), permitiram a construção de um banco de dados dos perfis de restrição dos antígenos flagelares padrões, assim como das amostras de *E. coli* HNT utilizadas neste trabalho;
- Através da identificação computacional dos perfis de restrição presentes no banco de dados, 10 amostras de *E. coli* HNT tiveram o antígeno H caracterizado.
- As amostras de *E. coli* HNT apresentaram apenas o operon flagelar *fliC*, e não amplificaram para os genes associados à variação de fase *fljA*, *flkA* e *flmA*.
- Dez amostras (1C, 5C, 6C, 7C, 9C, 3A, 5A, 7A, 14A, 15A) não apresentaram similaridade com nenhum perfil de restrição presente no banco de dados.
- Oito amostras (5C, 7C, 9C, 3A, 5A, 7A, 14A, 15A), não apresentaram similaridade com nenhum perfil de restrição presente no banco de dados, assim como, os anti-soros destas amostras, também não reagiram com os antígenos flagelares padrões, agrupando estas amostras como possíveis novos antígenos.
- Seis amostras (1C, 5C, 8C, 9C, 9A e 11A) possivelmente possuem um novo antígeno H, uma vez que, os resultados dos testes moleculares, não corroboraram com os resultados já descritos na literatura, e com os resultados fenotípicos.
- Treze amostras (1C, 5C, 6C, 7C, 8C, 9C, 3A, 5A, 7A, 9A, 11A, 14A, 15A) foram caracterizadas como sendo um provável novo antígeno H.
- A obtenção dos anti-soros das amostras HNT, demonstraram que seis soros (2C, 3C, 2A, 8A, 10A, 13A) reagiram com os antígenos padrões H2, H9, H11 H12, H16, H30, que estão associados à *E. coli* patogênica.

VII-REFERÊNCIAS BIBLIOGRÁFICAS

REFERÊNCIAS BIBLIOGRÁFICAS

ACHTMAN, M; HEUZENROEDER, M; KUSECEK, B. Clonal analysis of *Escherichia coli* O2:K1 isolated from diseased humans and animals. **Infect. Immun.**, v. 51, p. 268-276, 1986.

AMHAZ, J. M. K; ANDRADE, A; BANDO, S. Y; TANAKA, T. L; MOREIRA-FILHA C. A; MARTINEZ, M. B. Molecular typing and phylogenetic analysis of enteroinvasive *Escherichia coli* using the *fliC* gene sequence. **FEMS Microbiol. Lett.**, v. 235, p. 259-264, 2004.

AUSUBEL, F.M; BRENT, R.; KINGSTON, R.E.; MOORE, D.D; SEIDMAN, J.D.; SMITH, J.A.; STRUHL, K. (Ed.). *Escherichia coli*, Plasmids and Bacteriophages. **Curr. Protoc. Mol. Biol.**, New York: Greene-Wiley, 1993. v.1, p. 1-6.

BALDY-CHUDZIK, K & STOSIK, M. Diversity of *fliC* gene in commensal *Escherichia coli* derived from various mammals. **Folia Microbiol.**, v. 52 (3), p. 261-272, 2007.

BERG, H. C. The rotary motor of bacterial flagella. **Annu. Rev. Biochem.**, v. 72, p. 19-54, 2003.

BEUTIN, L & STRAUCH, E. Identification of sequence diversity in the *Escherichia coli fliC* genes encoding flagellar types H8 and H40 and its use in typing of Shiga toxin-producing *E. coli* O8, O22, O111, O174, and O179 strains. **J. Clin. Microbiol.**, v. 45(2), p. 333-339, 2007

BEUTIN, L; STRAUCH, E; ZIMMERMANN, S; KAULFUSS, S; SCHAUDINN, C; MANNEL, A; GELDERBLUM, H. R. Genetical and functional investigation of *fliC* genes

encoding flagellar serotype H4 in wildtype strains of *Escherichia coli* and in a laboratory *E. coli* K-12 expressing flagellar antigen type H48. **BMC Microbiol.**, v. 5 (4), 2005.

BLANCO M, BLANCO J, BLANCO J. E, GONZÁLEZ E. A, GOMES T. A, ZERBINI L. F, YANO T, DE CASTRO A. F. Genes coding for Shiga-like toxins in bovine verotoxin-producing *Escherichia coli* (VTEC) strains belonging to different O:K:H serotypes. **Vet. Microbiol.**, v. 42, (2-3), p. 105-10, 1994.

BOTELHO, B. A; BANDO, S. Y; TRABULSI, L. R, MOREIRA-FILHO, C. A. Identification of EPEC and non-EPEC serotypes in the EPEC O serogroups by PCR-RFLP analysis of the *fliC* gene. **J. Microbiol. Methods.**, v. 54, p. 87-93, 2003.

BRUTINEL, E. D & YAHR, T. L. Control of gene expression by type III secretory activity. **Curr. Opin. Microbiol.**, v. 11 (2), p. 128-133, 2008.

CAMPOS, L. C; FRANZOLIN, M. R; TRABULSI, L. R. Diarrheagenic *Escherichia coli* categories among the traditional enteropathogenic *E. coli* O serogroups--a review. **Mem. Inst. Oswaldo Cruz**, v. 6, p. 545-552, 2004.

CHEVANCE, F. F. V & HUGHES, K. Coordinating assembly of a bacterial macromolecular machine. **Nature**, v. 6, p. 455-465, 2008.

DAUGA, C; ZABROVSKAIA, A; GRIMONT, P. A. D. Restriction fragment length polymorphism analysis of some flagellin genes of *Salmonella entérica*. **J. Clin. Microbiol.**, v. 36 (10), p. 2835-2846, 1998.

DePAMPHILIS, M. L & ADLER, J. Purification of intact flagella from *Escherichia coli* and *Bacillus subtilis*. **J. Bacteriol.**, v. 105 (1), p. 376-383, 1971.

DOBBIN, H. S; HOVDE, C. J; WILLIAMS, C. J; MINNICH, S. A. The *Escherichia coli* O157 flagellar regulatory gene *flhC* and not the flagellin gene *fliC* impacts colonization of cattle. **Infect Immun**, v. 74 (5), p. 1734 - 1749, 2006.

DURSO, L. M; BONO, J. L; KEEN, J. E. Molecular serotyping of *Escherichia coli* O26:H11. **Appl. Environ. Microbiol.**, v. 71 (8), p. 4941 – 4944, 2005.

EWING, W.H; TATUM, H. W; DAVIS, B. R; REAVIS, R. W. Studies on the serology of the *Escherichia coli* group. Communicable Disease Center Publication, Atlanta, Georgia, p. 1-42, 1956.

EWING, W.H. The genus *Escherichia*. In: EDWARDS, P. R; EWING, W. H. Ewing's Identification of *Enterobacteriaceae*, 4th Ed. Minneapolis, Burgess, p. 93-134, 1986.

FENG, L; LIU, B; LIU, Y; RATINER, Y. A; HU, B; LI, D; ZONG, X; XIONG, W; WANG, L. A genomic islet mediates flagellar phase variation in *Escherichia coli* strains carrying the flagellin-specifying locus *flk*. **J. Bacteriol.**, v. 190 (13), 2008.

FÉRIA, C; MACHADO, J; CORREIA, J. D; GONÇALVES, J; GAASTRA, W. Distribution of *papG* alleles among uropathogenic *Escherichia coli* isolated from different species. **FEMS Microbiol. Lett.**, v. 202, p. 205-208, 2001.

FIELDS, P.I; BLOM, K; HUGHES, H.J; HELSEL, L.O; FENG, P; SWAMINATHAN, B. Molecular characterization of the gene encoding H antigen in *Escherichia coli* and development of a PCR-Restriction fragment length polymorphism test for identification of *E. coli* O157:H7 and O157:NM. **J. Clin. Microbiol.**, v. 35, p. 1066-1070, 1997.

GRIFFITHS, A. J. F; WESSLER, S. R; LEWONTIN, R. C; GELBART, W. M. SUZUKI, D. T. Introdução à Genética, 8^a ed, Guanabara Koogan, 2006.

JOHNSON, J. R & RUSSO, T. A. Extraintestinal pathogenic *Escherichia coli*: “The other bad *E. coli*”. **J. Lab. Clin. Med.**, v. 139 (3), p. 155-160, 2002.

JOHNSON, J. R & STELL. PCR for specific detection of H7 flagellar variant of *fliC* among Extraintestinal Pathogenic *Escherichia coli*. **J. Clin. Microbiol.**, v. 39 (10), p. 3712-3717, 2001.

JOYS, T. M & SCHODEL, F. Epitope mapping of the d flagellar antigen of *Salmonella muenchen*. **Infect. Immun.**, v. 59 (9), p. 3330-3332, 1991.

KAPER, J.B; NATARO, J.P; MOBLEY, H.L.T. Pathogenic *Escherichia coli*. **Nature**, v2. p 123-140, 2004.

KAUFFMANN, F. The serology of the *Coli* group. **J. Immunol.**, v. 57, p. 71-100, 1947.

KOSTRZYŃKA, M; BETTS, J.D; AUSTIN, J.W; TRUST, T. J. Identification, characterization and spatial localization of two flagellin species in *Helicobacter pylori* flagella. **J. Bacteriol.**, v. 173 (3), p. 937-946, 1991.

KUWAJIMA, G.; ASAKA, J. I; FUJIWARA, T; FUJIWARA, T; NODE, K; KONDO. E. Nucleotide sequence of the *hag* gene encoding flagellin of *Escherichia coli*. **J. Bacteriol.**, v. 168 (3), p. 1479-1483, 1986.

LAWN, A. M. Comparison of the flagellins form different flagellar morphotypes of *Escherichia coli*. **J. Gen. Microbiol.**, v. 101, p. 112-130, 1977.

LAWN, A. M; ORSKOV, I; ORKOV.F. Morphological distinction between different H serotypes of *Escherichia coli*. **J. Gen. Microbiol.**, v. 101, p. 11-119, 1977.

LINO, T. Genetics of structure and function of bacterial flagella. **Ann. Rev. Genet.**, v. 11, p. 61-82, 1977.

LINO, T; KOMEDA, Y; KUTSUKAKE, K; MACNAB, R. M; MATSUMURA, P; PARKINSON, J. S; SIMON, M. I; YAMAGUCHI, S. New unified nomenclature for the flagellar genes of *Escherichia coli* and *Salmonella typhimurium*. **Microbiol. Rev.**, v. 52 (4), 1988.

LOGAN, S. M; HARRIS, L. A; TRUST, T.J. Isolation and Characterization of *Campylobacter* flagellins. **J. Bacteriol.**, v. 169 (11), p. 5072-5077, 1987.

- MACHADO, J; GRIMONT, F; GRIMONT, P. A. D. Identification of *Escherichia coli* flagellar types by restriction of the amplified *fliC* gene. **Res. Microbiol.**, v. 151, p. 535-546, 2000.
- MACNAB, R. M. Flagella. In *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular biology, 1987.
- MACNAB, R. M. Genetics and Biogenesis of bacterial flagella. **Annu. Rev. Genet.**, v. 26, p. 131-158, 1992.
- MACNAB, R. M. How bacteria assemble flagella. **Annu. Rev. Microbiol.**, v. 57, p. 77-100, 2003.
- MORENO, A.C.R; GUTH, B. E. C; MARTINEZ, M.B. Can the *fliC* PCR-Restriction fragment length polymorphism technique replace classic serotyping methods for characterizing the H antigen of Enterotoxigenic *Escherichia coli* strains? **J. Clin. Microbiol.**, v. 44, p. 1453-1458, 2006.
- MYTELKA, D. S & CHAMBERLIN, M. J. *Escherichia coli* *fliAZY* operon. **J. Bacteriol.**, v. 178 (1), p. 24-34, 1996.
- NATARO, J. P & KAPER, J. B. Diarrheagenic *Escherichia coli*. **Clin. Microbiol. Rev.**, v. 11, p. 142-201, 1998.
- ØRSKOV, F. & ØRSKOV, I. Two new *Escherichia coli* O antigens, O162 and O 163, and one new H antigen, H56. Withdrawal of H antigen H50. **Acta Path. Microbiol. Scand.**, Sect (83), p. 121-124, 1975.
- ØRSKOV, I. & ØRSKOV, F. *Escherichia coli* in extra-intestinal infections. **J. Higiene.** Cambridge, v. 95, p. 551-575, 1985.
- ØRSKOV, I & ØRSKOV, F. *Escherichia coli* serotyping and in man and animals. **Can. J. Microbiol.**, v. 38, p. 699-704, 1992.

- PRAGER, R; STRUTZ, U; FRUTH, A; TSCHÄPE, H. Subtyping of pathogenic *Escherichia coli* strains using flagellar (H) – antigens: serotyping versus *fliC* polymorphisms. **Int. J. Med. Microbiol.**, v. 292, p. 477-486, 2003.
- RATINER, Y. A. Phase variation of the H antigen in *Escherichia coli* strain Bi7327-41, the standard strain for *Escherichia coli* flagellar antigen H3. **FEMS Microbiol. Lett.**, v. 15, p. 33-36, 1982.
- RATINER, Y. A. Presence of two structural genes determining antigenically different phase-specific flagellins in some *Escherichia coli* strains. **FEMS Microbiol. Lett.**, v. 19, p. 37-41, 1983.
- RATINER, Y. A. Two genetic arrangements determining flagellar antigen specificities in two diphasic in two diphasic *Escherichia coli* strains. **FEMS Microbiol. Lett.**, v. 29, p. 317-323, 1985.
- RATINER, Y. A. Different alleles of the flagelina gene *hagB* in *Escherichia coli* standard H tests strains. **FEMS Microbiol. Lett.**, v. 48, p. 97-104, 1987.
- RATINER, Y. A. New flagellin-specifying genes in some *Escherichia coli* strains. **J. Bacteriol.**, v. 180 (4), p. 979-984, 1998.
- RATINER, Y. A. Temperature-dependent flagellar antigen phase variation in *Escherichia coli*. **Res. Microbiol.**, v. 150, p. 457-463, 1999.
- RATINER, Y. A; SALMENLINNA, S; EKLUND, M; KESHIMAKI, M; SIITONEN, A. Serology and Genetics of the flagellar antigen of *Escherichia coli* O157: H7a, 7c. **J. Clin. Microbiol.**, v. 41 (3), p. 1033-1040, 2003.
- REID, S. D; SELANDER, R. K; WHITTAM, T. S. Sequence diversity of flagellin (*fliC*) alleles in pathogenic *Escherichia coli*. **J. Bacteriol.**, v. 181 (1), p.153-160, 1999.

SAMATEY, F. A; IMADA, K; NAGASHIMA, S; VONDERVISZT, F; KUMASAKA, T; YAMAMOTO, M; NAMBA, K. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. **Nature**, v. 410, p. 331-337, 2001.

SAMBROOK, J; FRITSCH, E. F; MANIATIS, T. Molecular cloning. A laboratory manual. 2ed. CSH, 1989.

SCHEUTZ, F; CHEASTY, T; WOODWARD, D; SMITH HR. Designation of O174 and O175 to temporary O groups OX3 and OX7, and six new *E. coli* O groups that include Verocytotoxin-producing *E. coli* (VTEC): O176, O177, O178, O179, O180 and O181. **APMIS**, v. 112 (9), p. 569-584, 2004.

SCHOENHALS. G & WHITFIELD, C. Comparative analysis of flagellin sequences from *Escherichia coli* strains possessing serologically distinct flagellar filaments with a shared complex surface pattern. **J. Bacteriol.**, v. 175 (17), p. 5395-5402, 1993.

SEAH. J.N & KWANG. J. Identification of H-specific determinants in flagellin of four *Escherichia coli* strains. **Arch Microbiol.**, v.174, p.28-34, 2000.

SILVERMAN, M & SIMON, M. I. Bacterial flagella. **Ann. Rev. Microbiol.**, v. 31, p. 397-419, 1977.

SMITH, N. H & SELANDER, R. K. Sequence invariance of the antigen-coding central region of the phase 1 flagellar filament gene (*fliC*) among strains of *Salmonella typhimurium*. **J. Bacteriol.**, v. 172 (2), p. 603-609, 1990.

STRAUCH, E & BEUTIN, L. Imprecise excision of insertion element *IS5* from the *fliC* gene contributes to flagellar diversity in *Escherichia coli*. **FEMS Microbiol. Lett.**, v. 256, p. 195-206, 2006.

SUSSMAN. M. *Escherichia coli* : Mechanism of virulence. Cambridge University Press, 1ed, 1997.

TOLEDO, M. R. F; FONTES, C. F.; TRABULSI, L. R. Um meio para a realização dos testes de motilidade, indol e lisina descarboxilase. **Rev. Microbiol.**, v. 13, p. 230-235, 1982.

TOMINAGA, A. Characterization of six flagellin genes in the H3, H53 and H54 standard strains of *Escherichia coli*. **Genes Genet. Syst.**, v. 79, p. 1-8, 2004.

TOMINAGA, A & KUTSUKAKE, K. Expressed and cryptic flagellin genes in the H44 and H55 type strains of *Escherichia coli*. **Genes Genet. Syst.**, v. 82, p. 1-8, 2007.

VONDERVISZT, F; KANTO, S; AIZAWA, S. I; NAMBA K. Terminal regions of flagellin are disordered in solution. **J. Mol. Biol.**, v. 209, p. 127-133, 1989.

VONDERVISZT, F; UEDAIRA, H; KIDOKORO, S. I; NAMBA, K. Structural organization of flagellin. **J. Mol. Biol.**, v. 214, p. 97-104, 1990.

WANG, L; ROTHEMUND, D; CURD, H; REEVES, P. R. Sequence diversity of the *Escherichia coli* H7 *fliC* genes: Implication for a DNA-based typing scheme for *E. coli* O157:H7. **J. Clin. Microbiol.**, v.38 (5), p.1786-1790, 2000.

WANG, L; ROTHEMUND, D; CURD, H; REEVES, P.R. Species-wide variation in the *Escherichia coli* flagellin (H-antigen) gene. **J. Bacteriol.**, v. 185 (9), p. 2936-2943, 2003.

WEI, L. N; & JOYS, T. M. Covalent structure of three phase-1 flagellar filament proteins of *Salmonella*. **J. Mol. Biol.**, v. 186 (4), p. 791-803, 1985.

WHITFIELD, C; WALKER, S. G; ATKINSON, C. F; LAM, J. S; MACDONALD, L. A; BEVERIDGE, T. J; ORSKOV, I; ORSKOV, F. Serotype-specific monoclonal antibodies against the H12 flagellar antigen of *Escherichia coli*. **J. Gen. Microbiol.**, v. 134, p. 1747-1753, 1988.

WINSTANLEY, C & MORGAN, J. A. W. The bacterial flagellin gene as a biomarker for detection, population genetics and epidemiological analysis. **Microbiology**, v. 143, p. 3071-3084, 1997.

WOOD, T. K; BARRIOS, A. F. G; HERZBERG, M; LEE, J. Motility influences biofilm architecture in *Escherichia coli*. **Appl Microbiol Biotechnol.**, v. 72, p. 361-367, 2006.

YAMAMOTO, S & KUTSUKAKE, K. FljA – mediated post-transcriptional control of phase 1 flagellin expression in flagellar phase variation of *Salmonella enterica* serovar *Typhimurium*. **J. Bacteriol.**, v.188 (3), p. 958-967, 2006.

VIII-ANEXOS

Anexo 1



Universidade Estadual de Campinas
Instituto de Biologia



CEEA-IB-UNICAMP

Comissão de Ética na Experimentação Animal
CEEA-IB-UNICAMP

CERTIFICADO

Certificamos que o Protocolo nº 781-1, sobre "CARACTERIZAÇÃO DE NOVOS ANTÍGENOS FLAGELARES DE ESCHERICHIA COLI DE ORIGEM HUMANA" sob a responsabilidade de Prof. Dr. Domingos da Silva Leite / Monique Ribeiro Tiba está de acordo com os Princípios Éticos na Experimentação Animal adotados pelo Colégio Brasileiro de Experimentação Animal (COBEA), tendo sido aprovado pela Comissão de Ética na Experimentação Animal (CEEA)-IB-UNICAMP em reunião de 17 de fevereiro de 2005.

CERTIFICATE

We certify that the protocol nº 781-1, entitled "NEW H ANTIGENS CHARACTERIZATION OF ESCHERICHIA COLI FROM HUMAN STRAINS", is in agreement with the Ethical Principles for Animal Research established by the Brazilian College for Animal Experimentation (COBEA). This project was approved by the institutional Committee for Ethics in Animal Research (State University of Campinas - UNICAMP) on February 17, 2005.

Campinas, 17 de fevereiro de 2005.

Profa. Dra. Liana Verinaud
Presidente - CEEA/IB/UNICAMP

Fátima Alonso
Secretária - CEEA/IB/UNICAMP

Anexo 2



COMISSÃO INTERNA DE BIOSSEGURANÇA - CIBio
INSTITUTO DE BIOLOGIA
UNIVERSIDADE ESTADUAL DE CAMPINAS

Ofic. CIBio 001/2005.

Cidade Universitária "Zeferino Vaz",
14 de março de 2005.

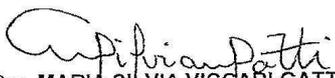
Prof. Dr. TOMOMASA YANO
Chefe do Departamento de Microbiologia e Imunologia
Instituto de Biologia
UNICAMP

Prezado Professor:

Informamos que o projeto abaixo relacionado, envolvendo OGM do tipo I, sob responsabilidade do Prof. Dr. DOMINGOS DA SILVA LEITE, foi aprovado pela CIBio-IB/UNICAMP na Reunião Ordinária de 09 de março de 2005, para ser desenvolvido nas dependências do Departamento de Microbiologia e Imunologia desta Unidade.

No. Projeto (Data da aprovação)	Data de recepção	Nome do Projeto	Prazo para envio de relatório à CIBio
CIBio 2005/01 (9/3/2005)	17/2/2005	Caracterização de novos antígenos flagelares de <i>Escherichia coli</i> de origem humana	agosto/2005

Atenciosamente,


Profa. Dra. MARIA SILVIA VICCARI GATTI
Presidente da CIBio/IB/UNICAMP