



ILMD
INSTITUTO LEONIDAS
& MARIA DEANE
Fiocruz Amazônia

UNIVERSIDADE FEDERAL DO AMAZONAS

INSTITUTO DE CIÊNCIAS BIOLÓGICAS

PROGRAMA MULTI-INSTITUCIONAL DE PÓS-GRADUAÇÃO

EM BIOTECNOLOGIA

JOÃO VICTOR DE MELO VERÇOSA

**GENÔMICA COMPARATIVA DE *SHIGELLA* SPP. ISOLADAS DE CRIANÇAS COM
DOENÇAS DIARREICAS NA AMAZÔNIA**

MANAUS - AM

2018

JOÃO VICTOR DE MELO VERÇOSA

**GENÔMICA COMPARATIVA DE *SHIGELLA* SPP. ISOLADAS DE CRIANÇAS COM
DOENÇAS DIARREICAS NA AMAZÔNIA**

Dissertação de Mestrado em Biotecnologia
para a obtenção do título de Mestre em
Biotecnologia pela Universidade Federal do
Amazonas do Programa Multi-institucional de
Pós-Graduação em Biotecnologia

Orientadora: Patrícia Puccinelli Orlandi

Co-Orientadora: Tainá Raiol Alencar

MANAUS - AM

2018

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

V482g Vercosa, João Victor de Melo
Genômica comparativa de *Shigella* spp. isoladas de crianças com doenças diarreicas na Amazônia / João Victor de Melo Vercosa. 2018
79 f.: il. color; 31 cm.

Orientadora: Patrícia Puccinelli Orlandi
Coorientadora: Tainá Raiol Alencar
Dissertação (Mestrado em Biotecnologia) - Universidade Federal do Amazonas.

1. Shigella. 2. genômica. 3. bioinformática. 4. shigelose. 5. sst3. I. Orlandi, Patrícia Puccinelli II. Universidade Federal do Amazonas III. Título

AGRADECIMENTOS

Agradeço a **Deus** pelo dom da vida, por todas oportunidades e portas que Ele tem aberto na minha vida.

Aos meus pais, **Wilson** e **Ivete**, por sempre acreditarem em mim incessantemente, até mesmo quando eu não acredito. À minha irmã **Rachel**, pelo companheirismo e por me ajudar sempre. Amo vocês!

Agradeço à minha orientadora **Dr^a. Patrícia Orlandi** por ter confiado em mim todos esses anos e compartilhado um pouco do seu conhecimento comigo. À minha co-orientadora, **Dr^a. Tainá Raiol**, por sempre me ajudar com toda paciência e acreditar na minha capacidade. Muito obrigado!

À **Paula**, que sempre se dispôs a ajudar no que pode, estando presente durante todo o processo. Você é demais!

Obrigado aos amigos do DCDIA, **Fernanda, Késsia, Juliane, Ruth, Natália, Paula, Edilene, Daniele, Wellington, Gleice, Ingrid, Aline, Lucas, Vani, Paloma, Emily, Ana Paula, e Vanda**. Também aos amigos que fiz na minha turma de mestrado, em meio às disciplinas e reuniões, **Aline Brilhante, Bruno, Dannel, Laís, Anny e Sabrina**.

Ao meu amigo **Ismael Rabelo**, que esteve acompanhando todo esse processo e me dando o apoio necessário. Obrigado por tudo!

Aos amigos da faculdade, **Polyane, Isabela, Alessandra, Marcos, Rommel, Andressa, Nani, Duda** e todos que passaram por essa etapa comigo, com quem eu pude dividir momentos inesquecíveis. Obrigado pelas conversas, mesmo que muitas vezes por videoconferência.

Ao **Dr. Paulo Afonso** que somou bastante nesse trabalho, possibilitando minha ida à Brasília e ao **Dr. Marcelo Brígido** da UNB, por ter aberto às portas do seu laboratório para mim.

Obrigado à **Andressa Valadares** da UNB, por ter sido sempre prestativa me ajudando com todas minhas dúvidas e ensinando sobre bioinformática. Muito obrigado!

Aos membros da banca da minha qualificação, Dr^a. **Márcia Neiva**, **Dr. João Vicente** e **Dr. Diogo Castro**, que me auxiliaram a melhorar esse trabalho e pelas sugestões.

Ao **Programa de Pós-Graduação em Biotecnologia (PPGBIOTEC)** e toda sua coordenação e secretaria, **Dr. Carlos Gustavo**, **Dr. Edmar Vaz de Andrade**, **Dr. Adolfo José da Mota**, **Nubiane** e **Jackeline**. Ao corpo docente e discente que tornaram essa experiência única.

À **Fiocruz Amazônia – ILM**, pela estrutura física e suporte financeiro na viagem à Brasília, bem como à **Fiocruz Brasília** que me recebeu em suas instalações.

Aos colaboradores do **Instituto Oswaldo Cruz (IOC)** e da Faculdade de Computação da **Universidade Federal do Mato Grosso do Sul (UFMS)**.

À **Fundação de Amparo à Pesquisa na Amazônia (FAPEAM)** pelo apoio financeiro fornecido no decorrer desse trabalho.

"In lumine Tuo, videbimus lumen"

(Salmos 36:9)

RESUMO

A diarreia é a segunda maior causa de mortalidade infantil no mundo. Dentre as doenças diarreicas, destaca-se a shigelose, uma doença caracterizada pela presença de muco e sangue nas fezes, além de febre, cólica e tenesmos. Atualmente no mundo, ocorrem cerca de 165 milhões de casos da doença com 1,5 milhões de mortes por ano. Essa doença tem como seu agente etiológico a *Shigella*, sendo quatro espécies pertencentes a este gênero: *S. flexneri*, *S. sonnei*, *S. boydii* e *S. dysenteriae*; sendo as duas primeiras as mais endêmicas no mundo. A patogenicidade da *Shigella* é dependente do Sistema de Secreção do Tipo 3 (SST3), localizado no pINV, um plasmídeo de invasão encontrado em todas as espécies do gênero. O SST3 expressa proteínas que driblam o sistema imunológico do hospedeiro, fazendo com que a bactéria consiga se alojar e multiplicar-se dentro das células epiteliais do intestino. Além do SST3, podem ser encontrados na *Shigella* os Sistemas de Secreção Tipo II (SST2) e VI (SST6), o primeiro em *S. boydii* e o segundo, mais recentemente descrito, em *S. sonnei*. Um estudo realizado pelo grupo Diagnóstico e Controle de Doenças Infeciosas na Amazônia (DCDIA) em 2009, foram coletadas amostras de fezes de 1339 crianças, no qual foram encontradas 30 cepas de *Shigella*. Destas, foram selecionadas 7, que seguiram para o sequenciamento completo de seus genomas no equipamento Illumina HiSeq 2500. Os contigs foram montados utilizando a abordagem *de novo*, no software Velvet e os *scaffolds* montados pelo ABACAS. A anotação foi feita pelo pipeline do Prokka. As amostras foram classificadas a partir da comparação de sequências em vários bancos de dados, como Genbank, utilizando a ferramenta BLAST, RDP e SILVA. Das 7 bactérias, 4 foram classificadas como *S. flexneri* e 3, como *S. boydii*, uma espécie raramente encontrada fora da Índia. Com uma anotação feita pelo KAAS, foram visualizadas vias metabólicas no KEGG, utilizando o pacote Pathway, no qual foi identificado o SST6 nas 3 representantes da *S. boydii*. Esse sistema garante uma vantagem adaptativa competitiva para a bactéria, podendo ser a resposta do porquê o número de casos com *S. boydii* fora da Índia tem crescido.

Palavras-chaves: *Shigella*; genômica; bioinformática; shigelose; SST3.

ABSTRACT

Diarrhea is the second leading cause of child mortality in the world. Diarrheal diseases include shigellosis, a disease characterized by the presence of mucus and blood in the faeces, in addition to fever, colic and tenesms. Currently, around 165 million cases of the disease occur worldwide, with 1.5 million deaths a year. This disease has as its etiologic agent the *Shigella*, being four species belonging to this genus: *S. flexneri*, *S. sonnei*, *S. boydii* and *S. dysenteriae*; the first two being the most endemic in the world. The pathogenicity of *Shigella* is dependent on the Type 3 Secretion System (T3SS), located in pINV, an invasion plasmid found in all species of the genus. T3SS expresses proteins that dribble the host's immune system, causing the bacteria to lodge and multiply within the intestinal epithelial cells. Besides the SST3, Type 2 (T2SS) and 6 (T6SS) Secretion Systems can be found in *Shigella*, the first in *S. boydii* and the second, more recently described, in *S. sonnei*. A study conducted by the group Diagnóstico e Controle de Doenças Infeciosas na Amazônia (DCDIA) in 2009, stool samples were collected from 1339 children, in which 30 strains of *Shigella* were found. From these, 7 were selected, which followed for the complete sequencing of their genomes in the Illumina HiSeq 2500 equipment. The contigs were assembled using the approach *de novo* in the Velvet software and the *scaffolds* assembled by ABACAS. The annotation was made by the Prokka pipeline. The samples were classified from the comparison of sequences in several databases, such as Genbank, using the tool BLAST, RDP and SILVA. Of the 7 bacteria, 4 were classified as *S. flexneri* and 3, as *S. boydii*, a species rarely found outside India. With an annotation made by KAAS, metabolic pathways were visualized in the KEGG using the Pathway package, in which the T6SS was identified in the 3 representatives of *S. boydii*. This system guarantees a competitive adaptive advantage for the bacteria and may be the answer as to why the number of *S. boydii* cases outside India has increased.

Keywords: *Shigella*; genomics; bioinformatics; shigellosis; T3SS.

LISTA DE ABREVIATURAS E SIGLAS

BLAST	Basic Local Alignment Search Tool
CIC	Departamento da Ciência da Computação
DCDIA	Laboratório de Diagnóstico e Controle de Doenças Infecciosas da Amazônia
DNA	Ácido desoxirribonucleico
DTA	Doenças transmitidas por alimentos
ECDC	European Centre for disease prevention and control
EIEC	<i>Escherichia coli</i> Enteroinvasiva
ESPGHAN	Sociedade Européia de Gastroenterologia pediátria, hepatologia e nutrição
ETEC	Escherichia coli Enterotoxigênica
FACOM	Faculdade de Computação da Universidade Federal do Mato Grosso do Sul
FIOCRUZ	Fundação Oswaldo Cruz
HIV	Vírus da Imunodeficiência Humana
HUS/SHU	Síndrome hemolítico-urêmica
ILMD	Instituto Leônicas e Maria Deane
IOC	Instituto Oswaldo Cruz
KAAS	KEGG Automatic Annotation Server
Kb	Quilobyte
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	KEGG Orthology
LB	Meio Luria-Bertani
LPS	Lipopolissacarídeo
LTD	L-triptofano desaminase
Meio EPM	Escola Paulista de Medicina
MILI	Motilidade, indol e lisina
mL	Mililitro
NCBI	National Center for Biotechnology Information
NGS	New Generation Sequencing (Sequenciamento de nova geração)

°C	Grau Celcius
OMS	Organização Mundial da Saúde
ORF	Open reading frame
Pb	Par de base
PCR	Reação em cadeia da polimerase
pINV	Plasmídeo de virulência
RBS	Sítio de ligação do ribossomo
RDP	RibosomalDatabase Project
RNA	Ácido ribonucleico
rRNA	RNA ribossomal
SBH	Single-directional best-hit
SS Ágar	Salmonella Shigella Ágar
ssp.	Espécies
SST2	Sistema de Secreção tipo 2
SST3	Sistema de Secreção tipo 3
SST6	Sistema de Secreção tipo 6
tRNA	RNA transportador
UFAM	Universidade Federal do Amazonas
uL	Microlitro
UNB	Universidade Nacional de Brasília

LISTA DE TABELAS

TABELA 1: GENOMAS DO GENBANK UTILIZADOS PARA A MONTAGEM DE REFERÊNCIA PELO ABACAS.....	40
TABELA 2: CEPAS UTILIZADAS COMO REFERÊNCIA PARA A MONTAGEM DOS CROMOSSOMOS COM O ABACAS.	42
TABELA 3: CEPAS UTILIZADAS PARA A MONTAGEM DOS PLASMÍDEOS DAS CEPAS IDENTIFICADAS COMO <i>SHIGELLA BOYDII</i>	44
TABELA 4: CEPAS UTILIZADAS PARA AS MONTAGENS DOS PLASMÍDEOS DAS CEPAS DE <i>SHIGELLA FLEXNERI</i>	44
TABELA 5: GALERIA BIOQUÍMICA DAS CEPAS SEQUENCIADAS.....	453
TABELA 6: ANÁLISE QUANTITATIVA DA PCR DE GENOTIPAGEM	464
TABELA 7: ESTATÍSTICAS GLOBAIS DO SEQUENCIAMENTO	51
TABELA 8: TABELA COMPARATIVA ENTRE OS BANCOS DE DADOS UTILIZADOS PARA CLASSIFICAÇÃO DAS AMOSTRAS.....	53
TABELA 9: MONTAGENS DOS PLASMÍDEOS DAS CEPAS <i>SHIGELLA BOYDII</i> STR. 192 E <i>SHIGELLA BOYDII</i> STR. 183	63
TABELA 10: MONTAGENS DOS PLASMÍDEOS DAS CEPAS DE <i>SHIGELLA FLEXNERI</i>	61
TABELA 11: ANÁLISE ETIO-EPIDEMIOLÓGICA DOS PACIENTES NO QUAL FORAM COLETADAS AS AMOSTRAS DE SANGUE.....	66

LISTA DE FIGURAS

FIGURA 1: FIGURA REPRESENTATIVA DA INVASÃO DA <i>SHIGELLA</i> EM UMA CÉLULA EPITELIAL INTESTINAL.	20
FIGURA 2: DISTRIBUIÇÃO GEOGRÁFICA DE INCIDÊNCIA DE SHIGELOSE.....	23
FIGURA 3: PROPORÇÃO DE <i>S. SONNEI</i> PARA <i>S. FLEXNERI</i> ISOLADAS DE 100 PAÍSES ENTRE 1990-2014.....	26
FIGURA 4: REPRESENTAÇÃO ESQUEMÁTICA DA ESTRUTURA DO ANTÍGENO-O	27
FIGURA 5: FLUXOGRAMA DE ATIVIDADES	38
FIGURA 6: PCR DE GENOTIPAGEM.....	46
FIGURA 7: PONTUAÇÃO PHRED DE CADA BASE DOS FRAGMENTOS SEQUENCIADOS.....	48
FIGURA 8: PONTUAÇÃO PHRED DE CADA BASE DOS FRAGMENTOS SEQUENCIADOS.....	49
FIGURA 9: PONTUAÇÃO PHRED DE CADA BASE DOS FRAGMENTOS SEQUENCIADOS.....	50
FIGURA 10: REPRESENTAÇÃO CIRCULAR DOS CROMOSSOMOS MONTADOS..	52
FIGURA 11: DIAGRAMA DE VENN PARA A ESPÉCIE <i>SHIGELLA FLEXNERI</i>	54
FIGURA 12: ALINHAMENTO MÚLTIPLO DOS GENOMAS. DE CIMA PARA BAIXO..	55
FIGURA 13: ÁRVORE FILOGENÔMICA GERADA PELO PIPELINE ORTHOLOGSORTER.....	57
FIGURA 14: MAPA METABÓLICO DE KEGG DAS AMOSTRAS DE <i>SHIGELLA BOYDII</i>	59
FIGURA 15: MAPA DE KEGG REPRESENTANDO A PRESENÇA DO SISTEMA DE SECREÇÃO DO TIPO 3 NAS CEPAS DE <i>SHIGELLA FLEXNERI</i>	60
FIGURA 16: MAPA DE KEGG REPRESENTANDO A PRESENÇA DOS GENES ENVOLVIDOS NOS SISTEMAS DE SECREÇÃO BACTERIANO.....	62

SUMÁRIO

INTRODUÇÃO	15
1. REVISÃO BIBLIOGRÁFICA	17
1.1. Doenças diarreicas.....	16
1.2. Shigelose.....	17
1.2.1 Patogênese.....	17
1.2.2. Sintomas.....	19
1.3. Epidemiologia.....	20
1.3.1. No Mundo.....	21
1.3.2. No Brasil.....	22
1.3.3. Shigella como indicador econômico.....	24
1.4. Shigella spp.....	25
1.5. Diferença entre a Shigella spp. e E. coli.....	27
1.6. Sistemas de Secreção.....	30
1.7. Genômica.....	31
1.8. Montagem.....	32
1.9 Anotação.....	32
2. JUSTIFICATIVA	36
3. OBJETIVO GERAL	37
3.1. OBJETIVOS ESPECÍFICOS	33
4. MATERIAL E MÉTODOS	38
4.1. Cepas bacterianas.....	36
4.2. Reativação das cepas.....	37
4.3. Extração de DNA genômico e sequenciamento.....	37
4.4 Isolamento do DNA e amplificação.....	37
4.5. Filtragem e montagem.....	38
4.6. Anotação gênica.....	39
4.7. Análise de vias metabólicas.....	40

4.8. Genômica comparativa.....	41
4.9. Montagem dos plasmídeos.....	41
5. RESULTADOS E DISCUSSÃO	45
5.1. Isolamento e identificação das cepas de Shigella.....	43
5.2. Reação em cadeia da polimerase (PCR).....	43
5.3. Sequenciamento, montagem e anotação.....	45
5.4. Classificação.....	51
5.5. Genômica comparativa.....	51
5.6. Análises filogenéticas.....	54
5.7. Sistemas de secreção.....	56
5.8. Montagem dos plasmídeos.....	61
5.9. Análise Global.....	63
6. CONCLUSÃO	69
7. PERSPECTIVAS FUTURAS	71
8. REFERÊNCIAS BIBLIOGRÁFICAS.....	72
9. ANEXOS.....	78

INTRODUÇÃO

A doença diarreica é considerada um problema de saúde pública em todo mundo, especialmente nos países em desenvolvimento, onde as taxas da doença são maiores (ALBERT et al., 2009). Entre as faixas etárias mais afetadas, encontram-se crianças menores que 5 anos de idade, e entre os agentes causadores, vírus, parasitas e bactérias, entre elas, a *Shigella* spp. destaca-se por ser o patógeno responsável pela shigelose (SOUSA et al., 2013).

A shigelose, descoberta há mais de 100 anos, ainda tem uma grande representatividade para a saúde pública, sendo uma das principais causas de mortalidade infantil. Ela é responsável por aproximadamente 160 milhões de casos e um milhão de mortes anuais, principalmente nos países em desenvolvimento, devido à falta de saneamento básico e higiene (PHALIPON; SANSONETTI, 2007).

O patógeno causador, a *Shigella* spp, faz parte da família enterobacteriaceae, Gram-negativa, não móvel e intracelular, com via de transmissão fecal-oral, podendo induzir uma infecção sintomática com dose baixa (<10 bactérias) (THE et al., 2016). A *Shigella* engloba quatro espécies: *Shigella flexneri*, *Shigella sonnei*, *Shigella dysenteriae* e *Shigella boydii*, sendo estes os agentes etiológicos da disenteria bacilar ou Shigelose. A *S. flexneri* e *S. sonnei* são os principais responsáveis pela doença (ANDERSON et al., 2017a; MARTEYN; GAZI; SANSONETTI, 2012).

Na época de sua descoberta, por Kioshi Shiga, esse patógeno foi caracterizado como uma *Escherichia coli*, contudo, no final da década de 1940, após a descoberta das 4 espécies, Ewin diferenciou-as em gêneros distintos (LIMA; HAVT; LIMA, 2015). Uma das principais diferenças da *Shigella* é o Sistema de Secreção do Tipo 3, o SST3. Esse sistema concede a ela a capacidade de invadir a mucosa intestinal, contornando a imunidade inata do hospedeiro, onde, promove a ruptura da mesma, causando uma diarreia severa com febre, cólicas abdominais e retais, muco e até sangue nas fezes do paciente, demonstrando a característica invasiva natural da *Shigella* (ANDERSON; SANSONETTI; MARTEYN, 2016).

O grande responsável por essa capacidade invasiva está localizado dentro de um grande plasmídeo de virulência, chamado pINV. Nesse plasmídeo, que contém cerca de 200 kb, é encontrado o SST3, formado por um conjunto de 22 proteínas e oligômero

que forma uma estrutura semelhante a uma agulha, que fica ancorado na membrana da célula do hospedeiro, e é por onde são injetados os fatores de virulência (MARTEYN; GAZI; SANSONETTI, 2012; NIU et al., 2017).

Esse minucioso mecanismo de invasão tem dificultado a produção de uma profilaxia eficiente, como uma vacina ou probiótico para a sociedade. Isso é somado com a necessidade de abranger todos ou a maioria dos sorotipos existentes do gênero (THOMPSON; DUY; BAKER, 2015).

No pINV, além de conter todos os aparatos necessários para realizar a patologia, também está nos centros de discussão sobre a origem da *Shigella*. Acredita-se que a *Shigella* evoluiu de uma *Escherichia coli* não patogênica ao adquirir esse plasmídeo, logo, a presença dele é uma das principais diferenças entre os gêneros, dificultando até mesmo a identificação e diferenciação, o que impede o paciente de ter um tratamento eficiente (NIU et al., 2017).

Além do SST3, outros sistemas de secreção são encontrados nas espécies de *Shigella*, como o Sistema de Secreção do Tipo 2, o SST2, e o Sistema de Secreção do Tipo 6, o SST6. O SST2 é encontrado em *Shigella dysenteriae* e *Shigella boydii*, enquanto o SST6 foi descrito até então nas cepas de *Shigella sonnei*. Esses sistemas adicionais não estão diretamente relacionados com a patogenicidade, contudo, eles podem estar ligados à virulência, atenuando ou agravando-a. O SST2 é conhecido por ser uma via de secreção geral, e está associado com a produção de biofilme, enquanto o T6SS concede uma vantagem adaptativa para a bactéria quando colocada em um ambiente com outras, o que facilita sua disseminação, além de ser uma hipótese que pode ajudar a explicar a mudança global entre a *S. flexneri* e *S. sonnei*, no qual a segunda tem se destacado em mais países do que as demais (HAZEN et al., 2016; MATTOCK; BLOCKER, 2017a).

Com o advento da genômica, ficou mais fácil identificar o aparecimento de novos sistemas de secreção, assim como acompanhar a resistência bacteriana e a troca de genes, principalmente em *Shigella*, que tem seu genoma extremamente plástico, ou seja, possui uma grande capacidade de receber genes de outros organismos. A

genômica também pode auxiliar no detalhamento da patogenicidade da bactéria e auxiliar no controle do avanço da doença (THE et al., 2016a).

Isso posto, esse trabalho irá utilizar da genômica como baluarte para auxiliar não só no conhecimento dessas cepas, mas bem como uma estratégia de identificação.

1. REVISÃO BIBLIOGRÁFICA

1.1. Doenças diarreicas

Diarreia é uma doença caracterizada por três ou mais passagens de fezes aquosas em um dia, com ou sem vômito. Ela normalmente é causada por uma infecção, seja por bactérias, parasitas ou vírus. A diarreia pode durar dias, podendo deixar o paciente desidratado e com déficit dos nutrientes necessários, e levando nos casos mais graves a óbito. Essa infecção pode ser adquirida através da água ou de alimentos contaminados, ou transmitida de uma pessoa para outra por falta de higiene. Normalmente, os casos de diarreia são mais frequentes em locais onde há falta de saneamento básico e água potável para beber, cozinhar e limpar (KOTLOFF, 2017; “WHO | Diarrhoeal disease”, 2017). Existem três tipos de diarreia, diferenciadas pelo tipo clínico: diarreia aguda, sem a presença de sangue; e a diarreia sanguinolenta, ou disenteria, com a presença de sangue e/ou muco, febre, cólicas e tenesmos. Existe ainda a diarreia persistente, que dura mais de 14 dias (KOTLOFF, 2017).

A diarreia é a segunda maior causa de mortalidade infantil no mundo. Cada criança deve ter média 3 episódios infecciosos de diarreia por ano, e destes, cerca de 40% não são tratados adequadamente (THE et al., 2016a). Esses episódios consecutivos de diarreia, quando não levarem a óbito, afetam a absorção de nutrientes causando consequências em longo prazo no desenvolvimento e crescimento infantil, uma vez que este quadro infeccioso causa danos ao epitélio intestinal do hospedeiro. Essas consequências englobam a desnutrição, baixa estatura e deficiência de aprendizado, criando um ciclo, pois a deficiência de aprendizado pode se refletir em oportunidades de vida futuras, estando mais suscetível a outras doenças infecciosas (“DEFEATDD”, 2017, “WHO | Diarrhoeal disease”, 2017; KOTLOFF, 2017;).

A diarreia possui inúmeros patógenos, entre eles os principais são os vírus e as bactérias. Dentre os primeiros, destacam-se o rotavírus, calicivírus, adenovírus entérico. Dentre os causadores bacterianos, a *Shigella* e *Escherichia coli* enterotoxigênica

(ETEC) destacam-se como os principais agentes da diarreia severa e fatal em crianças, contando com aproximadamente 15% das 500 mil mortes por diarreia a cada ano. Em virtude disso, o desenvolvimento de uma vacina para *Shigella* e ETEC são prioridades para a OMS (Organização Mundial da Saúde) (“DEFEATDD”, 2017; KOTLOFF, 2017).

O fato da doença diarreica poder ser atribuída a inúmeros patógenos, levam a um diagnóstico precário pelos serviços de saúde, que cuidam dos sintomas independentemente do patógeno (WHITE; KANG, 2017).

1.2. Shigelose

A shigelose é a principal causa de doença diarreica bacteriana em crianças, sendo também uma das principais causas de morbidade e mortalidade em todo o mundo, principalmente em crianças menores de 5 anos e em países de situação com baixa e média renda. Em 2012, a OMS, Organização Mundial de Saúde, relatou que a shigelose contribui com cerca de 3,6% da carga global de doenças, com mais de 238 milhões de casos por ano, ocorrendo com maior frequência em países em desenvolvimento (BARRY et al., 2013; RAM et al., 2008; THE et al., 2016, PATH – DefeatDD.org, 2017).

A doença tem como agente etiológico qualquer uma das quatro espécies de *Shigella*. (UD-DIN; WAHID, 2014). A transmissão da doença ocorre de pessoa para pessoa pela via fecal-oral, através da ingestão de alimentos ou água contaminada, , além da disseminação por vetores como as moscas. A patologia tem início a partir da ingestão de 10 a 100 células bacterianas, o que é considerado um número baixo, facilitando a disseminação. A *Shigella* tem baixa suscetibilidade aos ácidos do estômago, o que também facilita sua disseminação (LIMA; HAVT; LIMA, 2015; SOUSA et al., 2013; THE et al., 2016b).

1.2.1. Patogênese

A doença tem início entre 24 e 48 horas a partir da ingestão da bactéria, que irá alojar-se no intestino grosso do ser humano. A bactéria invade a região do cólon,

gerando uma diarreia aquosa, que progride para uma colite grave. (LIMA; HAVT; LIMA, 2015; SOUSA et al., 2013).

No cólon, a bactéria penetra no epitélio intestinal através das células M, sendo logo em seguida, fagocitadas por macrófagos residentes. Dentro desses macrófagos, ela consegue lisar a membrana do fagossomo, induzindo a apoptose do macrófago e conseguindo escapar. Após o escape, a bactéria invade os enterócitos e dentro do citoplasma da célula, ela é capaz de se multiplicar e disseminar-se de maneira basolateral para as células. Para isso, ela forma uma calda de actina que a propulsiona horizontalmente até a borda da célula, onde irá induzir a formação de protusões, propiciando assim, a penetração para as células adjacentes. O ciclo da infecção é finalizado pela lise da membrana celular, liberando o patógeno no citoplasma da célula vizinha. Repetindo-se este processo até a destruição total da monocamada do epitélio intestinal (PARSOT, 2005).

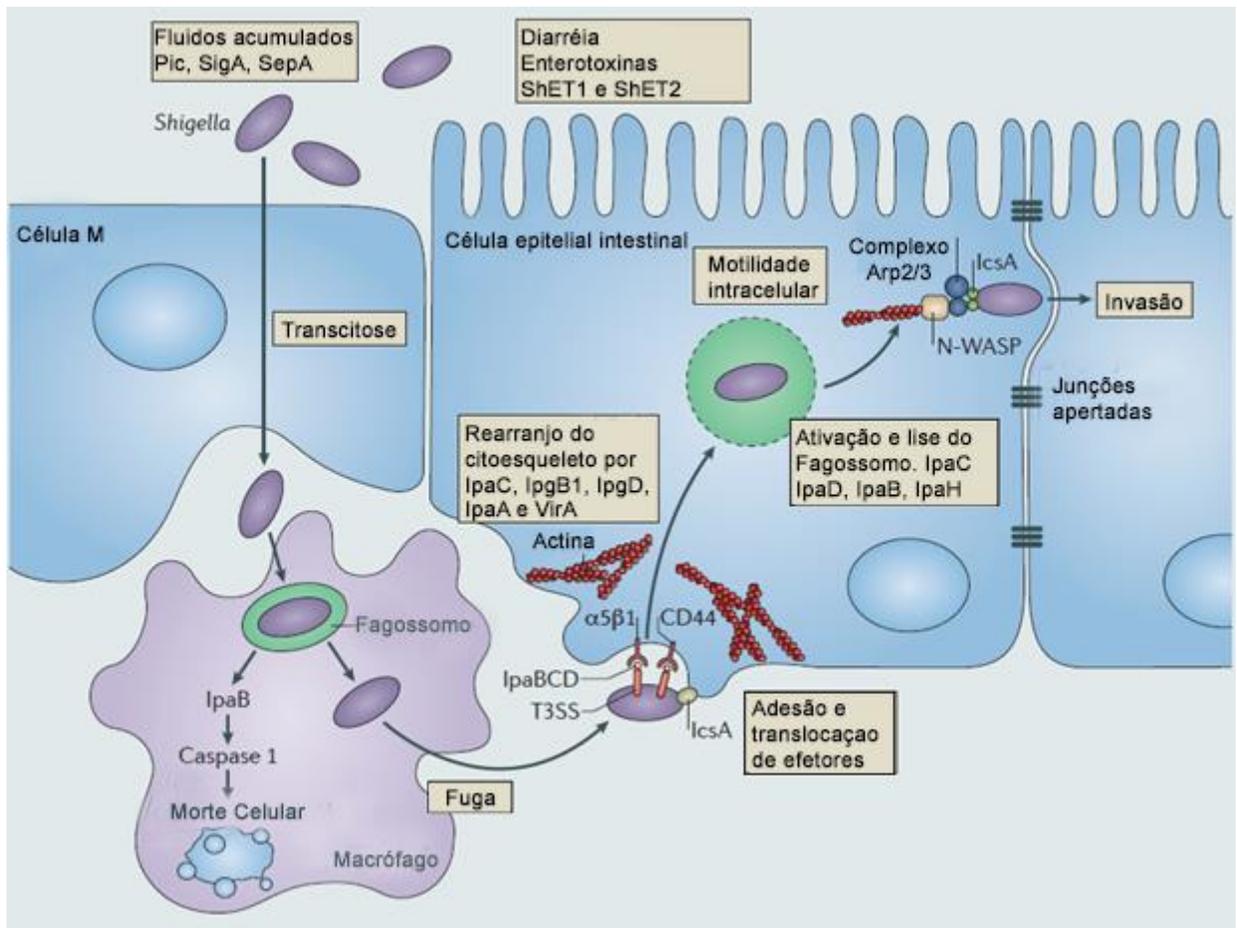


Figura 1: Figura representativa da invasão da *Shigella* em uma célula epitelial intestinal. A bactéria entra no tecido através das células M, através de transcitose e logo é englobada pelo Macrófago, que sofre morte celular devido a liberação de citocinas. Após a bactérias obter êxito nessa fuga, ela entra na célula epitelial, onde, com o auxílio do sistema de secreção do tipo 3, libera proteínas que causam rearranjo no citoesqueleto da célula, criando, portanto, uma causa de actina na bactéria, que agora consegue se propulsionar para a célula vizinha, passando basolateralmente.

1.2.2. Sintomas

A sintomatologia da shigelose pode variar entre uma diarreia com ou sem muco, febre, cólicas abdominais e tenesmo. Os sintomas podem durar entre 5 a 7 dias e a re-infecção é bem rara. Uma vez infectada com shigellose a pessoa se torna imune àquela cepa por muitos anos, porém, ainda são suscetíveis à infecção através de outra espécie de *Shigella* (RAM et al., 2008).

As complicações também podem causar uma artrite pós-infecciosa, com dores nas articulações, irritação nos olhos e micção dolorosa. Essa artrite atinge cerca de 2% dos infectados pela *S. flexneri*, e geralmente essas pessoas são geneticamente predispostas a ela. Mesmo que de forma rara, infecções no sangue também pode ocorrer, caso o intestino seja gravemente lesionado pela bactéria, que acaba entrando na corrente sanguínea. Esse tipo de infecção é mais comum em pacientes com sistema imunológico enfraquecido, como portadores de HIV ou câncer. As altas febres causadas pela *Shigella* também podem ocasionar convulsões, principalmente em crianças, além da Síndrome Hemolítico-Urêmica (HUS), que ocorre quando a bactéria entra no sistema digestivo, destruindo os glóbulos vermelhos pelas suas toxinas. Esse sintoma tem sua incidência dada mais comumente pela *S. dysenteriae*, produtora da toxina Shiga. (CENTERS FOR DISEASE CONTROL AND PREVENTION, 2016; PHALIPON; SANSONETTI, 2007; RAM et al., 2008).

O tratamento da shigelose não é similar aos de outros agentes enterotoxigênicos, com a reidratação oral. Logo, é preconizado o uso de antibióticos como tratamento padrão (BARRY et al., 2013; PARSOT, 2005). A OMS recomenda a utilização de Ciprofloxacina, enquanto a Sociedade Européia de Gastroenterologia Pediátrica, Hepatologia e Nutrição (ESPGHAN) recomenda o uso da azitromicina, além da cefalosporina de terceira geração (ceftriaxona), ácido nalidíxico e fluoroquinolonas como opções alternativas. Já no Brasil, o Ministério da Saúde recomenda a terapia combinada de trimetoprim com sulfametoxazol (CARRARI; TAHAN; MORAIS, 2012).

1.3. Epidemiologia

A diarreia é uma doença comum em todas as regiões do mundo. Em 2015 afetou cerca de 2,3 bilhões de pessoas, levando a óbito 1,3 milhões. Embora a doença acometa pessoas de todas as idades, 40% dos casos são em crianças menores de 5 anos, caracterizando-a como uma das principais doenças da infância, matando cerca de 520 mil crianças por ano. Somadas as mortes causadas pela Malária, Sarampo e AIDS, não superam o número de mortes causadas pela diarreia em crianças (KOTLOFF, 2017; LIMA; HAVT; LIMA, 2015; LIU et al., 2012).

Apesar desse número de mortes ter decaído 50% nos últimos 5 anos, a diarreia ainda é responsável por 8% das mortes em crianças (“Diarrhoeal Disease - UNICEF”, 2018). A maior ocorrência da shigelose, que é exclusiva de humanos, acontece nos países em desenvolvimento, chegando até 165 milhões de casos por ano, e embora esse número pareça alto, a expectativa é de que ele seja bem maior. Estima-se que a cada um caso de shigelose em uma comunidade, outras 25 pessoas também se infectem com a doença(LIMA; HAVT; LIMA, 2015; LIVIO et al., 2014; RAM et al., 2008)

Conquanto a morbidade pediátrica continue substancial, a mortalidade diminuiu, principalmente devido a diminuição dos casos de *S. dysenteriae* pelo mundo, espécie portadora da toxina Shiga. Outros motivos, incluem a implementação as diretrizes da OMS, que preconiza o uso de antibióticos para casos de diarreia clínica. Contudo o desenfreado uso desses medicamentos vem causando um crescimento na resistência das bactérias aos antibióticos que até então eram eficazes no tratamento da doença (LIVIO et al., 2014).

No tocante à distribuição geográfica das espécies, a *Shigella flexneri* e *Shigella sonnei* dividem o cargo de cepas endêmicas no mundo, entretanto, a *S. flexneri* se restringe aos países de baixo nível econômico, enquanto a *S. sonnei* ocupa os países mais desenvolvidos. A *S. boydii* possui poucos casos ao redor do mundo, sendo a forma endêmica da Índia, enquanto *S. dysenteriae* foi a grande causadora das epidemias de shigelose ao redor do globo, principalmente em países menos desenvolvidos, atualmente é a mais raramente isolada (MATTOCK; BLOCKER, 2017b).

1.3.1. No Mundo

A shigelose é uma doença endêmica em todo o globo, contudo nos países desenvolvidos, a grande maioria dos isolados de *Shigella* são provenientes de surtos de gastroenterites causadas por contaminações em alimentos ou água ou em casos de viajantes a países em que há uma grande taxa de shigelose (LIMA; HAVT; LIMA, 2015). As regiões com falta de saneamento básico, condições de superlotação e fraca infraestrutura de água e higiene geram condições ideais para os surtos de shigelose. Entre os exemplos destes surtos, destaca-se os surtos que ocorreram em meados de 2015, nos centros de refugiados na Europa, em que, somente na Áustria, entre os

meses de Julho a Novembro de 2015, foram reportados 23 casos dentro dos campos de refugiados, o que é um número significativo, tendo em vista que o país inteiro sofre em média com 70 casos anuais (EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL (ECDC), 2015; LEDERER et al., 2015)

Em um trabalho realizado na região sub-saariana da África e sul do continente asiático, a *Shigella* foi o 4º patógeno mais encontrado em diarreias, de um total de quase 10 mil crianças (KOTLOFF et al., 2013). Somente na Ásia, é estimado que ocorra anualmente 125 milhões de casos e 14 mil mortes. Isso se dá devido à baixa dose infecciosa necessária, a transmissão direta de pessoa para pessoa, transmissão através de água ou alimentos contaminados e sua baixa suscetibilidade aos ácidos estomacais (LIMA; HAVT; LIMA, 2015; THE et al., 2016b). A Figura 2 mostra um mapa epidemiológico, publicado em 2008, da incidência da shigelose no mundo, entre 1984 e 2005. (RAM et al., 2008)

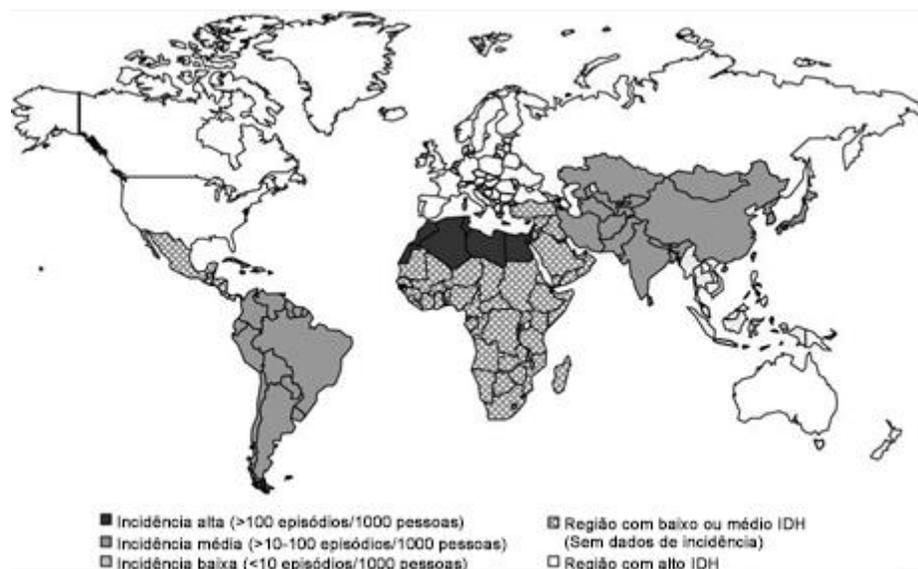


Figura 2: Distribuição geográfica de incidência de shigelose. É possível ver a grande incidência na região sub-saariana da África (cinza escuro), seguido de um nível médio de incidência na América do Sul e sul da Ásia (cinza). Os países hachurados em cinza são os países que não possuem dados, contudo a incidência é alta ou média, enquanto os países em branco são com IDH alto. O trabalho teve colaboração dos países: China, Tailândia, Egito, Bangladesh e Brasil. Fonte: RAM, et al., 2008.

1.3.2. No Brasil

Os dados epidemiológicos para o Brasil são escassos, principalmente para a shigelose, uma doença negligenciada no país, apesar de afetar muitas crianças. Na região nordeste, uma das regiões mais pobres do país, os dados são mais escassos ainda. Destaca-se um trabalho no estado do Piauí com 400 crianças, no qual 250 foram diagnosticadas com diarreia severa. Destas, 26 foram identificadas com *Shigella*, ou seja, 10% das crianças, sendo 21 com *S. flexneri*, representando 80,8%, e 5 com *S. sonnei*, com 19,2%. É importante salientar que 73% dos casos foi em crianças de até 2 anos, idade na qual o intestino ainda não está totalmente formado, ou seja, é extremamente prejudicial para a saúde para o bebê ter essa doença (NUNES et al., 2011)

Na região norte, mais especificamente no estado do Pará, de 377 crianças com diarreia, sendo que, 10% também destas era shigelose. No Amazonas, a *S. flexneri* também prevalece nos casos, similar ao Piauí. Em uma coleta com cerca de 1300 crianças, 30 amostras de *Shigella* foram isoladas, sendo 60% (18) de *S. flexneri*, 22,2% (8) de *S. sonnei*, 6,6% (2) de *S. dysenteriae* e *S. boydii* com a mesma porcentagem (CRUZ et al., 2014; DE OLIVEIRA SOUZA; CRISTINA DE MORAES SILVA; CARLOS BRITO LOUREIRO, 2012).

Ainda no norte do país, em Rondônia, a *Shigella* foi o quarto patógeno mais encontrado em crianças, ficando atrás do rotavírus, *E. coli* e *Salmonella*. 24 isolados de *Shigella* foram identificados, sendo 72% de *S. flexneri*, 12% *S. boydii*, 12% também de *S. sonnei* e 4% *S. dysenteriae* (SILVA et al., 2008). No mesmo estado, um outro trabalho isolou microrganismos de caldo de cana, muito popular em feiras regionais. De 15 amostras do suco, 80% delas estavam contaminadas com *Shigella* spp., seguido de 60% de amostras contaminadas com *Salmonella* spp. É importante salientar desses dados, que, 63% dos manipuladores do caldo afirmaram que nunca haviam feito um treinamento ou curso de manipulação de alimentos e boas práticas (ROCHA; BARCELOS; FARIA, 2016)

Um caso parecido foi identificado na região centro-oeste do Brasil, no estado do Mato Grosso, porém com frutas e hortaliças compradas de diferentes mercados da capital, Cuiabá. Foi encontrado *Shigella* spp. em 33,3% das frutas, como no mamão e melão. Enquanto 75% das hortaliças estavam contaminadas com *Shigella* spp., como no couve e cenoura. Nesse caso, o número de *Shigella* foi superior ao de *Salmonella* spp., contrastando com dados nacionais, que mostram que a *Salmonella* é campeã em doenças transmitidas por alimentos (DTA), enquanto a *Shigella* ocupa a 8ª posição, causando 2,44% dos casos de DTA (CRISTIANE HAUSCHILDT, 2014; DAL'MOLIN et al., 2014).

Em Minas Gerais, uma coleta realizada em um hospital de Belo Horizonte, na região sudeste do país, a *Shigella* foi encontrada em quase 11% das crianças que procuraram o hospital com diarreia, sendo que 88,2% foi de *S. sonnei*, se diferenciando do restante do país, onde a *S. flexneri* prevalece nos casos (SOUSA et al., 2013).

1.3.3. *Shigella* como indicador econômico

A *Shigella* é considerada um indicativo de desenvolvimento socioeconômico, tendo em vista que a *S. flexneri* é a forma mais comum a nível mundial, porém, mais frequentemente isolada em países em desenvolvimento, enquanto a *S. sonnei*, que também é bastante isolada mundialmente, é mais detectada em regiões mais desenvolvidas. Dos casos de shigelose identificados entre 2014 e 2015 na Europa, 49% foi de *S. sonnei*, seguido de 39% dos casos efetuados por *S. flexneri*, 7% por *S. boydii* e 4% por *S. dysenteriae*. (CHATTAWAY et al., 2017).

Porém, com a industrialização de muitos países, e com isso, o crescimento da renda, a *S. sonnei* vem emergindo nos países em desenvolvimento, como na Ásia, América Latina e Oriente Médio (Figura 3) (THOMPSON; DUY; BAKER, 2015).

Ao observar a renda *per capita* no estado do Amazonas e comparar com a renda do estado de Minas Gerais, é perceptível a diferença de desenvolvimento entre os estados. Enquanto o primeiro tem renda de R\$ 539,80, o segundo tem uma renda per

capita de R\$ 749,89, o que corrobora a ideia(ATLAS DO DESENVOLVIMENTO HUMANO NO BRASIL, 2013; LIMA; HAVT; LIMA, 2015).

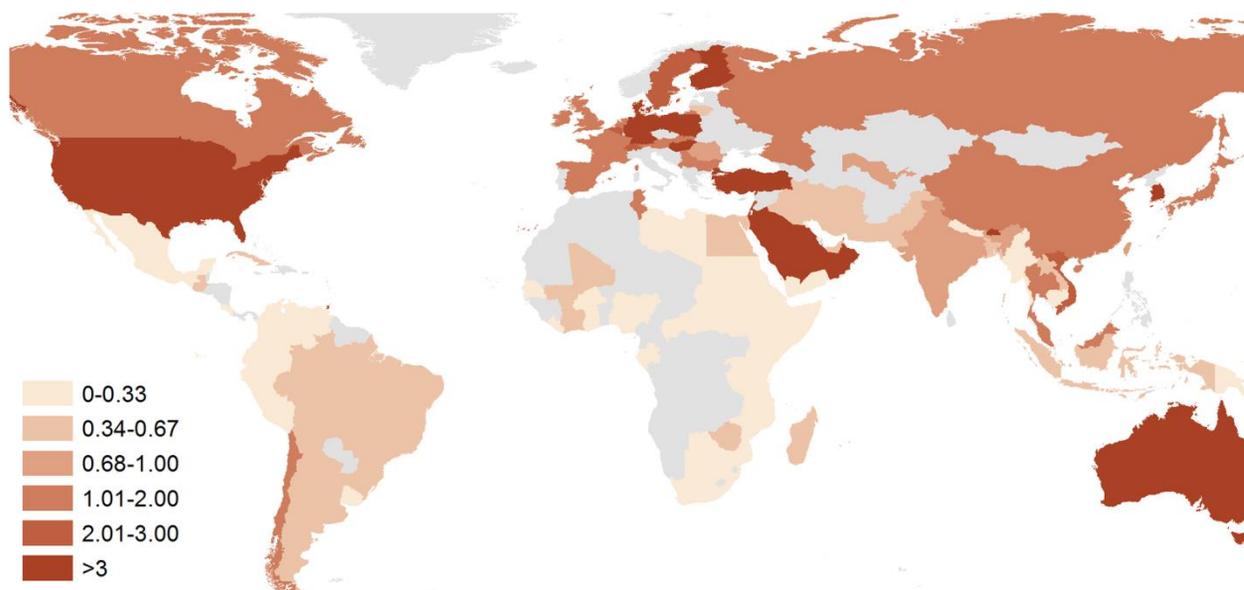


Figura 3: Proporção de *S. sonnei* para *S. flexneri* isoladas de 100 países entre 1990-2014. As cores mais escuras representam maior proporção de *S. sonnei*, enquanto as cores mais claras demonstram maior proporção de *S. flexneri*. Os países em cinza indicam que não há dados identificados sobre as espécies. Fonte: THOMPSON; DUY; BAKER, 2015

1.4. *Shigella* spp.

A *Shigella* foi identificada pela primeira vez no final do século XIX, em 1898, pelo cientista japonês Kioshi Shiga. Na ocasião, o médico examinou fezes de pacientes convalescentes após um surto de disenteria no Japão e a determinou seus achados como um *Bacillus*, devido sua semelhança com a então *Bacillus coli*, que só depois veio a ser chamada de *Escherichia coli*. Em 1940 que a *Shigella* se separou do gênero *E. coli* devido à importância patológica (LEVINE et al., 2007; THE et al., 2016a; UD-DIN; WAHID, 2014).

A bactéria isolada por Shiga foi a, hoje conhecida como, *Shigella dysenteriae* 1, porém dentro do gênero existem outras três espécies, que se diferenciam pelos seus sorotipos, que são identificados pela base da estrutura do antígeno-O, um

polissacarídeo (LPS): a *S. flexneri*, com 14 sorotipos, *S. sonnei*, um sorotipo, *S. boydii*, com 20 sorotipos, além da *S. dysenteriae*, com 15 sorotipos, totalizando 50 sorotipos (ANDERSON; SANSONETTI; MARTEYN, 2016; RAM et al., 2008).

É possível observar as diferenças dentro dos grupos sorológicos a partir de variações na composição do LPS (Figura 4). As regiões do antígeno-O são: as subunidades repetidas, que é a região variante, um núcleo interno e o lipídeo A (MARTEYN; GAZI; SANSONETTI, 2012).

A *Shigella* faz parte da família das enterobactérias, sendo gram-negativa, não móvel e lactose negativa, agente etiológico da desintéria bacilar ou shigelose. Ela é um patógeno intracelular, podendo induzir uma infecção sintomática com dose, contudo, o que a diferencia das demais enterobactérias é o fato dela ser um patógeno não afetar primatas, ou seja, exclusivo de seres-humanos (MATTOCK; BLOCKER, 2017b; THE et al., 2016a)

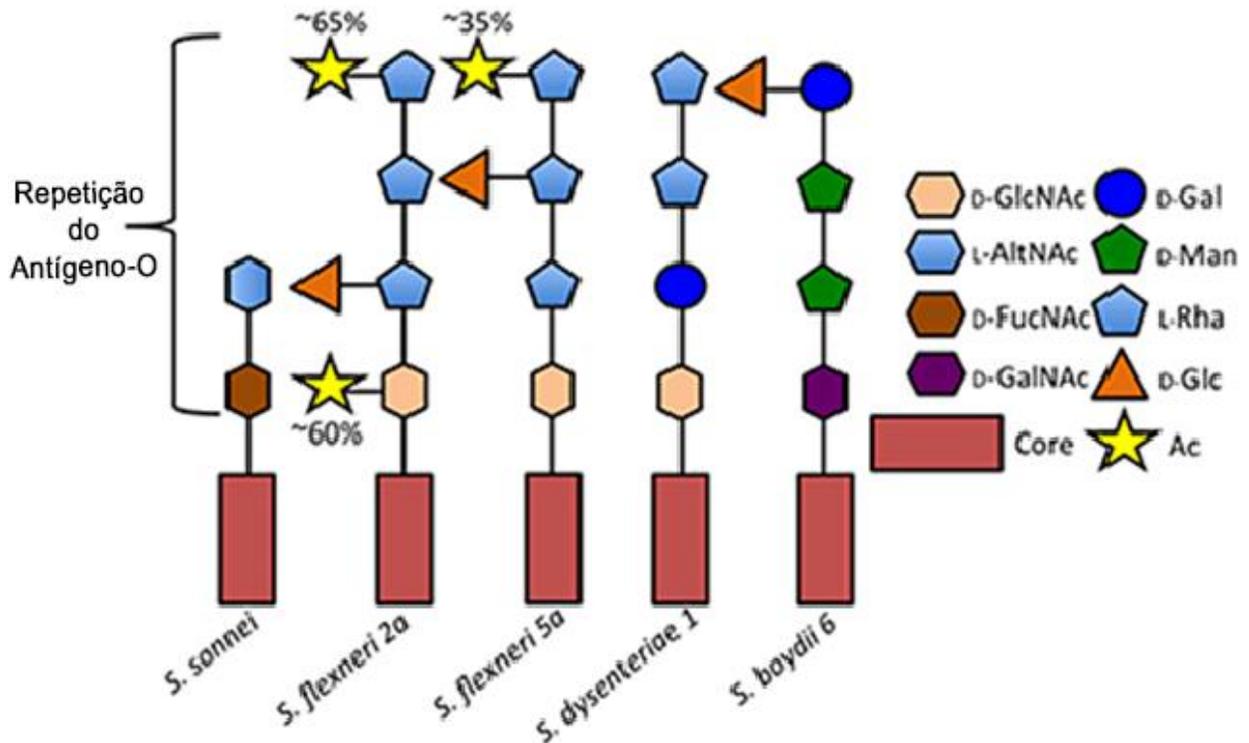


Figura 4: Representação esquemática da estrutura do antígeno-O: 2-acetamido-2-desoxi-D-glucose (D-GlcNAc); 2-acetamido-2-desoxi-L-ácido altruronic (L-AltNAc); 2-acetamido-4-amino-2, 4-

didesoxi-D-fucose (D-FucNac); 2-acetamido-2-desoxi-D- ácido galacturónico (D-GalNac), D-galactopirranose (D-Gal); D-manose (D-Man); L-ramnose (L-Rha); D-glucose (D-Glc); O-acetil (Ac); Oligossacarídeo que liga o antígeno-O ao lipídeo A (parte central). Fonte: Andreson et al., (2016).

1.5. Diferenças entre a *Shigella* spp. e *Escherichia coli*

Apesar de serem classificadas como gêneros diferentes, a *Shigella* e a *Escherichia coli* compartilham diversas semelhanças. Há quem as classifique até como o mesmo gênero. Uma hipótese que a *Shigella* surgiu de uma evolução da *Escherichia coli* e isso se deu devido a plasticidade que o genoma tem, com a possibilidade de aquisição de genes, plasmídeos, fagos e ilhas de patogenicidade por transferência genética horizontal, além da adaptação contínua que esses organismos sofrem através de mutações, rearranjos e deleções de genes desnecessários para adaptação no meio ambiente. No grupo da *Escherichia coli* existe a EIEC, a *E. coli* Enteroinvasiva, que também tem a capacidade invasiva e estudos mostram que a EIEC está em um cluster distinto das demais *E. coli*, e isso leva à uma outra hipótese de que essa diferenciação se deu em um evento de aquisição do plasmídeo de virulência, o pINV, que contém a maioria dos genes facilitadores da invasão, além de outros genes envolvidos na sobrevivência e disseminação intracelular (DI MARTINO et al., 2013).

Tendo em vista essas semelhanças, uma série de trabalhos tem sido desenvolvidos a fim de conhecer mais essa relação entre *Shigella* e EIEC. É sabido que, antes de se diferenciar em um gênero único, a *Shigella* foi identificada como uma *Escherichia coli* devido suas semelhanças, e com a *E. coli* enteroinvasiva essas semelhanças são maiores. Bioquimicamente, geneticamente e patogenicamente as duas são muito parecidas, quase que indistinguíveis em nível de espécie, somente em nível clínico, por exemplo, a *Shigella* spp. pode causar a HUS, síndrome hemolítica urêmica, caracterizada pela insuficiência renal relacionada a anemia hemolítica microangiopática e trombocitopenia, ao passo que não há relatos de EIEC causando HUS (KAPER; NATARO; MOBLEY, 2004; UD-DIN; WAHID, 2014).

A EIEC se diferencia das demais estirpes de *E. coli* patogênicas pois, enquanto a maioria permanece extracelular, a EIEC é um patógeno intracelular, capaz de invadir e

replicar dentro das células epiteliais e dos macrófagos. Algumas *E. coli* conseguem até ser internalizadas, mas em níveis muito mais baixos, além de não conseguirem se replicar dentro da célula (KAPER; NATARO; MOBLEY, 2004). Outra maneira de diferenciar a EIEC das *E. coli* comensais é pelo teste de Serény e através da identificação de proteínas ou genes relacionados a invasão. (UD-DIN; WAHID, 2014).

Apesar de serem bem similares, trabalhos demonstram que, do ponto de vista clínico, a distinção entre *Shigella* e EIEC é desnecessária, porém, na perspectiva da saúde pública, essa diferenciação pode ser importante. Na Holanda, por exemplo, a infecção causada por *Shigella* é considerada uma doença infecciosa, enquanto as por *Escherichia coli* não, com exceção da *Escherichia coli* Enterohemorrágica, a EHEC (VAN DEN BELD; REUBSAET, 2012).

Na busca da diferenciação entre *Shigella*, EIEC e *E. coli* não-invasivas, constataram que a EIEC é mais parecida com *Shigella* do que com as demais espécies do gênero *Escherichia*. É possível diferenciar *Shigella* e EIEC das demais *E. coli* através de um conjunto de técnicas, incluindo PCR do gene ipaH - que se faz presente nas duas primeiras -, através de ensaios fisiológicos e bioquímicos, além da tipagem sorológica (VAN DEN BELD; REUBSAET, 2012).

A primeira diferença entre as duas é a maneira de transmissão. *Shigella* é transmitida de modo fecal-oral, ou seja, através de um contato mais pessoal, ao passo que a EIEC está mais relacionada com águas e alimentos contaminados, embora a *Shigella* também seja um patógeno relacionado às DTA, as doenças transmitidas por alimentos (VIEIRA et al., 2007). Além disso, a EIEC ainda carrega uma elevada atividade metabólica, tendo em vista que ela consegue catabolizar diversos substratos utilizados pelas *E. coli*, contrastando mais uma vez com a *Shigella* (DI MARTINO et al., 2013).

Outro ponto que as diferencia é o nível de expressão dos genes de virulência, em que na *Shigella* é mais elevado (UD-DIN; WAHID, 2014). A *Shigella* consegue induzir uma infecção com uma baixa dose (<10 bactérias), ao contrário da *E. coli* que

para causar infecção tem a necessidade de no mínimo quatro vezes mais bactérias (THE et al., 2016).

Há uma metodologia que demonstra isso, é o teste de invasão pelo *Serény*, que demonstra a natureza virulenta da *Shigella* quando aplicada em cobaias, como coelhos ou porquinho-da-índia, causando uma ceratoconjutivite. Em EIEC a conjutivite se desenvolve muito mais lentamente e tardia do que a ceratoconjutivite causada por *Shigella*, porém a utilização deste teste como método diferencial na rotina laboratorial é inviável, devido as questões éticas e legais ao uso de animais, além de ser mais oneroso. Um dos motivos de ainda manterem a diferenciação taxonômica é o grau de severidade causado na disenteria da *Shigella* (VAN DEN BELD; REUBSAET, 2012).

Atualmente, não há um método de diferenciação de *Shigella* spp. e EIEC, logo, isso tem se dado de maneira clássica, por cultura convencional e com ensaios bioquímicos. Alguns testes bioquímicos podem distinguir EIEC de *Shigella*, como a fermentação do mucato e acetato, onde a EIEC pode ser positiva para ambas, contrastando com a *Shigella* que é negativa para ambas. Para complementar, Wahid propôs em 2014 um ensaio duplex de PCR em tempo real, utilizando os genes *uidA* e *lacY*, sendo a EIEC positiva para os dois, e *Shigella* negativa para *lacY* (KAPER; NATARO; MOBLEY, 2004; PAVLOVIC et al., 2011; UD-DIN; WAHID, 2014).

Uma metodologia utilizada para determinação de espécie é a amplificação do gene da girase, porém, foram encontradas mutações nesse gene na *Shigella dysenteriae*. O resultado do sequenciamento de diversas cepas completas e análises por estatística bayesiana constataram mutações e polimorfismo em diversos genes, incluindo o da girase, o que inviabilizou a metodologia (NJAMKEPO et al., 2016).

Além do uso falho do gene *gyr*, já foram adotadas metodologias usando *primers* em PCR multiplex (UD-DIN; WAHID, 2014), MALDI-TOF acoplado ao espectrômetro de massas (PAAUW et al., 2015), amplificação de genes específicos como o *CadA* e *ipaH* (REUNSAEY, 2011) e até a sorotipagem, mas como já citado acima, a associação dessas técnicas podem ser falhas ou onerosas, como a diferenciação utilizando a

reação de PCR quantitativa ou PCR Real-Time, que mostra a expressão diferencial entre os genes de virulência.

1.6. Sistema de Secreção

Grande parte da patogenicidade da *Shigella* se dá devido o sistema de secreção do Tipo 3, uma maquinaria localizada em um grande plasmídeo de virulência, encontrado em todas as espécies de *Shigella* e EIEC, e em outras bactérias Gram-negativas (KAPER; NATARO; MOBLEY, 2004). Esse sistema forma uma espécie de seringa ou agulha, que secreta múltiplas proteínas como IpaA, IpaB, IpaC e IpgD, que medeiam eventos de sinalização epitelial, rearranjos do citoesqueleto, captação celular, lise do vacúolo, entre outras reações. Essas proteínas são denominadas de proteínas efetoras, e, através dessa agulha, a bactéria consegue depositá-las no interior da célula-alvo. (PUHAR; SANSONETTI, 2014)

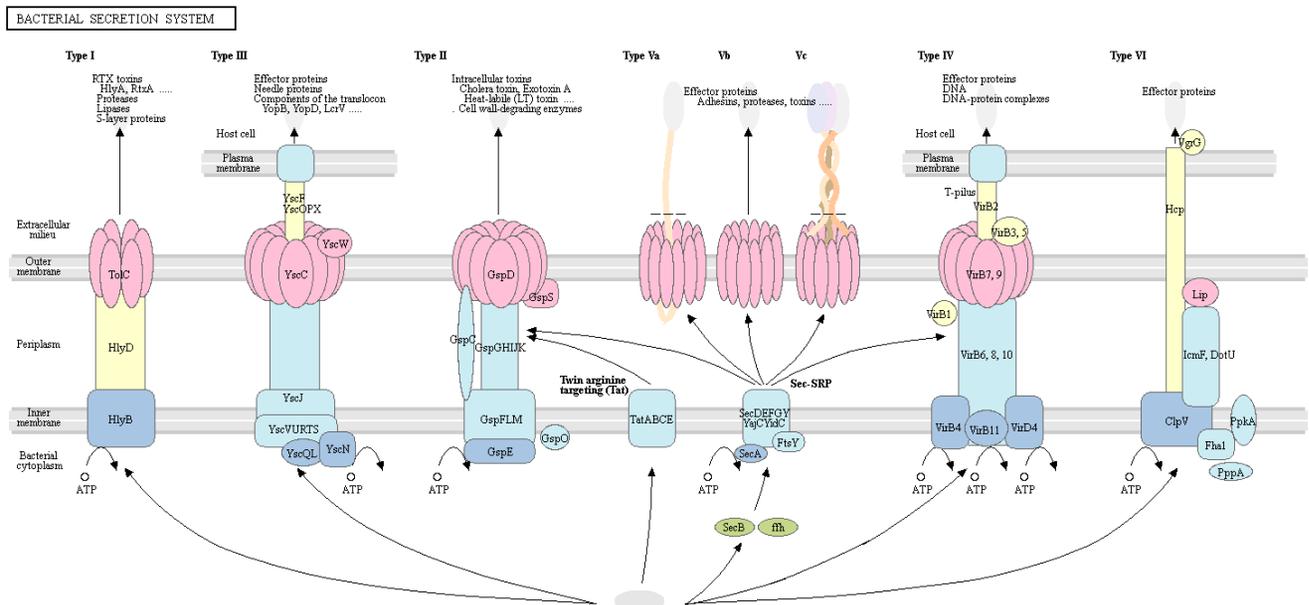


Figura 5: Sistemas de secreções existentes em bactérias. Os sistemas 3, 2 e 6 que já foram previamente descritos em cepas de *Shigella*. Com destaque para o SST3, que é o responsável pela virulência das espécies de *Shigella*. Ele é encontrado nos plasmídeos.

Além do SST3, existem outros sistemas de secreção, como o sistema secreção do tipo 2, SST2, que está comumente presente em bactérias Gram-negativas como *Pseudomonas aeruginosa* e *Vibrio cholerae*, porém, o mesmo já foi descrito em *Shigella boydii* e *Shigella dysenteriae* (KOROTKOV; SANDKVIST; HOL, 2012; YANG et al., 2005). Contudo, esse sistema é considerado inativo devido a uma mutação no gene *gspD* de um deslocamento em *gspC*. DUNSTAN e colaboradores, 2013, mostraram que *Escherichia coli* e *Shigella* devem ter adquirido esse sistema de *Vibrio*, devido a uma homologia no gene *AspS*, essencial para a funcionalidade do mesmo.

O sistema de secreção do tipo VI, SST6, foi encontrado recentemente em *Shigella sonnei*, e da mesma maneira que o SST3 e o SST2, esse sistema é utilizado pelas bactérias para inserir proteínas dentro da célula alvo, por exemplo, célula eucariótica do epitélio intestinal, no caso da *Shigella* e outras enterobactérias. Todavia, o SST6 tem a capacidade de fornecer uma vantagem competitiva de colonização dentro do intestino em relação à outras bactérias, como as da própria microbiota. (STARLING, 2017). O funcionamento desse Sistema é comparado à uma arma de fogo, que bombardeia as proteínas efetoras para dentro da célula alvo, que até 2010, era pensado que só atuava contra células eucariontes, contudo, posteriormente foi comprovado que ele age como uma arma contra outras bactérias também. Isso gera uma vantagem de ocupação contra bactérias toxigênicas no intestino. Um exemplo é o caso do *Bacteriodes fragilis*, que utiliza desse sistema de secreção contra *Bacteriodes fragilis* toxigênicos que interagem com eles no lúmen do intestino, matando-os. Outro exemplo é uma competição nutricional, caso que acontece entre as *Salmonella Typhimurium*, que competem contra *Klebsiella oxitoca* devido a compostos nutricionais. O mesmo caso já foi comprovado com a *S. flexneri* e *S. sonnei*, em que a segunda consegue “vencer” essa “luta”, como o autor descreveu, e ocupar o espaço da microbiota. (ANDERSON et al., 2017b; SANA; LUGO; MONACK, 2017)

1.7. Genômica

No que tange a identificação de microrganismos e seus estudos, ainda é aplicado o uso da microbiologia clássica, mas o uso das novas tecnologias de biologia molecular tem se tornado indispensável, principalmente no uso do diagnóstico, que deve ser rápido, particularmente no caso de estudo da *Shigella*, tem sua diferenciação com a *E. coli* de maneira dificultosa. (LIMA; HAVT; LIMA, 2015)

Entre essas novas tecnologias, o sequenciamento de alto desempenho (*High throughput sequence*) tem seu uso cada vez mais frequentemente utilizado devido a rapidez e baixo custo, considerando o volume de informações geradas para um genoma. Um genoma completo de um ser humano, por exemplo, pode sair por até U\$ 1000, o que há cerca de 20 anos atrás, durante o projeto Genoma Humano, saiu por quase U\$ 3 bilhões (VENTER, 2001). Isso se dá pela quantidade de plataformas existentes hoje, como 454, Illumina, ABISOLiD, ionTorrent, entre outras, que, em um único ensaio podem ser feitas análises para descobrir polimorfismo de nucleotídeo simples (SNP), inserções, deleções ou mutações nos genes, rearranjos estruturais, variação no número de cópias, entre outros (BAKER; HANAGE; HOLT, 2010; EKBLÖM; WOLF, 2014).

Essas novas metodologias estão sendo bastante utilizadas em trabalhos relacionados à saúde pública, tanto para isolados clínicos, quanto cepas laboratoriais e mutantes. Exemplos incluem o estudo de patogenicidade e resistência antimicrobiana (EDWARDS; HOLT, 2013). Os genomas completos também estão sendo empregados para análises de surtos e estudos filogenéticos para indicativos de disseminação (HOLT et al., 2012).

A genômica é uma área que vem crescendo, permitindo uma melhor compreensão dos sistemas biológicos e organismos, onde existem grupos de pesquisadores focados em desenvolver *softwares* eficientes para melhorar a montagem e análise dos genomas, além de algoritmos para uma validação rigorosa do desempenho dos programas (OLSON et al., 2015).

O uso da genômica poderá ajudar a caracterizar as características exclusivas da *Shigella*. Trabalhos mostram que há uma similaridade nucleotídica de 80 a 90% entre os genomas de *Shigella* e EIEC, contudo, ainda há cerca de 10% de diferenças abertas para estudos, e serão algumas dessas diferenças abordadas nesse trabalho, não só de outro gênero, mas entre as espécies também (VAN DEN BELD; REUBSAET, 2012).

1.8. Montagem

A montagem é uma das primeiras e principais etapas de análise do sequenciamento. É também uma etapa crucial para o trabalho, pois se não for feita uma boa montagem, o trabalho posterior será ineficiente, portanto, nessa etapa é preciso demandar tempo, recursos e experiência. É na montagem que as *reads*, o resultado do sequenciamento são sobrepostas formando *contigs*. Sequências muito longas, como genomas completos, geralmente são sequenciados a partir de ambas as extremidades para conferir uma maior confiança na hora da montagem, sem perder conteúdo. Após a montagem dos *contigs*, estes também são unidos para formar sequências maiores ainda, os *scaffolds* (EKBLÖM; WOLF, 2014)

Para definir o montador ou os montadores a serem utilizados é feita uma análise do tipo do material a ser sequenciado, buscando sempre ter o melhor resultado no menor tempo possível. Para uma montagem *de novo*, ou seja, uma montagem sem o uso de genoma de referência, existe, por exemplo, o montador Velvet. A montagem *de novo* é aplicada quando uma referência não gera um bom resultado de montagem ou quando o objeto de estudo é uma espécie nova. O Velvet é baseado no uso dos grafos de Bruijn, que, a partir das *reads*, remove os erros e tenta solucionar as repetições, além de outros parâmetros que são definidos de acordo com o tipo de montagem a ser feita (ZERBINO, 2010). Após a montagem, o próximo passo crucial é a anotação.

1.9. Anotação

Existem vários *softwares* para anotação de genomas, mas em sua maioria eles funcionam de maneira similar, divididas em duas etapas principais. A primeira etapa, a fase computacional, é caracterizada por ter marcadores de sequências expressas, as ESTs, e proteínas alinhadas ao genoma, através da predição de genes. Na segunda

etapa, a anotação propriamente dita, os genes são anotados através da utilização de diversas ferramentas diferentes. É nessa fase que os genes codificadores de proteínas são anotados, assim como RNAs também (YANDELL; ENCE, 2012)

Com o advento do sequenciamento de nova geração, o NGS, os softwares de montagem e anotação foram se atualizando também, e com isso, hoje estão disponíveis diversas ferramentas. Os programas atuais de anotação envolvem a predição de genes, localização dos RBS, os sítios de ligação do ribossomo, além de sítios terminais e domínios proteicos conservados. (RICHARDSON; WATSON, 2013)

O *software* de anotação utilizado nesse trabalho foi o Prokka, que funciona com base no uso da linha de comando em sistemas Unix, como Linux ou Mac. O Prokka gera um conjunto de ferramentas, resultando em uma anotação rápida e confiável para genomas bacterianos. O software recebe as sequências já montadas, no formato FASTA(SEEMANN, 2014).

2. JUSTIFICATIVA

A *Shigella* spp. foi descoberta há mais de 120 anos, conseguindo ultrapassar suas próprias barreiras, pois até então, era uma bactéria restrita às grandes guerras e epidemias de diarreia, porém, até hoje ela se perpetua e atinge os mais vulneráveis e menos favorecidos, como crianças e populações de países carentes. Contudo, o fato desse gênero ter inúmeras semelhanças com a *Escherichia coli*, isso abriu caminhos para o estudo da *Shigella* spp., pois muitas metodologias podem ser aplicadas de semelhante modo (LAMPEL; FORMAL†; MAURELLI, 2018).

Em trabalhos anteriores do grupo de pesquisa DCDIA (Laboratório de diagnóstico e controle de doenças infecciosas na Amazônia), foram isoladas 30 cepas de *Shigella*, e com isso, foram realizados diversos experimentos na tentativa de elucidar corretamente as espécies e caracterizá-las, porém, muitos desses resultados foram incertos para 7 destas isolados.

Entre as metodologias aplicadas estão o MALDI-TOF, o sequenciamento do DNA ribossomal 16S, galeria de bioquímicos da PROBAC, VITEK 2 (galeria bioquímica feita em sistema automatizado), sorotipagem da PROMICRO, além da galeria de bioquímicos feita em laboratório (indol, motilidade e Lisina). Os resultados desses experimentos foram majoritariamente inconclusivos, desta forma, esses 7 isolados foram selecionados para um sequenciamento completo de seu genoma, no qual estes foram montados, anotados e caracterizados, pois, apesar de os estudos da patogênese e mecanismo de invasão serem bem explorados, ainda há dificuldades na identificação das espécies e no tratamento. Tanto é que até hoje não existe uma vacina eficaz para combater a bactéria (LEVINE et al., 2007; PAAUW et al., 2015). Esse trabalho, portanto, será um avanço nesse sentido, além de nos permitir conhecer melhor as peculiaridades das cepas circulantes na região.

3. OBJETIVO GERAL

- Realizar genômica comparativa entre os genomas isolados de *Shigella* spp. de crianças com diarreia em Manaus, AM

3.1. OBJETIVOS ESPECÍFICOS

- Montar e anotar os genomas de diferentes espécies de *Shigella*;
- Comparar os genomas através da aplicação de ferramentas de bioinformática;

4. MATERIAL E MÉTODOS

4.1. Cepas Bacterianas

As cepas de *Shigella* utilizadas nesse trabalho foram provenientes de um estudo epidemio-etiológico realizada entre 2007 e 2009 no Hospital da Criança das zonas Leste, Oeste e Sul. Foram coletadas fezes – por evacuação espontânea – de cerca de 1339 crianças entre 0 a 10 anos de idade. O material foi processado e armazenado na Fiocruz Amazônia ILMD. Destas, 30 bactérias com aspectos fenotípicos do gênero *Shigella* foram isoladas, contudo, 7 tiveram resultados inconclusivos acerca de suas identificações.

Essas amostras foram selecionadas por já serem utilizadas em trabalhos no Laboratório de Diagnóstico e Controle de Doenças Infecciosas da Amazônia (DCDIA) da Fiocruz Amazônia - Instituto Leônidas e Maria Deane (ILMD). Essas amostras já passaram por inúmeras metodologias diferentes a fim de caracteriza-las.

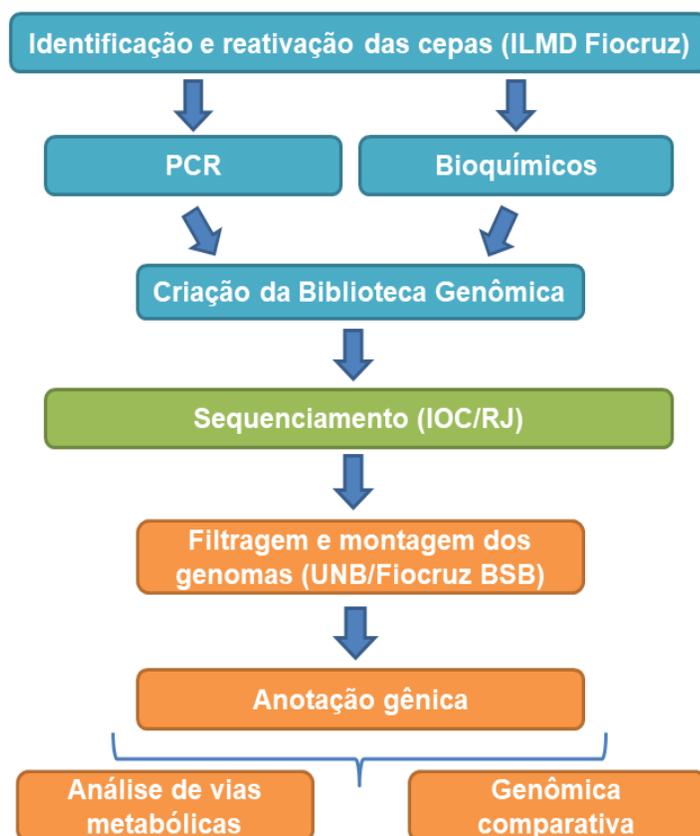


Figura 6: Fluxograma de atividades, desde a identificação das cepas até às análises de bioinformática a fim de determinar marcadores moleculares de diferenciação.

4.2. Reativação das cepas

As cepas de *Shigella* foram reativadas do meio definitivo, sendo cultivadas em caldo Luria-Bertani (LB), a 37°C, *overnight*, e em seguida foram estriadas em placas de Petri contendo meio SS Ágar (*Salmonella-Shigella*), também incubadas a 37°C para isolamento. A seguir as colônias isoladas foram caracterizadas em galeria bioquímica para enterobactérias, sendo eles o meio EPM (hidrólise de ureia, desaminação do triptofano, H₂S e produção de gás), meio MILi (motilidade, prova de indol e descarboxilação de lisina) e o meio Citrato de Simons, para detectar a sua utilização ou não pela bactéria.

4.3. Extração de DNA genômico e amplificação

As cepas de *Shigella* tiveram seus DNAs extraídos utilizando uma metodologia sem o uso de fenol e clorofórmio, pois ambos podem interferir no sequenciamento. A extração usa 300µL de uma cultura crescida de bactérias, em LB líquido, e é baseada no '*salting out*', que precipita proteínas em solução de alta concentração de sais, no caso, o Acetato de Potássio (POP da metodologia anexado no Suplemento)

Após esse procedimento, foi realizado a reação em cadeia da Polimerase (PCR), com iniciadores específicos que auxiliam na identificação dessas espécies. Foram utilizados genes específicos para genotipagem de *Shigella*, utilizando uma metodologia proposta por OJHA (2013). Os *primers* utilizados foram os mesmos propostos no artigo e já haviam sido padronizados e utilizados em trabalhos anteriores.

4.4. Isolamento do DNA e Sequenciamento

Bibliotecas de DNA genômico foram construídas utilizando o kit *NEXTERA XT DNA Sample Preparation* (Illumina, San Diego, CA, USA) e foram sequenciadas na Plataforma de Sequenciamento de Ácidos Nucléicos de Nova Geração, no IOC (Instituto Oswaldo Cruz), Rio de Janeiro, Brasil, no equipamento IlluminaHiSeq 2500, de acordo as instruções do fabricante.

4.5. Filtragem e Montagem

A qualidade das *reads* foi avaliada por meio da ferramenta FASTQC v0.11.6 (www.bioinformatics.babraham.ac.uk/projects/fastqc/) e filtradas utilizando a ferramenta PRINSEQ (SCHMIEDER; EDWARDS, 2011).

A montagem dos *contigs* foi realizada utilizando o *software* Velvet 1.2.10 (ZERBINO; BIRNEY, 2008), através da abordagem *de novo*, onde não é utilizado um genoma de referência, e *scaffolds*, utilizando o programa ABACAS 1.3.1 (*Algorithm Based Automatic Contiguation of Assembled Sequences*) (ASSEFA et al., 2009), utilizando os *contigs* obtidos e genomas de referência disponíveis no banco de dados Genbank, do NCBI, *Nacional Center for Biotechnology Information*.

O ABACAS é um *script* em linguagem de programação *perl*, com o qual foram montados os *scaffolds* dos cromossomos. Para essa montagem, foram utilizados genomas de referência.

Os genomas usados como referência para a montagem no ABACAS foram selecionados a partir da comparação dos maiores *contigs* da montagem *de novo* utilizando o programa BLASTn (*Basic Local Alignment Search Tool*) e o banco de dados GenBank e selecionados os que obtiveram a maior porcentagem de nucleotídeos idênticos (Tabela 1).

Antes da montagem final, foram feitas diversas montagens com outras cepas do banco de dados, inclusive com *Shigella dysenteriae* e *Shigella sonnei*, até obter uma montagem com um bom resultado, onde o número de nucleotídeos foi muito similar a uma *Shigella* padrão (excluindo nucleotídeos indefinidos, N). Os nucleotídeos não utilizados são armazenados em um arquivo.

Tabela 1: Genomas do Genbank utilizados para a montagem de referência pelo ABACAS.

Cepa	Nome	Número de Acesso
837	<i>Shigella flexneri</i> 1a strain 0439	NZ_CP020342
36P	<i>Shigella flexneri</i> 1a strain 0439	NZ_CP020342
192	<i>Escherichia coli</i> strain H7	CP010171
183	<i>Escherichia coli</i> CFSAN029787	CP010171
113	<i>Shigella flexneri</i> 1a strain 0439	NZ_CP020342

80	<i>Escherichia coli</i> 09-00049	CP015228.1
201	<i>Shigella flexneri</i> 1a strain 0439	NZ_CP020342

Os *scaffolds* montados pelo ABACAS foram visualizados com ferramenta a CGView, uma plataforma *online* que gera uma imagem circular do genoma. A montagem foi analisada pelo MUMmerplot, também instalado, que tem como entrada o cromossomo montado (STOTHARD; WISHART, 2005).

A porcentagem de GC foi encontrada a partir de um script em Shell, assim como o valor de N50, encontrado a partir da aplicação de um *script* AWK.

4.6. Anotação Gênica

Para a anotação foi utilizado o pipeline Prokka (SEEMANN, 2014), um conjunto de ferramentas que resulta em uma anotação rápida para genomas bacterianos (SEEMANN, 2014). Esse pipeline foi instalado na própria máquina e executado no terminal do Linux.

Posteriormente, uma anotação manual também foi realizada para complementar a anotação automatizada, na qual criou-se um banco com os principais genes de virulência de *Shigella*. Esse banco foi comparado e utilizando o programa BLASTx, em um script rodado na linha de comando do computador com os genes preditos, e com isso foram identificadas as sequências entre os genes preditos anteriormente, sendo então acrescentados à anotação do Prokka. Os resultados foram filtrados para não ficarem em duplicidade na hora da adição à anotação automática.

Os rRNAs e tRNAs foram preditos pela utilização do RNAmmer 1.2 (LAGESEN et al., 2007) e tRNAscan-SE 1.21 (LOWE; CHAN, 2016), e posteriormente feitos alinhamentos contra outros genomas da mesma espécie, utilizando o programa BLASTx e o banco de dados Swiss-Prot, para análise de identidade. As sequências de RNA de 16S predito pelo RNAmmer foram submetidas à análise de homologia em dois bancos de dados: o RDP Classifier (Ribosomal Database Project) (WANG et al., 2007) e o SILVA (*High Quality Ribossomal RNA Databases*), através do alinhador SINA (*SILVA Incremental Aligner*) (PRUESSE; PEPLIES; GLÖCKNER, 2012). Juntamente

com o resultado do BLAST foi possível identificar, a nível de espécie, as cepas sequenciadas.

4.7. Análise de vias metabólicas

Uma nova anotação também foi gerada através do KAAS (*KEGG Automatic Annotation Server*)(MORIYA et al., 2007), que gera uma caracterização funcional aos genes por meio de uma anotação automática de ortólogos, utilizando o método de atribuição SBH (*single-directional best-hit*).O KAAS atribui um KO (*KOGG Orthology*), uma identificação em um banco de dados curado que armazena *funções moleculares de genes e proteínas, permitindo uma comparação com o banco de dados do KEGG (Kyoto Encyclopedia of Genes and Genomes)*(KANEHISA et al., 2016).

Foram selecionados três mapas, disponíveis no *KEGG Pathway Maps* que foram gerados com o pacote Pathview (<https://bioconductor.org/packages/release/bioc/html/pathview.html>) no ambiente R: KO 05131, *Shigellosis*; KO 03070, *Bacterial Secretion System*; e o KO 05100, *Bacterial Invasion of Epithelial Cells*.

Para essa análise, cepas padrões também foram utilizadas. A sequência codificadora de proteínas foi obtida do banco de dados do NCBI, tanto do cromossomo, quanto do plasmídeo, tendo em vista a necessidade de identificar os sistemas de secreção, e alguns deles são localizados no plasmídeo, como o SST3. A Tabela 2 descreve as cepas utilizadas, assim como seus respectivos números de acesso.

Tabela 2: Cepas utilizadas como referência para a montagem dos cromossomos com o ABACAS, bem como seu número de acesso do cromossomo e plasmídeos utilizados.

Cepa	Nº acesso cromossomo	Nº acesso plasmídeo
<i>Shigella flexneri</i> 2a str. 301	NC_004337.2	NC_004851.1
<i>Shigella flexneri</i> 2002017	NC_004337.2	NC_004851.1
<i>Shigella boydii</i> Sb227	CP000036.1	CP000037.1 CP001058.1

<i>Shigella boydii</i> CDC 3083-94	CP001063.1	CP001062.1
		CP001059.1
		CP001060.1
		CP001061.1

4.8. Genômica comparativa

Foram gerados diversos diagramas de Venn, utilizando o pacote *Venn Diagram* no R. Esses diagramas oportunizam uma análise entre as diferenças gênicas, permitindo obter a informação de quantos genes as cepas compartilham entre si e quais os genes específicos em cada uma delas. Além disso, um alinhamento múltiplo de genomas foi realizado através do *software* MAUVE, que, além de alinhar, representa eventos evolutivos como rearranjos e inversão entre cromossomos.

Para a construção da árvore filogenômica foi utilizado o *pipeline* automático Orthologsorter desenvolvido pela FACOM/Universidade Federal do Mato Grosso do Sul (Faculdade de Computação), que tem como produto final árvores filogenéticas a partir da clusterização de sequências (FARIAS, 2013). Esse pipeline é baseado na comparação de sequências de proteínas preditas de cada organismo a ser comparado, gerando dois *outputs* distintos: uma árvore gerada através do agrupamento de famílias proteicas, obtendo informações funcionais e evolutivas; e outra árvore gerada para inferência evolutiva das espécies. O pipeline se inicia com a identificação das proteínas por meio do BLAST, seguido de uma criação de conjunto ou famílias de proteínas pelo OrthoMCL. Essas proteínas são adicionadas em um banco de dados MySQL, e então alinhadas com o MUSCLE e Gblocks, que filtra as famílias. Esse resultado segue para o PhyML e NJplot, que constroem as árvores, utilizando 1000 repetições de Bootstrap.

4.9. Montagem dos plasmídeos

Para montar os plasmídeos foi utilizado o arquivo *unused contigs*, gerado pelo ABACAS, durante a montagem referenciada. Para as cepas de *Shigella boydii*, foi

utilizado como referência alguns plasmídeos baixados do banco de dados do NCBI (Tabela 3).

Tabela 3: Cepas utilizadas para a montagem dos plasmídeos das cepas identificadas como *Shigella boydii*, bem como seus respectivos códigos de acesso no NCBI

Plasmídeo	Número de acesso
<i>Shigella boydii</i> CDC 3083-94 plasmid pBS512_2	CP001058.1
<i>Shigella boydii</i> CDC 3083-94 plasmid Pbs512_211	CP001062.1
<i>Shigella boydii</i> Sb227 plasmid Psb4_227	CP000037.1

Já para a montagem dos plasmídeos de *Shigella flexneri* foram utilizadas outras referências, também armazenadas no NCBI (Tabela 4).

Tabela 4: Cepas utilizadas para as montagens dos plasmídeos das cepas de *Shigella flexneri*, bem como o número de acesso de seus respectivos.

Plasmídeo	Número de acesso
<i>Shigella flexneri</i> 2a str. 301 plasmid pCP301	NC_004851.1
<i>Shigella flexneri</i> 2002017 plasmid pSFxv_1	CP001384.1
<i>Shigella flexneri</i> G1663 plasmid pG1663	CP007038.1
<i>Shigella flexneri</i> 0670_plasmid_unnamed1	CP020087.1
<i>Shigella flexneri</i> 1602_plasmid_unnamed1	CP020337.1

5. RESULTADOS E DISCUSSÃO

5.1. Isolamento e identificação das cepas de *Shigella*

As amostras isoladas passaram por análises de microbiologia básica, como testes bioquímicos (Tabela 5), no qual é importante ressaltar que as mesmas apresentaram algumas diferenças que dificultaram a identificação, levando essas cepas a serem escolhidas para o sequenciamento. As cepas 192, 113 e 201, possuem a produção de indol positivo, e as cepas 183 e 80, possuem descarboxilação da lisina positivo, dois fatores que para o gênero *Shigella* são negativos (ZAGHLOUL et al., 2007). Outro aspecto diferente do comum, observado na amostra 183, foi a fermentação de glicose. A glicose e indol positivos são características bioquímicas de *Escherichia coli*.

Tabela 5: Galeria bioquímica das cepas sequenciadas

Bioquímicos	SH 837	SH 36P	SH 192	SH 183	SH 113	SH 80	SH 201
Citrato	-	-	-	-	-	-	-
Lisina	-	-	-	+	-	+	-
Indol	-	-	+	-	+	-	+
H ₂ S	-	-	-	-	-	-	-
Gás	+/-	-	-	-	-	-	-
LTD	-	-	-	-	-	-	-
Glicose	-	-	-	+	-	-	-
Urease	-	-	-	-	-	-	-
Motilidade	-	-	-	-	-	-	-

5.2. Reação em cadeia da polimerase (PCR)

A técnica da PCR foi utilizada como forma adicional para auxiliar a identificação das cepas escolhidas através de uma reação pentaplex, onde 5 pares de iniciadores, amplificam genes de gênero e espécie. A Figura 6 mostra o resultado dessa reação,

onde, o primer ICD amplifica para o gene *ompA*, um controle positivo interno, com 1329pb, que amplifica para todas as bactérias.

O primer SGen amplifica para o gênero *Shigella*, identificando o gene *invC*. Já o SFlex amplifica o gene *rfc*, da *S. flexneri* e Sdv para o *rfpB*, exclusivo da *S. dysenteriae*.

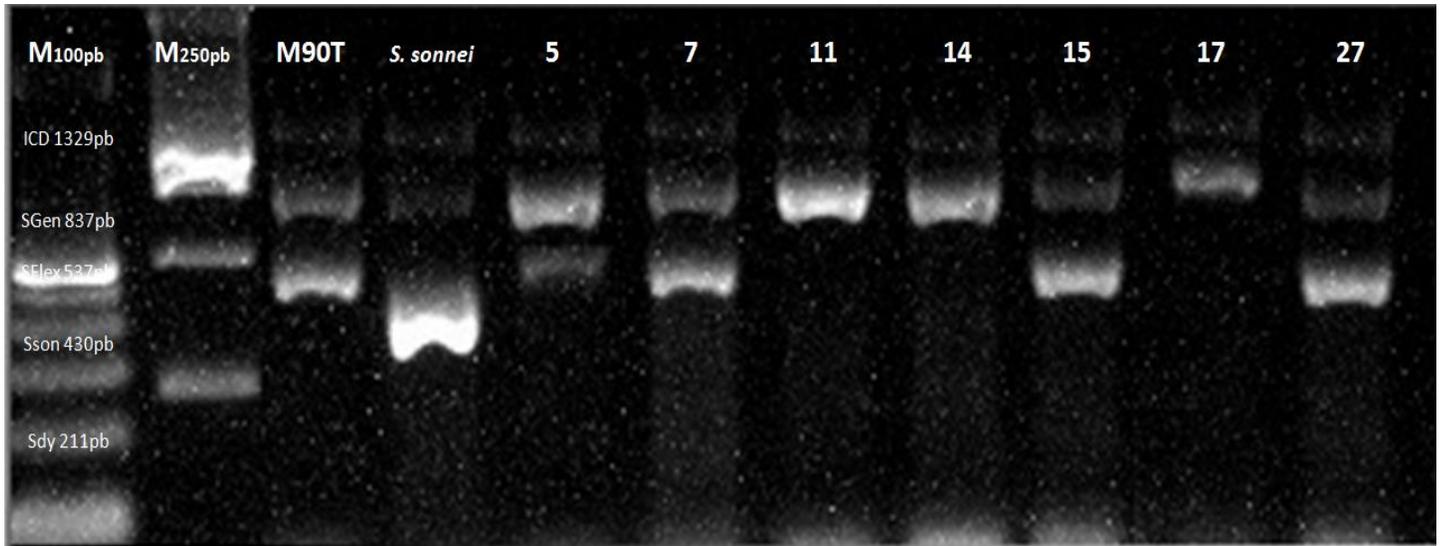


Figura 7: PCR de Genotipagem; a linha de cima representa as 11 amostras, dois marcadores de 100pb e 250pb, respectivamente, as amostras subsequentes são uma *Shigella flexneri* padrão, M90T, *S. sonnei* padrão, e as amostras 5, 7, 11, 14, 15, 17 e 27. A coluna da esquerda mostra o tamanho dos genes a serem amplificados, ICD, controle da reação, SGen para o gênero, Sflex para *S. flexneri* e Sdv para *S. dysenteriae*

Tabela 6: Análise quantitativa da PCR de Genotipagem
Resultado PCR Gênero-Espécie

	ICD	Sgen	Sflex	Sson	Sdy	Resultado
M90T	+	+	+	-	-	<i>S. flexneri</i>
05	+	+	+	-	-	<i>S. flexneri</i>
07	+	+	+	-	-	<i>S. flexneri</i>
11	+	+	-	-	-	<i>S. boydii</i>

14	+	+	-	-	-	<i>S. boydii</i>
15	+	+	+	-	-	<i>S. flexneri</i>
17	+	+	-	-	-	<i>S. boydii</i>
27	+	+	+	-	-	<i>S. flexneri</i>

5.3. Sequenciamento, montagem e anotação

Para uma melhor caracterização as amostras seguiram para o sequenciamento completo do genoma.

A etapa de análise de qualidade das sequências é essencial ser realizada antes de prosseguir com trabalho posterior. A qualidade é dada através da escala Phred, em forma de gráficos. Utilizando a escala de Phred, de 0 a 40 - onde quanto maior a pontuação, melhor a qualidade - podemos inferir que o sequenciamento foi de boa qualidade, uma vez que grande parte dos fragmentos obtidos possuíam uma pontuação alta (Figura 7).

Apesar de o conjunto de *reads* R2, que representa a sequência 3' - 5' (*reverse*) do sequenciamento, ter ficado com qualidade menor, esse é um resultado esperado. Há uma limitação da tecnologia nos sequenciamentos por síntese, contudo a qualidade considerando o conjunto total de *reads* foi bastante elevada.

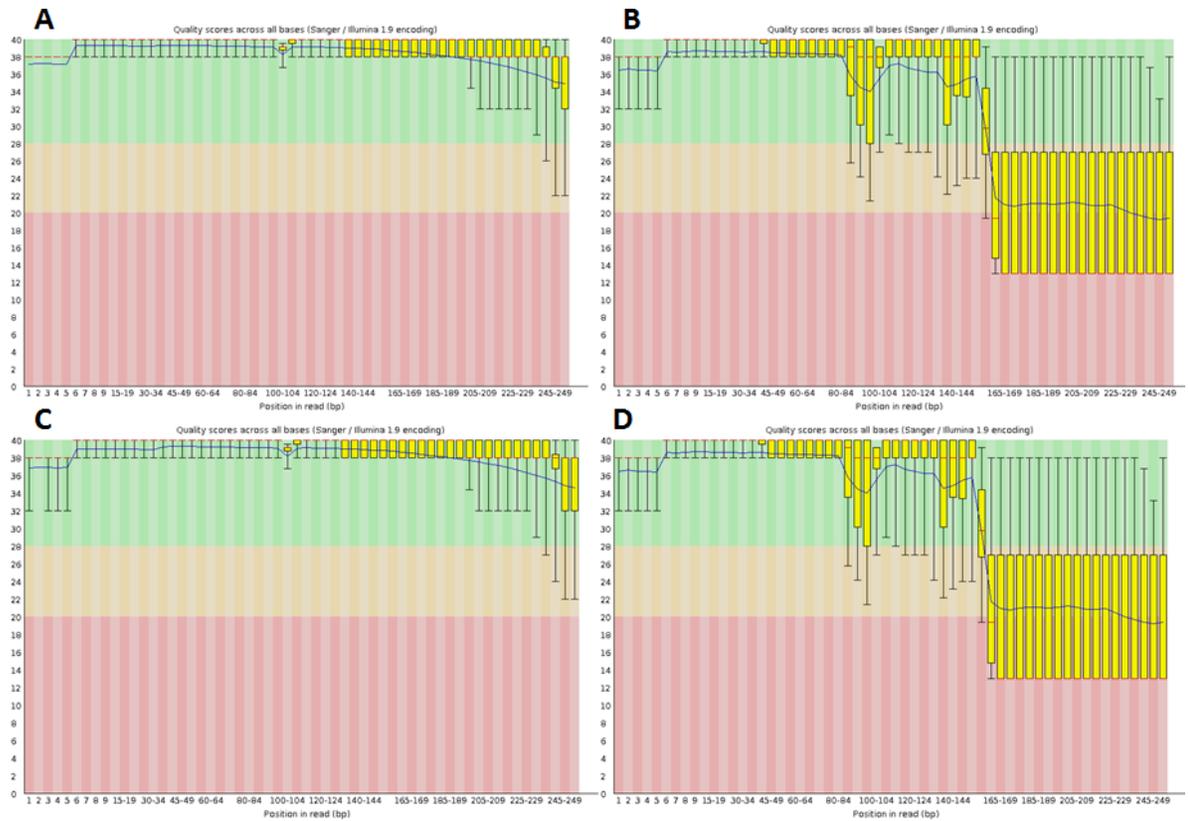


Figura 8: Pontuação Phred de cada base dos fragmentos sequenciados. As figuras A e B são da amostra 837, enquanto C e D são da amostra 36P. Sendo as duas da direita, as reads R1 e as subsequentes, reads R2. A linha azul representa a qualidade média das *reads*, enquanto a linha vermelha representa a mediana da qualidade.

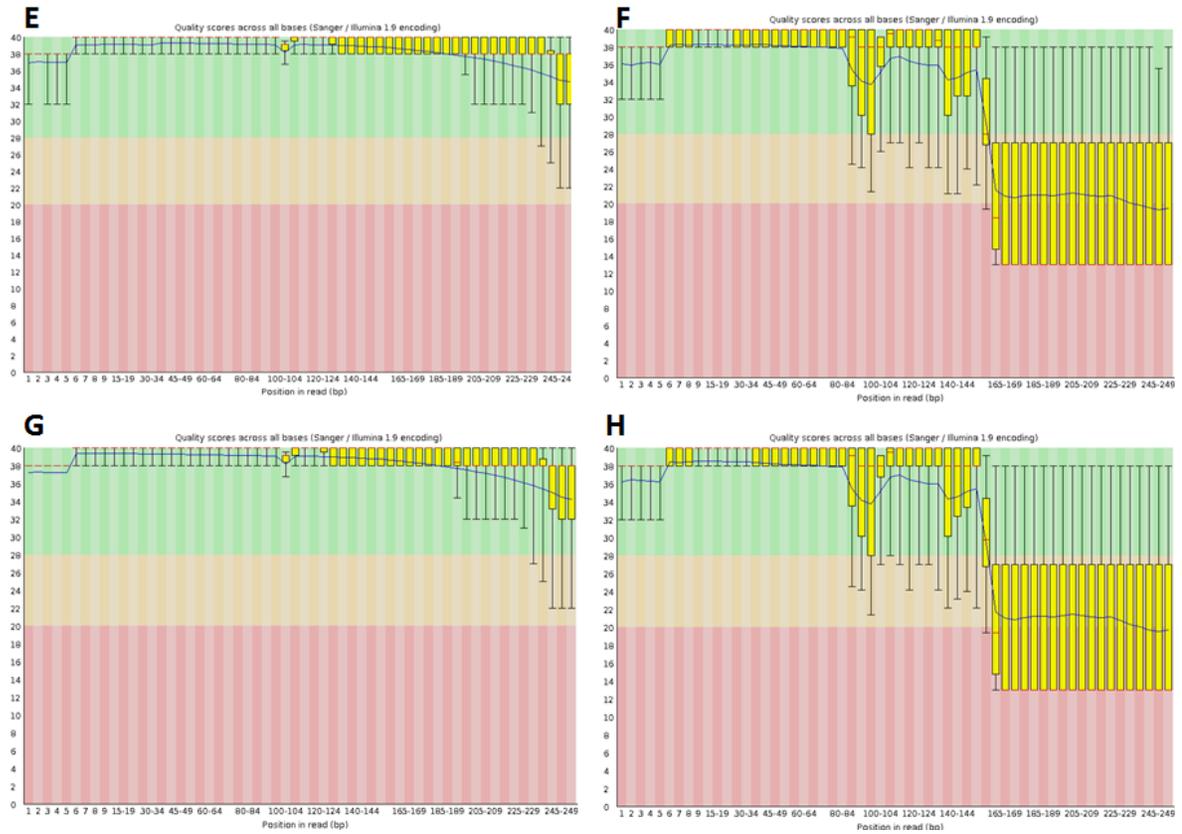


Figura 9: Pontuação Phred de cada base dos fragmentos sequenciados. As figuras E e F são da amostra 192, enquanto G e H são da amostra 183. Sendo as duas da direita, as reads R1 e as subsequentes, reads R2. A linha azul representa a qualidade média das *reads*, enquanto a linha vermelha representa a mediana da qualidade.

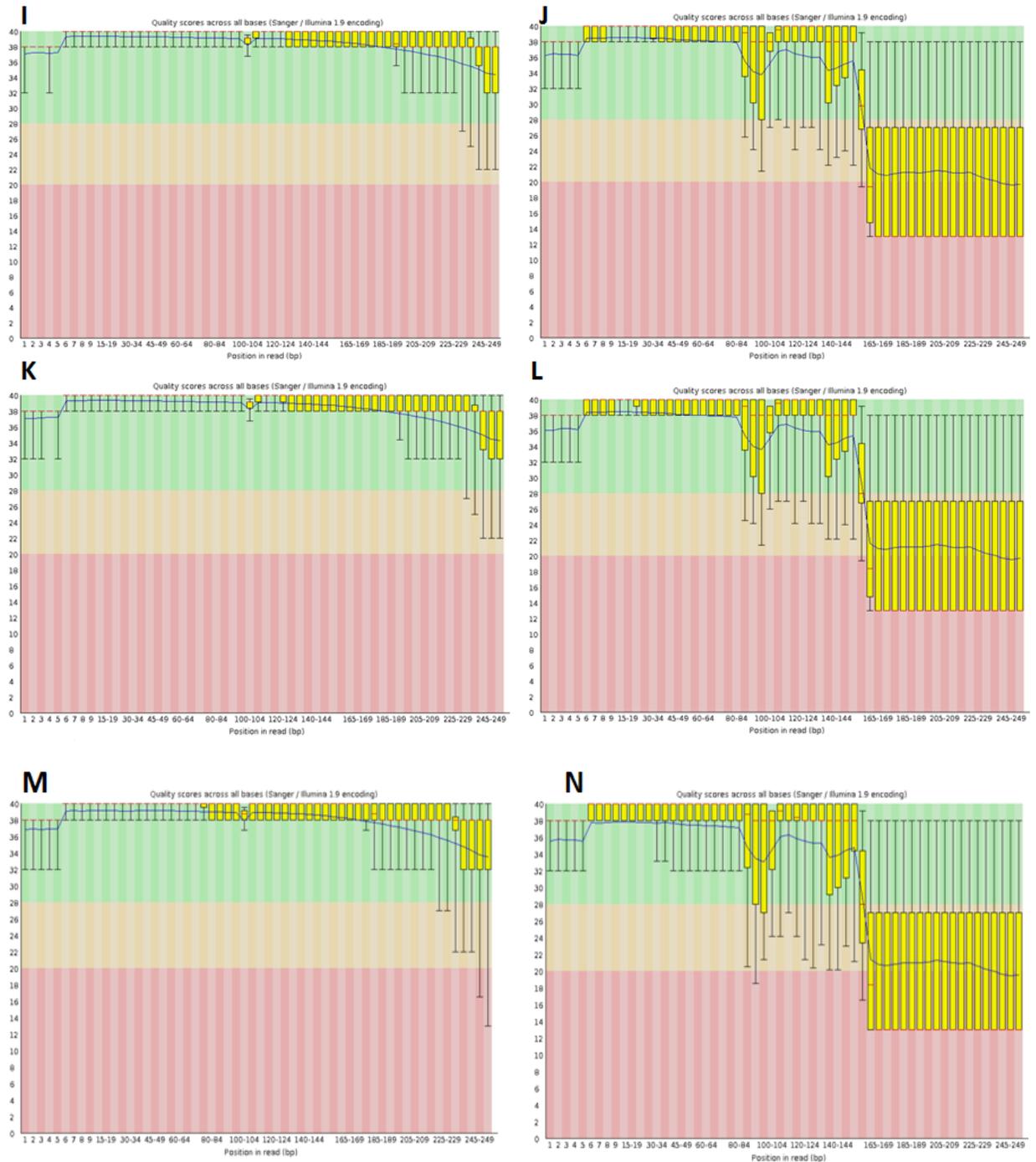


Figura 10: Pontuação Phred de cada base dos fragmentos sequenciados. As figuras I e J são da amostra 113, K e L são da amostra 80, e M e N da amostra 201. Sendo as três da direita, as reads R1 e as subsequentes, reads R2. A linha azul representa a qualidade média das *reads*, enquanto a linha vermelha representa a mediana da qualidade.

Apesar do número de *reads* sequenciadas ter variado bastante - entre 1001805 na cepa 36P e 4983112 para a amostra 192 - o número de genes preditos foi similar e anotados foi acima de 70% para todas as amostras (Tabela 7). A anotação manual feita com o BLAST mostra a relevância de fazer uma acurácia da anotação feita

Tabela 7: Estatísticas globais do sequenciamento

	Nº				Genes	Genes	Anotação
	Nº reads	Contigs	N50	%GC	Preditos	Anotados	Manual
36P	1001805	338	33718	50,40%	4440	3357	39
837	2076358	320	33633	50,60%	4380	3318	38
201	3093765	330	35100	50,60%	4459	3337	52
192	4983112	262	63158	50,70%	4935	3558	50
183	1989616	259	63158	50,70%	4902	3554	45
113	3988632	340	35069	50,50%	4449	3341	57
80	4845488	319	37836	50,50%	4909	3452	51

A montagem dos *scaffolds* dos cromossomos foi possível com a utilização do programa ABACAS, utilizando genomas de referência, sendo visualizado pelo CGView (Figura 10). A linha mais externa do cromossomo circular representa a quantidade de GC, enquanto a linha interna, representa o GC Skew, ou a inclinação de GC, que prediz a origem e término da replicação.

O MUMmerplot consegue indicar as regiões do cromossomo que conseguiram ser mapeadas com o genoma de referência. Diversas montagens foram feitas, utilizando genomas diferentes, inclusive de espécies diferentes. Essa plotagem representa a melhor montagem obtida, e é possível notar a quantidade de *gaps* encontrados, principalmente na amostra 80 (Letra F, Figura 10), o que nos leva a crer no quão diferente essa cepa é de *uma Shigella boydii* convencional.

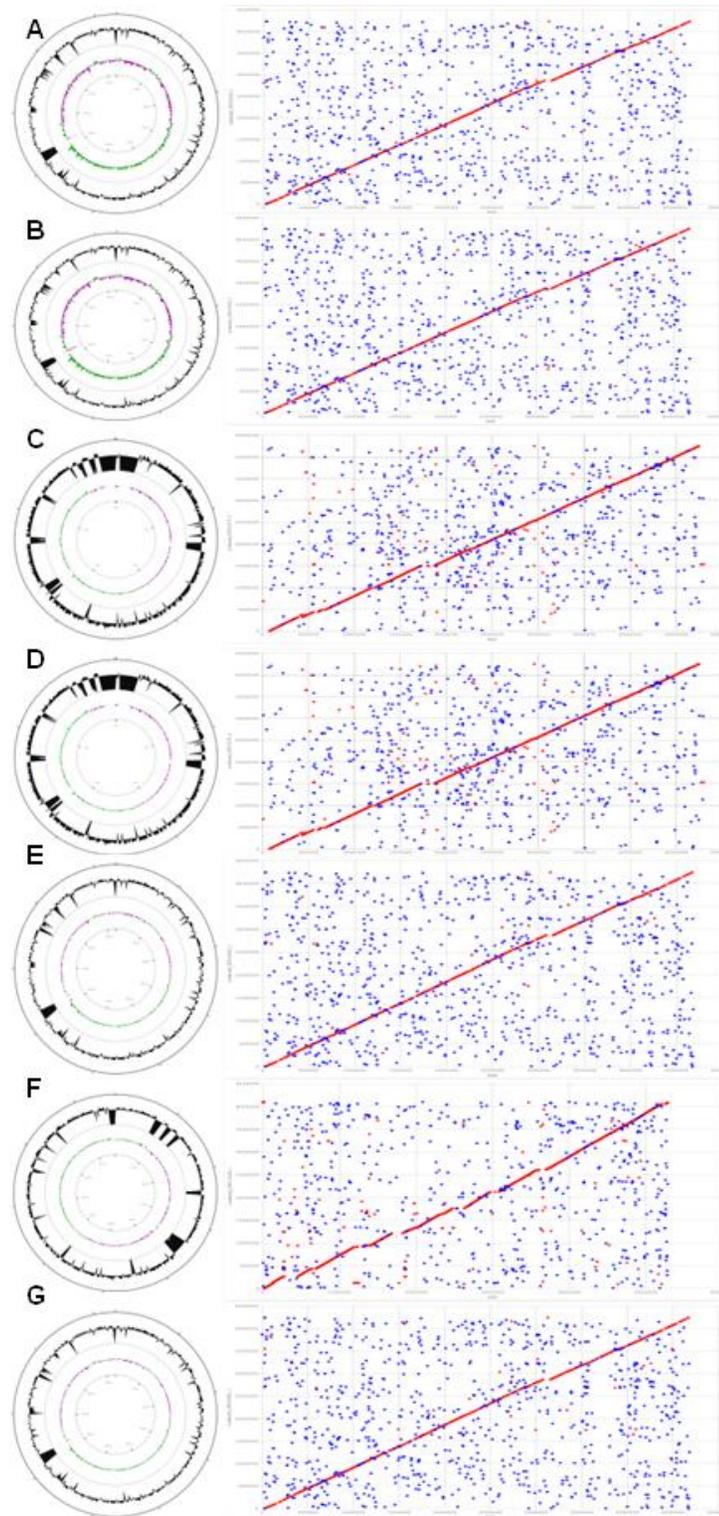


Figura 11: Representação circular dos cromossomos montados pelo ABACAS e visualizados pelo CGView Server. Seguindo a ordem alfabética, eles representam as cepas: 837, 36P, 192, 183, 113, 80 e 201. Ao lado do cromossomo é o alinhamento gerado pelo MUMerplot, que analisa o pareamento dos *contigs* com o genoma de referência

5.4. Classificação

Para uma primeira classificação a nível de espécie das amostras, foi utilizada a sequência do 16S comparado a diferentes bancos de dados: RDP, SILVA e BLAST. O resultado foi inconclusivo para as cepas 192 e 183 (Tabela 6).

Tabela 8: Tabela comparativa entre os bancos de dados utilizados para classificação das amostras

	RDP	Silva	Blast	Identidade do BLAST (%)
837	<i>Escherichia/Shigella</i>	<i>Escherichia/Shigella</i>	<i>Shigella flexneri</i>	100,00%
36P	<i>Escherichia/Shigella</i>	<i>Escherichia/Shigella</i>	<i>Shigella flexneri</i>	100,00%
192	<i>Escherichia/Shigella</i>	-	-	-
183	<i>Escherichia/Shigella</i>	-	-	-
113	<i>Escherichia/Shigella</i>	<i>Escherichia/Shigella</i>	<i>Shigella flexneri</i>	100,00%
80	<i>Escherichia/Shigella</i>	<i>Escherichia/Shigella</i>	<i>Escherichia coli</i>	100,00%
201	<i>Escherichia/Shigella</i>	<i>Escherichia/Shigella</i>	<i>Shigella flexneri</i>	100,00%

Mesmo após anos de descoberta do gênero *Shigella*, ainda há dificuldades na diferenciação entre *Shigella* e *E. coli*, tanto, que dos três bancos utilizados, dois classificam as bactérias como *Escherichia/Shigella*, ou seja, não são capazes de diferencia-las somente com a sequência ribossomal do 16S.

5.5. Genômica Comparativa

As quatro cepas de *S. flexneri* compartilham entre si 3141 genes, sendo dessas, a cepa 36P a que mais difere, com 25 genes exclusivos. Já entre as cepas de *S. boydii*, que são 3, somente a 192 e a 183 compartilham entre si 3433 genes. Comparando essas duas últimas, a 192 e 183, com a cepa 80, que também é *S. boydii*, elas partilham 2959 genes e 376 exclusivos, tornando, portanto, *S. boydiistr.* 80 a que mais difere entre as demais seqüenciadas (Figura 11).

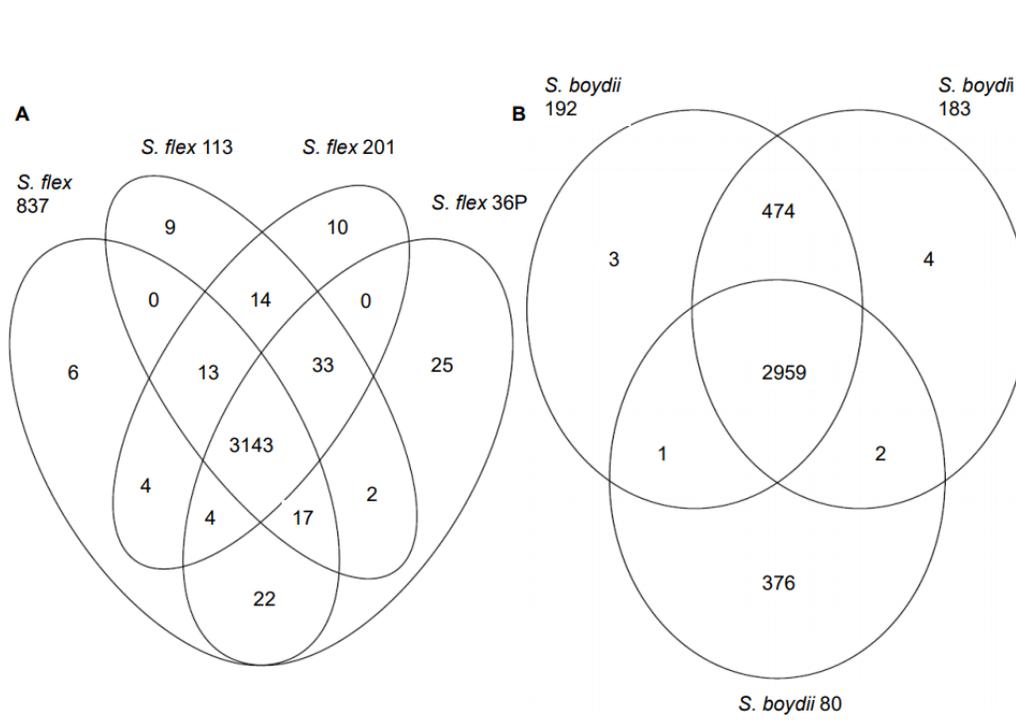


Figura 12: Diagrama de Venn para a espécie *Shigella flexneri*; B: Diagrama de Venn com as representantes da espécie *Shigella boydii*. O número maior, de ambas as imagens, representa o número de genes compartilhados entre todas amostras do grupo, já os valores menores os genes compartilhados entre duas amostras ou genes exclusivos

Quando comparado os genomas de *S. flexneri* com um genoma de *S. boydii*, a diferença chegou a quase 650 genes distintos, e essa diferença pôde ser também observada em um alinhamento múltiplo de genomas, utilizando o programa MAUVE.

A Figura 3 traz o resultado desse alinhamento e é patente a diferença entre os cromossomos. A parte inferior da imagem representa as quatro cepas de *S. flexneri*, que são muito semelhantes entre si. De maneira geral, pode se dizer que os genomas de *S. flexneri* possuem um alto nível de conservação estrutural, o que já havia sido provado em outros trabalhos (NIE et al., 2006). Porém bem distintos dos genomas de *S. boydii* (os 3 primeiros: 80, 192 e 183). Ainda dentro do quadro de *S. boydii*, a amostra 80 (primeira) difere da própria espécie, corroborando o dado visto anteriormente, no diagrama de Venn, com um grande número genes exclusivos.

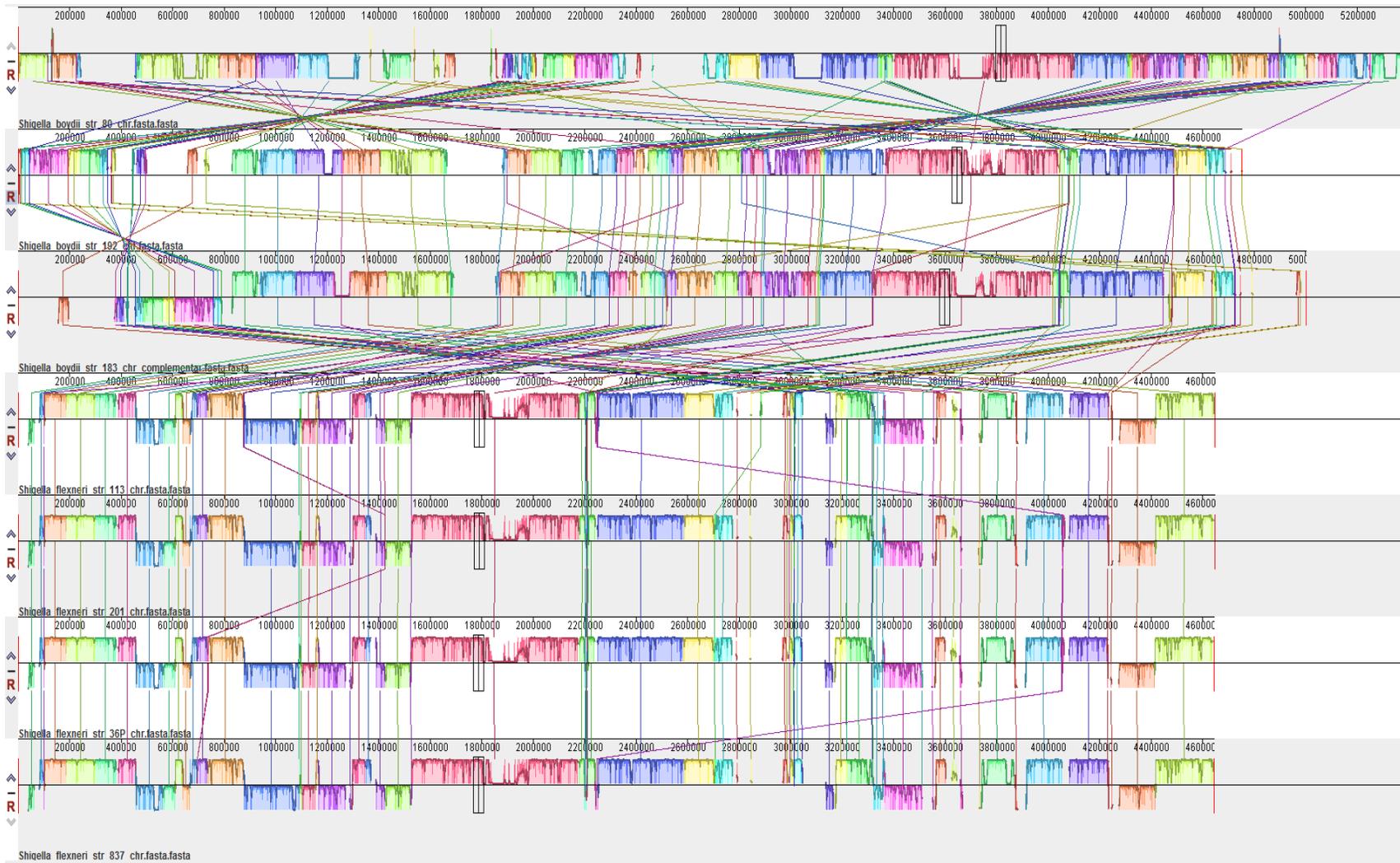


Figura 13: Alinhamento múltiplo dos genomas. De cima para baixo: as 3 amostras de *S. boydii*, seguido das 4 amostras de *S. flexneri*. É possível ver alinhamentos bem satisfatórios entre as cepas de *S. flexneri*, enquanto quando alinhados *S. flexneri* com *S. boydii* já é possível identificar inúmeras inversões, da mesma maneira dentro do nicho das *S. boydii* que mesmo sendo da mesma espécie possuem diferenças entre si.

Essas grandes diferenças podem estar relacionadas com a grande quantidade de pseudogenes encontrados. Em média, cerca de 15% destes genomas foram anotados como pseudogenes, o que dificulta o estudo, pois o esperado era pelo menos metade disso (NIE et al., 2006).

O número de elementos transponíveis também foi elevado, ultrapassando os 10% do genoma. Contudo, essa é uma característica do genoma de *Shigella*, sendo este um grande gargalo na diferenciação com *Escherichia coli*. O genoma da *Shigella* tem uma plasticidade, ou seja, uma facilidade de receber genes através da transferência horizontal de genes. Esses elementos recebidos também podem ser capazes de causar vários rearranjos genéticos como deleções ou inversões e translocações, tais como foi observado na Figura 12 (DI MARTINO et al., 2013).

5.6. Análises Filogenéticas

É sabido que a *Shigella* é bastante similar a outras bactérias, como a *Escherichia coli*, e também compartilha características com a *Salmonella*. Portanto, o pipeline Orthologsorter foi aplicado com algumas cepas escolhidas por similaridade, resultando em uma árvore filogenômica (Figura 13).

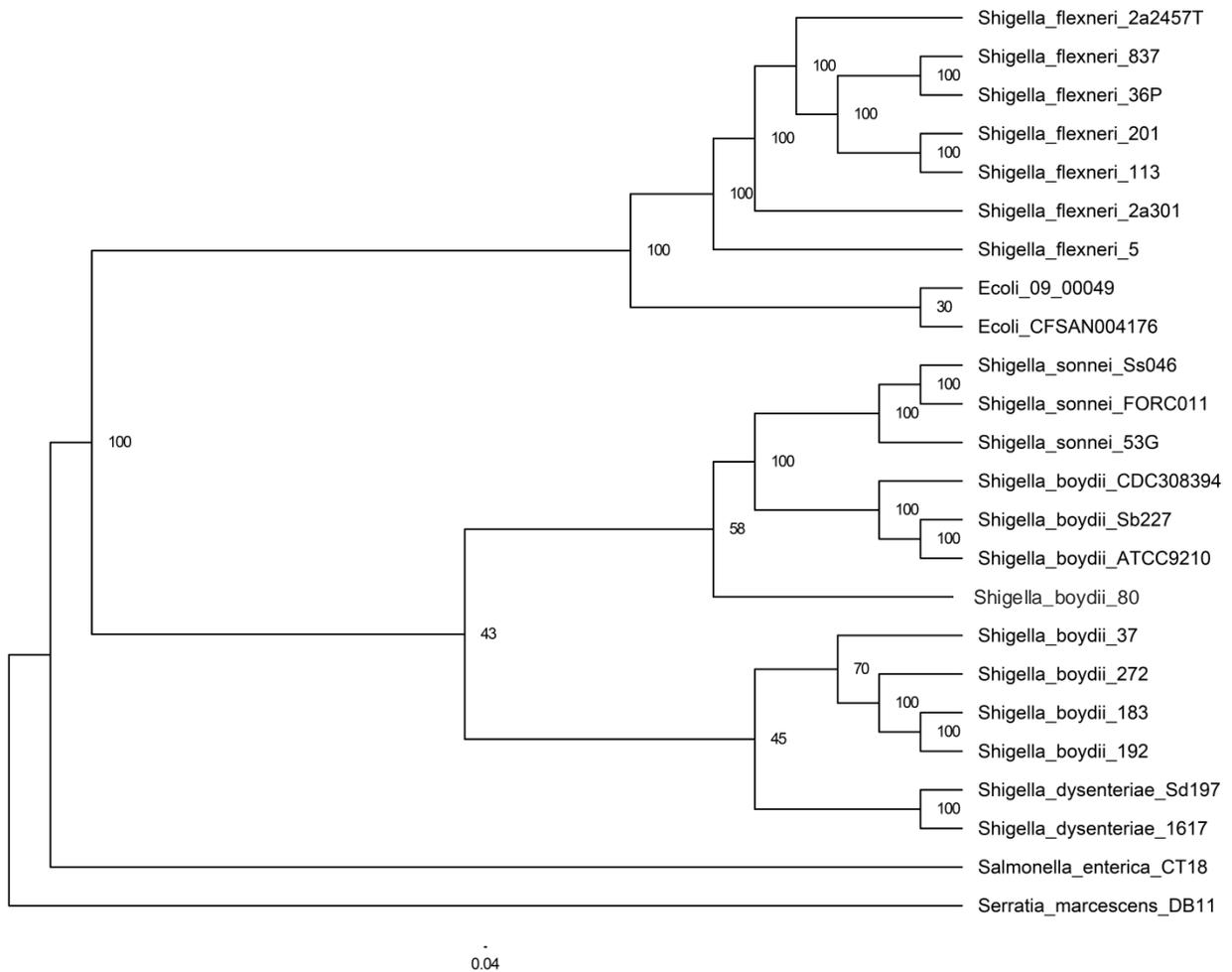


Figura 14: Árvore filogenômica gerada pelo pipeline Orthologsorter. As amostras com asterisco após o nome são as sequenciadas neste trabalho, utilizando 1000 repetições de Bootstrap

Nenhuma das amostras se agruparam com as cepas de *E. coli* ou outras enterobactérias, como *Salmonella* e *Serratia*. As quatro cepas de *S. flexneri* foram dispostas em um cluster juntamente com outras representantes da espécie, selecionadas do banco de dados Genbank do NCBI, já a cepa *S. boydii* str. 80 alocou-se em um cluster próxima da espécie *Shigella sonnei* e *Shigella boydii*, porém, na etapa de montagem com o genoma de referência, a cepa 80 exibiu maior similaridade ao genoma de *S. boydii* do que *S. sonnei*.

A aplicação da análise filogenômica foi essencial para a classificação da amostra *S. boydii* str. 80, pois o genoma desta se difere bastante de um genoma padrão de *S.*

boydii, porém, a diferença com os genomas de outras espécies é superior, levando a crer que ela é de fato uma cepa de *Shigella boydii*.

As demais cepas de *S. boydii*, 192 e 183, agruparam-se em outro *cluster*, mais próximas de outras sequências de *S. boydii* do banco de dados e da *Shigella dysenteriae*.

5.7. Sistemas de Secreção

Uma das principais características da *Shigella* é a expressão do Sistema de Secreção do Tipo 3. O T3SS é o grande responsável por contornar a imunidade inata do hospedeiro e causar a disenteria bacilar (ANDERSON et al., 2017b). Para verificar a presença deste e de outros sistemas de secreção, uma análise de ortólogos foi feita utilizando o programa KAAS e observada a partir de mapas do KEGG, mapas de vias metabólicas.

Das 7 amostras, foi possível verificar a presença do T3SS em 6 amostras, exceto na cepa de *S. flexneri* str. 837 (Figura 15). Contudo, estudos anteriores a este comprovam a presença de alguns dos genes do T3SS nesta cepa, o que nos leva a considerar algum erro no sequenciamento ou na montagem, impossibilitando a anotação devida desse sistema. O fato de o T3SS ter sido identificado nas amostras sequenciadas corrobora com a presença do plasmídeo nas amostras, tendo em vista que, esse sistema é localizado em um grande plasmídeo em *Shigella*, diferente de *Salmonella* spp. e *Yersinia* spp, onde a ilha de patogenicidade está no cromossomo (VONGSAWAN et al., 2015).

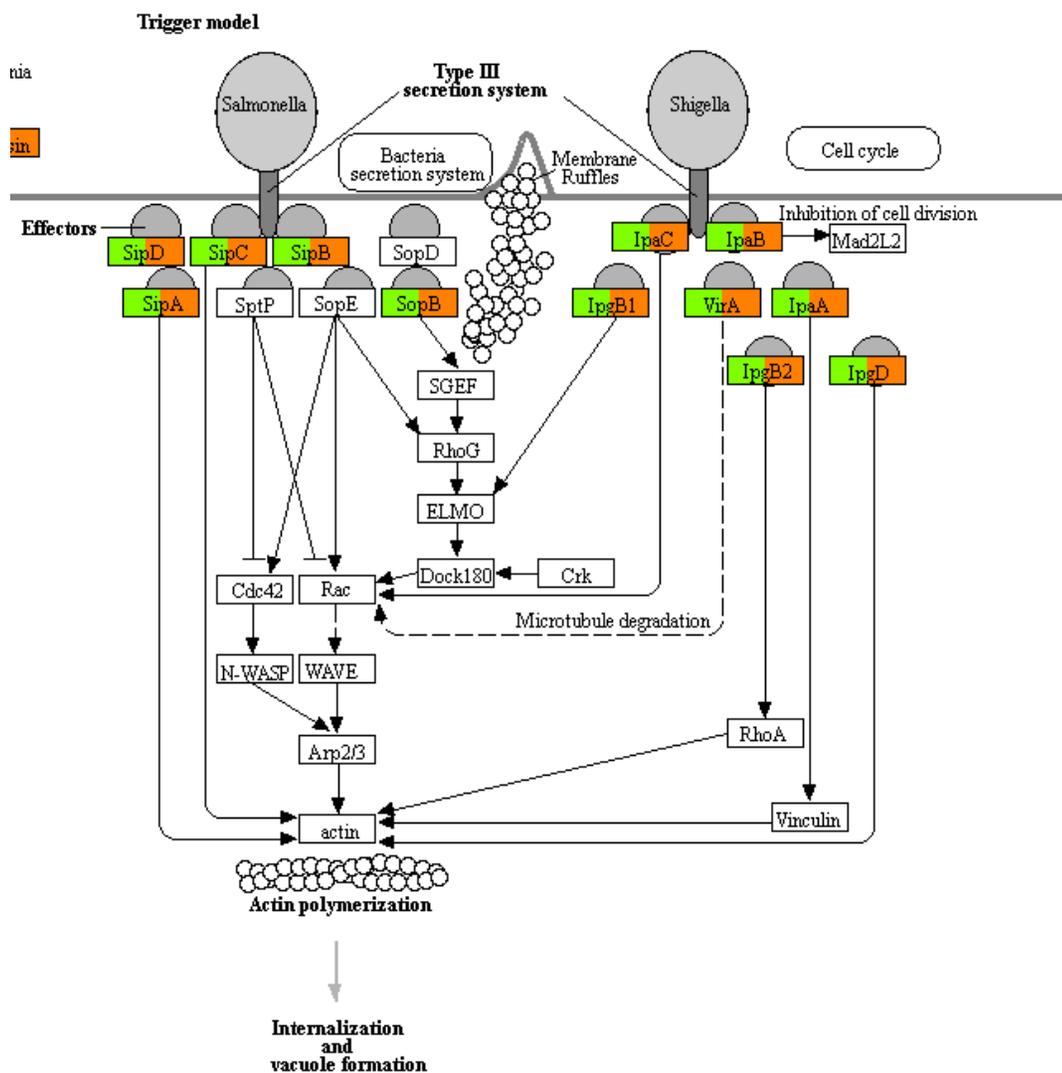


Figura 15: Mapa metabólico de KEGG das amostras de *Shigella boydii*, destacando os genes participantes do sistema de secreção do tipo 3, das amostras *S. boydii* str. 192 (m verde) e *S. boydii* str. 183 (laranja). Ambas as cepas possuem esse sistema completo. A cepa *S. boydii* str. 80 tem o mesmo padrão, contudo não está representada nessa imagem.

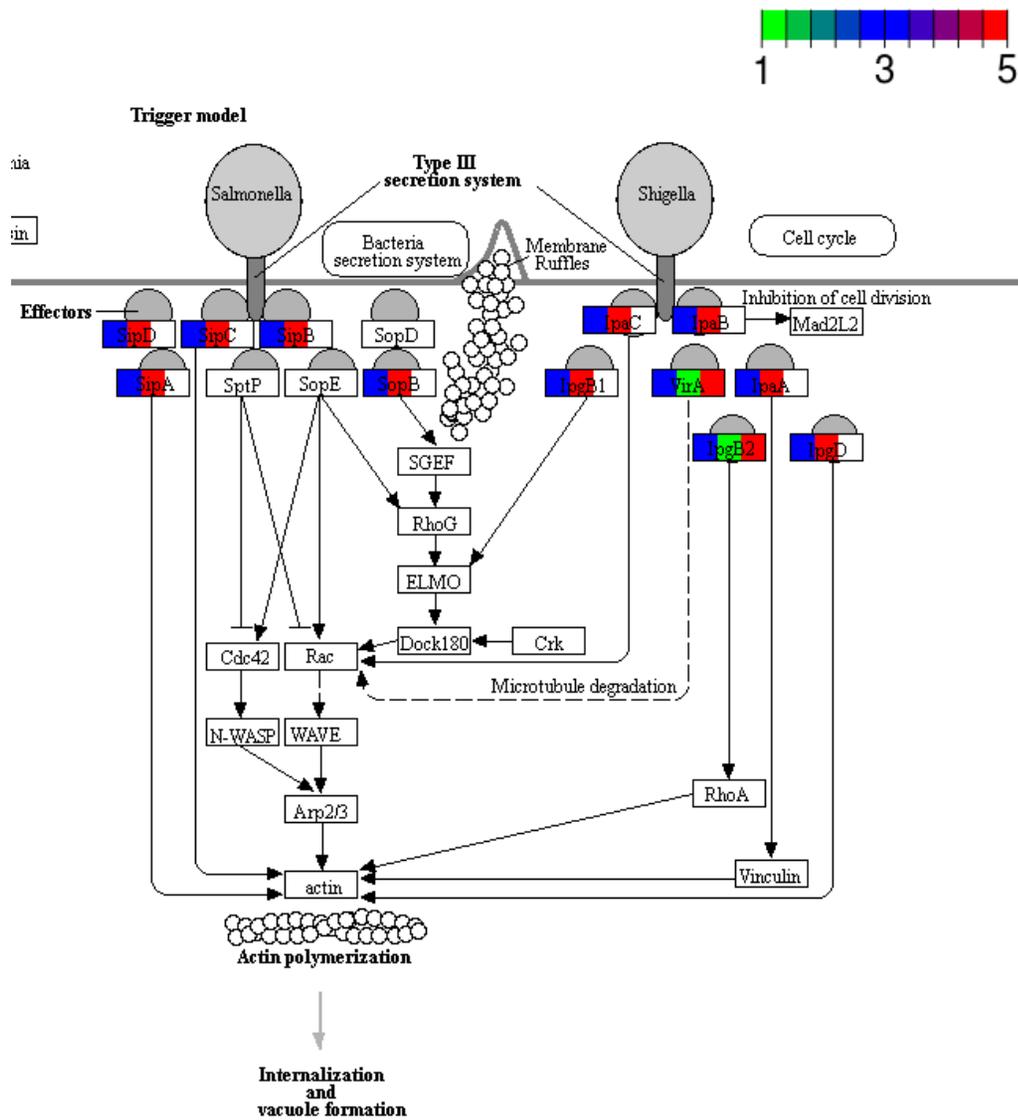


Figura 16: Mapa de KEGG representando a presença do sistema de secreção do tipo 3 nas cepas de *Shigella flexneri*. A cor em verde representa a *S. flexneri* str. 837; a azul, *S. flexneri* str. 36P; e a vermelha a cepa *S. flexneri* str. 113. A *S. flexneri* str. 201 não está representada na imagem, porém contém a mesma representação das cores azul e vermelho, portanto, o sistema de secreção do tipo 3 completo, diferente da 837.

Dentre os outros sistemas de secreção anotados está o T2SS, encontrado nas três *S. boydii* sequenciadas (Figura 14). São raros os relatos sobre a presença de T2SS em *Shigella*, somente em algumas cepas de *Shigella dysenteriae* e poucas cepas de *Shigella boydii*, tendo essas, um conjunto de genes bem semelhante ao T2SS de *Escherichia coli* Enterotoxigênica (ETEC) e *Vibrio cholerae*. Em 2005, YANG e

colaboradores descreveram esse sistema em *S. boydii*. Acredita que ele seja inativo nesta espécie devido a uma mutação no gene *gspD* e a um deslocamento em *gspC*.

Para a confirmação se esse sistema também é inativo em nossas amostras serão necessários estudos adicionais.

Nas amostras *S. boydii* 192, 183 e 80, 13 dos 12 genes que codificam proteínas do T2SS foram encontradas. A única ORF não identificada foi a GspS. Já na cepa de *Shigella boydii* istr. 227 (YANG et al., 2005) foram identificados 12 ORFS, enquanto na cepa de *S. dysenteriae* strain Sd1617, 11 ORFS foram anotadas (VONGSAWAN et al., 2015).

YANG e colaboradores, 2005, também trazem a informação de que o conjunto de genes *gsp* está localizado em uma região do cromossomo chamada de ilha *pheV*tRNA, podendo esta ser uma região de inserção de genes, facilitando a transferência horizontal de genes, fator observado na anotação deste trabalho, tendo algumas amostras contendo mais de 10% do genoma composto de elementos transponíveis.

Além disso, nas três cepas sequenciadas de *S. boydii* também foram anotados genes que codificam o SST6, sendo este, em *Shigella*, encontrado apenas na espécie *S. sonnei*. Estudos recentes mostram que o SST6 pode fornecer uma vantagem para a bactéria que possui esse sistema em uma competição dentro do intestino, principalmente contra a microbiota comensal do homem, como a *E. coli* (ANDERSON et al., 2017b).

A *S. flexneri* e *S. sonnei* são responsáveis pelo maior número de casos de disenteria no mundo, sendo a *S. flexneri* a espécie mais comum a nível mundial, porém, mais frequentemente isolada em países em desenvolvimento, enquanto a *S. sonnei* mais detectada em regiões mais desenvolvidas (ANDERSON et al., 2017a; THOMPSON; DUJ; BAKER, 2015). Entretanto, a *S. sonnei* vem emergindo nos países em desenvolvimento, como na Ásia, América Latina e Oriente Médio. O SST6 pode ser a resposta do porquê que o número de casos de shigelose com infecção por *S. sonnei* tem se igualado aos casos de *S. flexneri*, e no caso desse estudo, uma possível hipótese do aumento da incidência de *Shigella boydii*, que até então era mais comumente encontrada na Índia (DEKKER; FRANK, 2015).

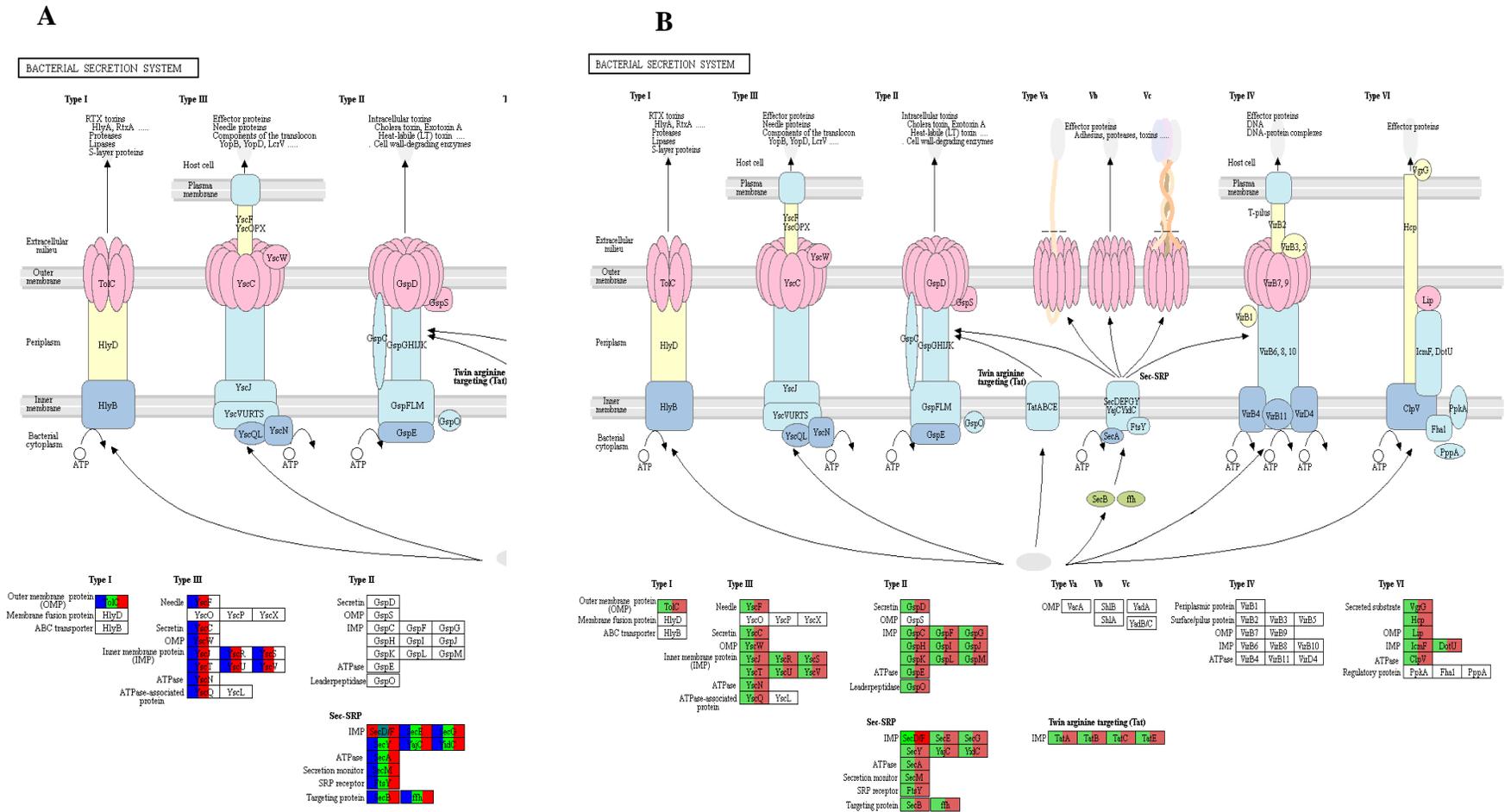


Figura 17: Mapa de KEGG representando a presença dos genes envolvidos nos sistemas de secreção bacteriano, entre eles, o tipo 1, tipo 3, tipo 4 e 6. A figura A trás, na ordem, os organismos 05, 07 e 15, *S. flexneri*, destacando a presença do T3SS na cepa 07 e 15. A figura B representa as cepas 14 e 11, que possuem genes envolvidos nos sistemas T3SS, T2SS e T6SS. A cepa 17, não representada nas imagens, possui um perfil idêntico ao das cepas 11 e 14.

5.8. Montagem dos Plasmídeos

A montagem dos plasmídeos não gerou resultados tão completos como a montagem dos cromossomos. É sabido que a *Shigella* possui um plasmídeo grande, chamado pINV, com cerca de 200kb. É nele que se localiza o sistema de secreção do Tipo 3 e com isso grande parte da virulência da bactéria. As Tabelas 9 e 10 resumem algumas das montagens realizadas dos plasmídeos utilizando referências.

Tabela 9: Montagens dos plasmídeos das cepas *Shigella boydii* str. 192 e *Shigella boydii* str. 183, utilizando 3 plasmídeos diferentes, depositados no banco de dados do NCBI.

Referência	Tamanho do plasmídeo	<i>S. boydii</i> str. 192		<i>S. boydii</i> str. 183	
		Com N	Sem N	Com N	Sem N
<i>Shigella boydii</i> CDC 3083-94 plasmid Pbs512_211	210919pb	195016	44410	217443	98563
<i>Shigella boydii</i> Sb227 plasmid Psb4_227	126697	112685	24090	130537	36029

Tabela 10: Montagem dos plasmídeos de *Shigella flexneri*, utilizando 5 plasmídeos diferentes como referência, todos encontrados no banco de dados do NCBI.

Referência	Nº de pb	<i>S. flexneri</i> str. 837		<i>S. flexneri</i> str. 36P		<i>S. flexneri</i> str. 113		<i>S. flexneri</i> str. 201	
		Com N	Sem N	Com N	Sem N	Com N	Sem N	Com N	Sem N
<i>S. flexneri</i> 2a str. 301 plasmid pCP301	221618	228937	97123	228379	139837	228241	130711	225311	151958
<i>S. flexneri</i> 2002017 plasmid pSFxv_1	223364	230712	97152	230155	139867	230016	130740	227086	151987

<i>S. flexneri</i>									
G1663									
plasmid	222586	229736	97136	229364	138403	229225	129276	225509	150510
Pg1663									
<i>S. flexneri</i> 1a									
0670									
plasmid_unn	228834	224244	93369	224244	93369	230319	130745	231297	152056
amed1									
<i>S. flexneri</i> 4c									
1602									
plasmid_unn	223354	219433	93369	219433	93369			225770	151966
amed1									

Outros plasmídeos como referência também foram utilizados, inclusive de outras espécies, como plasmídeos de *S. sonnei*, contudo os resultados apresentados foram os que produziram maior tamanho de sequência. Os dados de *S. boydii* são bem escassos, por isso, só foram montados utilizando 2 plasmídeos como referência, obtendo sequências de tamanhos reduzidos. A melhor montagem foi de um plasmídeo da cepa *S. boydii* str. 183 com o plasmídeo *Shigella boydii* CDC 3083-94 plasmid pBS512_211, obtendo 44% de similaridade. A cepa *S. boydii* CDC 3083-94 possui outro plasmídeo depositado, o plasmid pBS512_2, mas não apresentou similaridade nucleotídica com as cepas sequenciadas.

Os plasmídeos que obtiveram melhor montagem foram os montados da *S. flexneri* str. 201, que ultrapassaram os 150 mil pares de base, chegando até 68% com a referência *S. flexneri* 2a str. 301 plasmid pCP301.

As montagens obtidas das cepas 36P e 113 ficaram no intermédio, com cerca de 130 mil pares de bases similares às referências, enquanto a cepa str. 837 não chegou aos 100 mil pares de base, ou seja, menos de 50% dos nucleotídeos foram iguais. À essa última, esse resultado já era esperado tendo em vista que não foi identificado o sistema de secreção do Tipo 3 nesse isolado, e esse sistema é encontrado no

plasmídeo. O que pode ter ocorrido com essa cepa é um problema de sequenciamento, no qual o plasmídeo não foi englobado.

5.9. Análises clínicas

Ao observar a etiologia das bactérias é possível inferir inúmeras correlações sobre as cepas acima estudadas. A Tabela 11, abaixo resume algumas das principais características extraídas na coleta.

Dos sete isolados, cinco estavam com coinfeção por outros patógenos, que podem ser os causadores da diarreia. A *Shigella flexneri* str. 837 teve os sintomas clássicos de uma disenteria bacilar, com 6 dias de diarreia, febre, sangue visível e oculto nas fezes e muco, conseqüentemente, fezes líquidas. É importante observar as informações dadas pela mãe do bebê ao afirmar que tanto a higiene da criança, quanto a higiene com os alimentos são feitos habitualmente, o que leva corrobora o diagnóstico de shigelose, que tem sua transmissão dada de maneira fecal-oral, ou seja, através da ingestão e/ou contato de alimentos e água contaminados com fezes de outras pessoas contaminadas (DEKKER; FRANK, 2015; KOTLOFF, 2017).

De mesmo modo a criança de onde foi isolada a *Shigella flexneri* str. 201, no qual a criança também teve relatos de falta de higiene, assim como a limpeza dos alimentos na casa. Com apenas 3 meses, a criança estava com febre e fezes líquidas com muco. Todos sintomas clínicos de shigelose.

As outras representantes do gênero *Shigella flexneri*, a strain 36P e strain 113, não foram patógenos únicos no paciente, ambos estavam com ameba, causado pelo protozoário *Entamoeba histolytica*, que também causa uma colite aguda, similar a shigelose, com muco e sangue, por um ou dois dias, duração essa, semelhante a que os pacientes relataram estar com os sintomas (SANTOS; SOARES, 2008; SOUSA et al., 2013). Nesses casos não é possível afirmar qual o patógeno foi o causador da diarreia. É possível apenas afirmar que havia uma coinfeção, onde, pode haver uma soma dos sintomas, ou um deles prevalecer sobre o outro.

Tabela 11: Análise etio-epidemiológica dos pacientes no qual foram coletadas as amostras de sangue. Foram selecionados para esse trabalho 7 cepas, todas provenientes de uma coleta realizada na cidade de Manaus, em 3 hospitais diferentes. Nesse número amostral possuem crianças de idade entre 1 e 6 anos, onde foram coletadas informações como internações anteriores, condições de higiene da criança e dos alimentos ingeridos, além da sintomatologia da criança. Também há os resultados sobre a presença de uma co-infecção.

Cepa	Idade (anos)	Duração (Dias)	Internação Anterior	Higiene Da Criança	Higiene Com Alimento	Febre	Vômito	Sangue Visível	Consistência	Muco	PSO	Protozoários	Parasitas	Rotavirus
837	0.8	6	Negativo	Habitualmente	Habitualmente	Positivo	Positivo	Positivo	Líquido	Sim	Positivo	Negativo	Negativo	Negativo
36P	2.0	1	Negativo	As Vezes	As Vezes	Sim	Sim	Desconhece	Pastosa	Sim	Positivo	<i>Entamoebahy stolitica</i>	Negativo	Negativo
192	9.5	5	Positivo	Habitualmente	Habitualmente	Sim	Não	Não	Pastosa	Não	Negativo	Negativo	<i>Ascaris lumbricoides</i>	Positivo
183	9.1	2	Negativo	Habitualmente	Habitualmente	Sim	Sim	Não	Pastosa	Não	Negativo	Negativo	Negativo	Positivo
113	0.7	1	Negativo	Sempre	Sempre	Não	Sim	Desconhece	Pastosa	Sim	Positivo	<i>Entamoebahy stolitica</i>	Negativo	Negativo
80	4.2	3	Positivo	As Vezes	Sempre	Sim	Sim	Não	Líquida	Não	Positivo	Negativo	<i>Enterobiusve rmicularis</i>	Negativo
201	0.3	3	Negativo	As Vezes	As Vezes	Sim	Não	Não	Líquida	Sim	Negativo	Negativo	Negativo	Negativo

Contudo, ambos os patógenos, *Shigella flexneri* e *Entamoeba histolytica* são endêmicos de regiões com baixo ou nenhum saneamento básico. Outra característica similar é a maneira de transmissão, que não é dada a partir de um vetor – o que tornaria mais difícil de erradicar – mas uma obrigatoriedade da ingestão do organismo, seja na água ou alimentos, ou seja, uma melhoria no tratamento de efluentes e inserção de sistemas de saneamento básicos eficientes conseguiria reduzir drasticamente os casos dessas doenças. (MOTTA; SILVA, 2002)

Além das cepas de *Shigella flexneri*, foram identificadas 3 cepas de *Shigella boydii*: a strain 192, 183 e strain 80. A *S. boydii* str. 192 foi encontrada em uma criança com 9 anos e meio de idade, sendo a mais velha do grupo. Essa idade não é característica da shigelose, que normalmente atinge crianças de até 5 anos, contudo essa criança estava com sintomas de diarreia e febre. Estes sintomas já perduravam por 5 dias, porém, a falta de muco e sangue, pode descartar a shigelose do diagnóstico, além do mais, a criança portava a presença do helminto *Ascaris lumbricoides* e de Rotavírus.

É bastante provável que essa diarreia tenha sido causada pelo rotavírus, já que o outro *Ascaris l.* não é causador de diarreia (MOTTA; SILVA, 2002). A criança apresentava febre, um dos principais sintomas da infecção causada por rotavírus, que pode se estender até o terceiro dia de doença. Já a diarreia pode durar de 5 a 8 dias, período, este que a criança estava com diarreia. De mesmo modo segue a criança de onde foi isolada a cepa *S. boydii* str. 183, que teve diagnóstico positivo para rotavírus também, e sintomas com ausência de muco e sangue, contudo, apresentando vômitos, um sintoma característico da infecção causada por rotavírus, principalmente no início do período de incubação, entre 24 e 48 horas, além de febre. (OLIVEIRA; LINHARES, 1999)

A *S. boydii* str. 80, teve resultado negativo para rotavírus e protozoários, porém teve um helminto, o *Enterobius vermicularis*, ou oxiúros, causador da oxiurose, que causa uma coceira anal, principalmente à noite. Ele também é transmitido a partir de alimentos e se aloja no intestino humano. É possível que esse caso seja uma shigelose, pela presença de sangue nas fezes, mas pode ser também por conta da oxiurose, que

pode voltar à tona outras vezes, devido a eclosão dos ovos do parasita, que são depositados na região perianal, o que justificaria as 4 internações no ano (CHU et al., 2012).

Essas análises de sintomatologia são importantes pois, ao que tudo indica temos casos de coinfeção, ou seja, a *Shigella* infectando o paciente juntamente com outro patógeno, onde juntos refletem sua virulência nos sintomas, ou um sobressai ao sintoma do outro. No caso das cepas de *S. boydii* é crível que as 3 crianças com essa espécie estavam com sintomas sendo causados por outro patógeno, e não pela *Shigella*, o que leva a crer que a *S. boydii* tem se tornado uma bactéria que consegue permanecer latente dentro do intestino humano, ainda mais com a descoberta do sistema de secreção do tipo VI, que concede uma vantagem competitiva com outras bactérias, podendo assim, permitir que ela faça parte da microbiota do ser-humano, não necessariamente agindo como patógeno.

Outro aspecto importante que pode ser extraído dessa coleta, foi o fato de nenhuma das crianças ter tomado antibiótico. Isso nos remete também a falta do diagnóstico da shigelose, pois, das sete crianças, duas estavam exclusivamente com o patógeno da shigelose, ou seja, era necessário um tratamento específico. No Brasil, o Ministério da Saúde recomenda o uso combinado de sulfametoxazol com trimetoprima, e em casos de resistência, o uso de ciprofloxacina. Contudo, OMS recomenda o uso imediato de ciprofloxacina, um antibiótico da família das quinolonas. Na Europa, a recomendação dada é o uso de azitromicina, e a ceftriaxona ou ácido nalidíxico, como alternativos, porém, nunca o sulfametoxazol combinado com trimetoprima, igual ao Brasil (CARRARI; TAHAN; MORAIS, 2012; RAM et al., 2008).

É sabido que as cepas podem ter resistências diferentes em cada região ou país, entretanto, é necessário que seja feito um estudo de análise de resistência das cepas de *Shigella* circulantes na região amazônica, além de uma atualização por parte do Ministério da Saúde para o antibiótico recomendado.

6. CONCLUSÃO

Com essa caracterização foi possível identificar todas as cepas que não haviam sido identificadas pelo teste bioquímico e por outras metodologias. Foi possível também corroborar os dados da PCR proposta por OHJA em 2013, tendo, portanto, uma metodologia eficiente e barata que pode ser utilizada para a identificação futura de outros isolados de *Shigella*.

Isso foi possível pois o sequenciamento dessas amostras foi de alta qualidade, sendo possível obter até 75% de genes anotados. Das 7 amostras, 4 foram classificadas como *Shigella flexneri* e três como *Shigella boydii*. O pipeline Orthologsorter, que gerou árvores filogenômicas, foi essencial para a confirmação destas espécies.

A montagem dos cromossomos das cepas de *Shigella flexneri* se mostraram bem fidedignas com genomas de outras *Shigella* depositadas em bancos de dados, porém, os cromossomos de *S. boydii* não alcançaram tanto êxito, o que mostra a diferença das espécies isoladas no estado das demais espécies circulantes no mundo, principalmente a cepa 17 que se caracterizou bastante distinta das demais.

Contudo, a montagem dos plasmídeos não pôde ser concluída, variando entre 40% e 68% de identidade com outros plasmídeos depositados em bancos de dados do NCBI. Esse dado também carrega a informação de que, não só os cromossomos dessas espécies são diferentes, mas também o plasmídeo, confirmando o porquê essas bactérias têm se comportado de maneira tão peculiar mediante as conhecidas na literatura e cepas padrões.

Uma das principais diferenças detectadas foi nas vias metabólicas. Entre elas, importante salientar a presença do sistema de secreção do tipo VI, presente nas representantes da espécie *S. boydii*. Esse sistema só havia sido registrado até então em *S. sonnei*, e é responsável por garantir uma vantagem competitiva para a bactéria, podendo ser, portanto, a causa destas *S. boydii* terem sido encontradas nas crianças, pois a *S. boydii* não tem tantos casos registrados, especialmente em regiões fora da Índia, região no qual é endêmica.

Esse resultado pode nos trazer a hipótese que a *Shigella*, principalmente a *S. boydii* tem evoluído para se tornar uma espécie partícipe da microbiota do intestinal

conseguindo agora, não só driblar o sistema imunológico, mas também resistir à outros patógenos e permanecer no habitat, e isso é confirmado ao analisar os dados etio-epidemiológico desta coleta, quando a maioria dos pacientes estavam com co-infecção, ou seja, doença causada não só pela *Shigella*, mas muito provavelmente por outros microrganismos.

7. PERSPECTIVAS FUTURAS

Serão necessárias análises adicionais para estas cepas, como verificar se há genes que configuram a resistência a antibióticos. Já existem trabalhos que comprovam também a aquisição de genes de resistência em cepas de *S. boydii* (VENKATESAN, 2015), e este tema tem sua total relevância no tratamento, tendo em vista que o mesmo é realizado com o uso de antibióticos, logo, estudando o perfil de resistência das cepas circulantes na região será possível definir um tratamento melhor para o Estado.

Deverá ser feito também um estudo para identificar se os sistemas de secreção neste trabalho identificados são funcionais e ativos para estas cepas, como o sistema de secreção do tipo VI. Sendo positivo, será preciso verificar quais os benefícios ou malefícios que isso têm trazido para a disseminação da doença.

Além disso será necessário adotar uma metodologia mais eficiente para a montagem dos plasmídeos, pois a caracterização dos mesmos é de extrema importância, pois, no que tange a patogenicidade, grande parte do aparato de virulência da *Shigella* se encontra neles.

8. REFERÊNCIAS BIBLIOGRÁFICAS

ALBERT, M. J. et al. Diarrhoeagenic *Escherichia coli* are not a significant cause of diarrhoea in hospitalised children in Kuwait. **BMC Microbiology**, v. 9, n. 62, p. (30 March 2009), 2009.

ANDERSON, M. C. et al. *Shigella sonnei* Encodes a Functional T6SS Used for Interbacterial Competition and Niche Occupancy. **Cell Host & Microbe**, v. 21, n. 6, p. 769–776.e3, jun. 2017a.

ANDERSON, M. C. et al. *Shigella sonnei* Encodes a Functional T6SS Used for Interbacterial Competition and Niche Occupancy. **Cell Host & Microbe**, v. 21, n. 6, p. 769–776.e3, 14 jun. 2017b.

ANDERSON, M.; SANSONETTI, P. J.; MARTEYN, B. S. *Shigella* Diversity and Changing Landscape: Insights for the Twenty-First Century. **Frontiers in Cellular and Infection Microbiology**, v. 6, 19 abr. 2016.

ASSEFA, S. et al. ABACAS: algorithm-based automatic contiguation of assembled sequences. **Bioinformatics**, v. 25, n. 15, p. 1968–1969, 1 ago. 2009.

ATLAS DO DESENVOLVIMENTO HUMANO NO BRASIL. **Início | Atlas do Desenvolvimento Humano no Brasil**. Disponível em:

<<http://www.atlasbrasil.org.br/2013/>>. Acesso em: 22 fev. 2017.

BAKER, S.; HANAGE, W. P.; HOLT, K. E. Navigating the future of bacterial molecular epidemiology. **Current Opinion in Microbiology**, v. 13, n. 5, p. 640–645, 2010.

BARRY, E. M. et al. Progress and pitfalls in *Shigella* vaccine research. **Nature Publishing Group**, v. 10, n. 10, p. 245–25512, 2013.

CARRARI, M. H. C.; TAHAN, S.; MORAIS, M. B. Antibiotic therapy in acute diarrhea associated with *Shigella*: what is the best option? **Jornal de Pediatria**, v. 88, n. 4, p. 366–367, 20 ago. 2012.

CENTERS FOR DISEASE CONTROL AND PREVENTION. **Shigella – Shigellosis | Shigella – Shigellosis | CDC**. Disponível em:

<<https://www.cdc.gov/shigella/index.html>>. Acesso em: 9 mar. 2017.

CHATTAWAY, M. A. et al. Whole-Genome Sequencing for National Surveillance of *Shigella flexneri*. **Frontiers in microbiology**, v. 8, p. 1700, 2017.

- CHU, T.-B. et al. Enterobius vermicularis infection is well controlled among preschool children in nurseries of Taipei City, Taiwan. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 45, n. 5, p. 646–648, out. 2012.
- CRISTIANE HAUSCHILDT. Perfil epidemiológico das doenças transmitidas por alimentos no Brasil : uma análise crítica de dados secundários publicados entre 1983 e 2014. 2014.
- CRUZ, C. B. N. DA et al. Virulence Factors Associated with Pediatric Shigellosis in Brazilian Amazon. **BioMed Research International**, v. 2014, p. 1–9, 2014.
- DAL'MOLIN, L. F. C. S. et al. AVALIAÇÃO MICROBIOLÓGICA E FÍSICA QUÍMICA DE PRODUTOS MINIMAMENTE PROCESSADOS COMERCIALIZADOS NA REGIÃO DE CUIABÁ-MT. **CONNECTIONLINE**, v. 0, n. 10, 4 abr. 2014.
- DE OLIVEIRA SOUZA, C.; CRISTINA DE MORAES SILVA, M.; CARLOS BRITO LOUREIRO, E. Fatores de risco e etiologia infecciosa da doença diarreica no Município de Juruti, Estado do Pará, Brasil. **Rev Pan-Amaz Saude**, v. 3, n. 4, p. 49–60, 2012.
- DEFEATDD**. Disponível em: <<https://www.defeatdd.org/>>. Acesso em: 2 mar. 2018.
- DEKKER, J. P.; FRANK, K. M. Salmonella, Shigella, and Yersinia. **Clinics in Laboratory Medicine**, v. 35, n. 2, p. 225–246, 2015.
- DI MARTINO, M. L. et al. Molecular evolution of the nicotinic acid requirement within the Shigella/EIEC pathotype. **International Journal of Medical Microbiology**, v. 303, n. 8, p. 651–661, 2013.
- Diarrhoeal Disease - UNICEF**. Disponível em: <<https://data.unicef.org/topic/child-health/diarrhoeal-disease/#>>. Acesso em: 16 mar. 2018.
- EDWARDS, D. J.; HOLT, K. E. Beginner's guide to comparative bacterial genome analysis using next-generation sequence data. **Microbial Informatics and Experimentation**, v. 3, n. 1, p. 2, 2013.
- EKBLOM, R.; WOLF, J. B. W. A field guide to whole-genome sequencing, assembly and annotation. **Evolutionary applications**, v. 7, n. 9, p. 1026–42, nov. 2014.
- EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL (ECDC). **News - Shigellosis among refugees in the EU**. Disponível em: <http://ecdc.europa.eu/en/press/news/_layouts/forms/News_DispatchForm.aspx?ID=1334&list=8db7286c-fe2d-476c-9133-

18ff4cb1b568&Source=http://ecdc.europa.eu/en/Pages/home.aspx>. Acesso em: 22 fev. 2017.

HAZEN, T. H. et al. Investigating the relatedness of enteroinvasive *Escherichia coli* to Other *E. coli* and *Shigella* isolates by using comparative genomics. **Infection and Immunity**, v. 84, n. 8, p. 2362–2371, 2016.

HOLT, K. E. et al. *Shigella sonnei* genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. **Nature Genetics**, v. 44, n. 9, p. 1056–1059, 5 ago. 2012.

KANEHISA, M. et al. KEGG as a reference resource for gene and protein annotation. **Nucleic acids research**, v. 44, n. D1, p. D457–62, 4 jan. 2016.

KAPER, J. B.; NATARO, J. P.; MOBLEY, H. L. T. Pathogenic *Escherichia coli*. **Nature Reviews Microbiology**, v. 2, n. 2, p. 123–140, fev. 2004.

KOTLOFF, K. L. et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. **The Lancet**, v. 382, n. 9888, p. 209–222, jul. 2013.

KOTLOFF, K. L. The Burden and Etiology of Diarrheal Illness in Developing Countries. **Pediatric Clinics of North America**, v. 64, n. 4, p. 799–814, ago. 2017.

LAGESEN, K. et al. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. **Nucleic Acids Research**, v. 35, n. 9, p. 3100–3108, maio 2007.

LAMPEL, K. A.; FORMAL†, S. B.; MAURELLI, A. T. A Brief History of *Shigella*. **EcoSal Plus**, v. 8, n. 1, 9 jan. 2018.

LEDERER, I. et al. Shigellosis in refugees, Austria, July to November 2015. **Eurosurveillance**, v. 20, n. 48, p. 30081, 3 dez. 2015.

LEVINE, M. M. et al. Clinical trials of *Shigella* vaccines: two steps forward and one step back on a long, hard road. **Nature reviews. Microbiology**, v. 5, n. 7, p. 540–53, 2007.

LIMA, I. F. N.; HAVT, A.; LIMA, A. A. M. Update on molecular epidemiology of *Shigella* infection. **Current Opinion in Gastroenterology**, v. 31, n. 1, p. 30–37, jan. 2015.

LIU, L. et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. **The Lancet**, v. 379, n. 9832, p. 2151–2161, 9 jun. 2012.

LIVIO, S. et al. *Shigella* Isolates From the Global Enteric Multicenter Study Inform

- Vaccine Development. **Clinical Infectious Diseases**, v. 59, n. 7, p. 933–941, 1 out. 2014.
- LOWE, T. M.; CHAN, P. P. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. **Nucleic Acids Research**, v. 44, n. W1, p. W54–W57, 8 jul. 2016.
- MARTEYN, B.; GAZI, A.; SANSONETTI, P. Shigella: a model of virulence regulation in vivo. **Gut microbes**, v. 3, n. 2, p. 104–20, 2012.
- MATTOCK, E.; BLOCKER, A. J. How Do the Virulence Factors of Shigella Work Together to Cause Disease? **Frontiers in Cellular and Infection Microbiology**, v. 7, p. 64, 24 mar. 2017a.
- MATTOCK, E.; BLOCKER, A. J. How Do the Virulence Factors of Shigella Work Together to Cause Disease? **Frontiers in Cellular and Infection Microbiology**, v. 7, p. 64, 24 mar. 2017b.
- MORIYA, Y. et al. KAAS: an automatic genome annotation and pathway reconstruction server. **Nucleic Acids Research**, v. 35, n. Web Server, p. W182–W185, 8 maio 2007.
- MOTTA, M. E. F. A.; SILVA, G. A. P. DA. Diarréia por parasitas. **Revista Brasileira de Saúde Materno Infantil**, v. 2, n. 2, p. 117–127, ago. 2002.
- NIE, H. et al. Complete genome sequence of Shigella flexneri 5b and comparison with Shigella flexneri 2a. **BMC genomics**, v. 7, p. 173, 2006.
- NIU, C. et al. Role of the virulence plasmid in acid resistance of Shigella flexneri. **Scientific reports**, v. 7, p. 46465, 25 abr. 2017.
- NJAMKEPO, E. et al. Global phylogeography and evolutionary history of Shigella dysenteriae type 1. **Nature Microbiology**, n. March, p. 16027, 2016.
- NUNES, M. DO R. C. M. et al. Diarrhea associated with <i>Shigella</i> in children and susceptibility of the bacterium to antimicrobials: a study in Teresina, Piauí, Northeast of Brazil. **Jornal de Pediatria**, 16 nov. 2011.
- OLIVEIRA, C. S. DE; LINHARES, A. C. Rotavírus : aspectos clínicos e prevenção. **Jornal de pediatria**, v. 75, n. Supl. 1, p. 91–102, 1999.
- OLSON, N. D. et al. **Best practices for evaluating single nucleotide variant calling methods for microbial genomics** *Frontiers in Genetics* Frontiers Media SA, , 2015. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/26217378>>. Acesso em: 31 jan.

2017

PAAUW, A. et al. Rapid and reliable discrimination between *Shigella* species and *Escherichia coli* using MALDI-TOF mass spectrometry. **International Journal of Medical Microbiology**, v. 305, n. 4–5, p. 446–452, 2015.

PARSOT, C. ***Shigella* spp. and enteroinvasive *Escherichia coli* pathogenicity factors** **FEMS Microbiology Letters**, 2005.

PAVLOVIC, M. et al. Development of a duplex real-time PCR for differentiation between *E. coli* and *Shigella* spp. **Journal of Applied Microbiology**, v. 110, n. 5, p. 1245–1251, maio 2011.

PHALIPON, A.; SANSONETTI, P. J. *Shigella*'s ways of manipulating the host intestinal innate and adaptive immune system: a tool box for survival? **Immunology and Cell Biology**, v. 85, n. 2, p. 119–129, 1 fev. 2007.

PRUESSE, E.; PEPLIES, J.; GLÖCKNER, F. O. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. **Bioinformatics**, v. 28, n. 14, p. 1823–1829, 15 jul. 2012.

PUHAR, A.; SANSONETTI, P. J. Type III secretion system. **Current Biology Vol**, v. 24, n. 17, 2014.

RAM, P. K. et al. Part II. Analysis of data gaps pertaining to *Shigella* infections in low and medium human development index countries, 1984-2005. **Epidemiology and Infection**, v. 136, n. 5, p. 577–603, maio 2008.

RICHARDSON, E. J.; WATSON, M. The automatic annotation of bacterial genomes. **Briefings in bioinformatics**, v. 14, n. 1, p. 1–12, jan. 2013.

ROCHA, F.; BARCELOS, T.; FARIA, N. Avaliação Microbiológica e Microscópica do Caldo de Cana Comercializado em Ji-Paraná. **Ensaio e Ciência: Ciências Biológicas, Agrárias e da Saúde**, v. 20, n. 2, p. 111–115, 2016.

SANA, T. G.; LUGO, K. A.; MONACK, D. M. T6SS: The bacterial "fight club" in the host gut. **PLOS Pathogens**, v. 13, n. 6, p. e1006325, 8 jun. 2017.

SANTOS, F. L. N.; SOARES, N. M. Mecanismos fisiopatogênicos e diagnóstico laboratorial da infecção causada pela *Entamoeba histolytica*. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, v. 44, n. 4, p. 249–261, ago. 2008.

SCHMIEDER, R.; EDWARDS, R. Quality control and preprocessing of metagenomic

- datasets. **Bioinformatics**, v. 27, n. 6, p. 863–864, 15 mar. 2011.
- SEEMANN, T. Prokka: rapid prokaryotic genome annotation. **Bioinformatics**, v. 30, n. 14, p. 2068–2069, 15 jul. 2014.
- SILVA, T. et al. Characterization of *Shigella* spp. by antimicrobial resistance and PCR detection of ipa genes in an infantile population from Porto Velho (Western Amazon region), Brazil. **Mem Inst Oswaldo Cruz Rio de Janeiro**, v. 103, n. 7, p. 731–733, 2008.
- SOUSA, M. Â. B. et al. *Shigella* in Brazilian children with acute diarrhoea: prevalence, antimicrobial resistance and virulence genes. **Memorias do Instituto Oswaldo Cruz**, v. 108, n. 1, p. 30–5, fev. 2013.
- STOTHARD, P.; WISHART, D. S. Circular genome visualization and exploration using CGView. **Bioinformatics**, v. 21, n. 4, p. 537–539, 15 fev. 2005.
- THE, H. C. et al. The genomic signatures of *Shigella* evolution, adaptation and geographical spread. **Nature Reviews Microbiology**, v. 14, n. 4, p. 235–250, 29 abr. 2016a.
- THE, H. C. et al. The genomic signatures of *Shigella* evolution, adaptation and geographical spread. **Nature Reviews Microbiology**, v. 14, n. 4, p. 235–250, 29 abr. 2016b.
- THOMPSON, C. N.; DUY, P. T.; BAKER, S. The Rising Dominance of *Shigella sonnei*: An Intercontinental Shift in the Etiology of Bacillary Dysentery. **PLOS Neglected Tropical Diseases**, v. 9, n. 6, p. e0003708, 11 jun. 2015.
- UD-DIN, A.; WAHID, S. Relationship among *Shigella* spp. and enteroinvasive *Escherichia coli* (EIEC) and their differentiation. **Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]**, v. 45, n. 4, p. 1131–8, 2014.
- VAN DEN BELD, M. J. C.; REUBSAET, F. A. G. **Differentiation between *Shigella*, enteroinvasive *Escherichia coli* (EIEC) and noninvasive *Escherichia coli*** **European Journal of Clinical Microbiology and Infectious Diseases**, 2012.
- VENKATESAN, M. M. A novel protein-based subunit *Shigella* vaccine candidate. **Immunology and Cell Biology**, v. 93, p. 603–604, 2015.
- VENTER, J. C. The Sequence of the Human Genome. **Science**, v. 291, n. 5507, p. 1304–1351, 16 fev. 2001.

VIEIRA, N. et al. High prevalence of enteroinvasive Escherichia coli isolated in a remote region of northern coastal Ecuador. **The American journal of tropical medicine and hygiene**, v. 76, n. 3, p. 528–33, mar. 2007.

VONGSAWAN, A. A. et al. The genome of Shigella dysenteriae strain Sd1617 comparison to representative strains in evaluating pathogenesis. **FEMS microbiology letters**, v. 362, n. 5, mar. 2015.

WANG, Q. et al. Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. **Applied and Environmental Microbiology**, v. 73, n. 16, p. 5261–5267, 15 ago. 2007.

WHITE, A. C.; KANG, G. Multiplex molecular diagnostic tests and the management of diarrhea. **Current Opinion in Infectious Diseases**, v. 30, n. 5, p. 471–472, out. 2017.

WHO | Diarrhoeal disease. **WHO**, 2017.

YANDELL, M.; ENCE, D. A beginner's guide to eukaryotic genome annotation. **Nature Reviews Genetics**, v. 13, n. 5, p. 329–342, 18 abr. 2012.

YANG, F. et al. Genome dynamics and diversity of Shigella species, the etiologic agents of bacillary dysentery. **Nucleic acids research**, v. 33, n. 19, p. 6445–58, 2005.

ZAGHLOUL, L. et al. The distribution of insertion sequences in the genome of Shigella flexneri strain 2457T. **FEMS Microbiology Letters**, v. 277, n. 2, p. 197–204, 2007.

ZERBINO, D. R. Using the Velvet de novo assembler for short-read sequencing technologies. **Current protocols in bioinformatics**, v. Chapter 11, p. Unit 11.5, set. 2010.

ZERBINO, D. R.; BIRNEY, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. **Genome research**, v. 18, n. 5, p. 821–9, maio 2008.

9. ANEXOS

Anexo I - EXTRAÇÃO DE DNA FEM FENOL/CLOROFÓRMIO (Gram-negativo)

1. Crescer bactéria *overnight*;
2. Centrifugar 300uL por 10 minutos/12mil RPM;
3. Descartar sobrenadante;
4. 95uL de tampão osmótico (100mM Tris e 50mM EDTA);
- 40uL de Lisozima (20mg/mL); (somente p/ gram-positivo)
5. Incubar por 30min/37°C em banho seco (homogeneizar periodicamente);
6. Adicionar 1% de SDS 10% (11uL) e incubar* 30 min no banho seco a 60°C
(homogeneizar periodicamente e observar a turbidez diminuir e ficar límpido)
7. Adicionar v/v de acetato de potássio 5M (100uL) e verificar se forma uma gosma;
8. Colocar no gelo por 30 minutos;
9. Centrifugar 10mil RPM por 15 minutos a temp. ambiente;
10. Recuperar o sobrenadante e colocar em novo tubo;
11. Adicionar v/v (+- 200uL) de álcool isopropílico gelado e deixar agir por 30 minutos no congelador (essa fase pode ser *overnight*);
12. Centrifugar por 15 minutos/14mil RPM;
13. Descartar o sobrenadante e adicionar 500uL de EtOH 70% gelado e misturar por inversão (vortex);
14. Centrifugar por 15 minutos a 10mil RPM a temp. ambiente;
15. Descartar sobrenadante e deixar o evaporar o álcool (SpeedVac);
16. Ressuspender em 30uL de TE (10mm) e 1uL de RNA (10mg/mL);
17. Incubar no banho seco por 30 min a 37°C;
18. Correr gel de agarose 1%;

Obs.: Lisozima se usa para Gram+, porém pode usar se quiser;
Se colocar todos os reagentes gelados, pode diminuir o tempo no gelo (obs.: colocar o acetato de potássio gelado e diminui no gelo);

- O tempo de colocar o SDS e incubar em banho seco NÃO pode ultrapassar 30 segundos!