



UNIVERSIDADE FEDERAL DO CEARÁ
FACULDADE DE MEDICINA
DEPARTAMENTO DE MORFOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS MORFOFUNCIONAIS

DIVISON DE MELO PACÍFICO

**ANÁLISE EXOPROTEÔMICA COMPARATIVA DO SOBRENADANTE DE UMA
CEPA DE *Clostridium difficile* ISOLADA EM UM HOSPITAL CEARENSE**

FORTALEZA

2018

DIVISON DE MELO PACÍFICO

**ANÁLISE EXOPROTEÔMICA COMPARATIVA DO SOBRENADANTE DE UMA
CEPA DE *Clostridium difficile* ISOLADA EM UM HOSPITAL CEARENSE**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Ciências Morfofuncionais da Faculdade de Medicina da Universidade Federal do Ceará, como requisito parcial para a obtenção do título de Mestre em Morfologia.

Área de concentração: Morfologia.

Orientadora: Profa. Dra. Gerly Anne de Castro Brito

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

P125a Pacífico, Dvison de Melo.

Análise exoproteômica comparativa do sobrenadante de uma cepa de *Clostridium difficile* isolada em um hospital cearense / Dvison de Melo Pacífico. – 2018.
110 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Faculdade de Medicina, Programa de Pós-Graduação em Ciências Morfofuncionais, Fortaleza, 2018.

Orientação: Profa. Dra. Gerly Anne de Castro Brito.

1. *Clostridium difficile*. 2. Exoproteômica. 3. Toxinas. 4. Proteínas de superfície. 5. Flagelina. I. Título.

CDD 611

DIVISON DE MELO PACÍFICO

**ANÁLISE EXOPROTEÔMICA COMPARATIVA DO SOBRENADANTE DE UMA
CEPA DE *Clostridium difficile* ISOLADA EM UM HOSPITAL CEARENSE**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Ciências Morfofuncionais da Faculdade de Medicina da Universidade Federal do Ceará, como requisito parcial para a obtenção do título de Mestre em Morfologia.

Área de concentração: Morfologia.

Orientadora: Profa. Dra. Gerly Anne de Castro Brito

Aprovada em: ____/____/____.

BANCA EXAMINADORA

Profa. Dra. Gerly Anne de Castro Brito (Orientadora)
Universidade Federal do Ceará (UFC)

Profa. Dra. Eliane de Oliveira Ferreira
Universidade Federal do Rio de Janeiro (UFRJ)

Prof. Dr. Renato de Azevedo Moreira
Universidade de Fortaleza (Unifor)

“Para conquistar o sucesso, você precisa aceitar todos os desafios que vierem na sua frente. Você não pode apenas aceitar os que você preferir.”

Mike Gafka

AGRADECIMENTOS

Agradeço a Deus primeiramente, pois muitos momentos de dificuldades que passei ao longo dessa caminhada foi Ele quem me manteve sempre de pé dando forças e mostrando uma luz no final do túnel para eu poder seguir em frente e hoje estar finalizando o meu mestrado. Agradeço por nunca ter me abandonado nem se quer por um segundo nessa minha trajetória.

A minha orientadora, Profa. Dra. Gerly Anne de Castro Brito, por ter aceitado me orientar mesmo não sabendo se eu seria um bom pós-graduando. Agradeço pela paciência, amizade, pelos conselhos científicos e por todos os ensinamentos que recebi, além de ter confiado em minha pessoa. Profa. Gerly, a senhora é exemplo de pesquisadora e professora que todo e qualquer pós-graduando sempre sonha em ter. Sei que tenho muito que aprender ainda a respeito desse ramo da pesquisa que foi proposto pela senhora, mas uma coisa é certa, nada é impossível para aquele que crê. Obrigado mais uma vez pela sua confiança.

Ao Prof. Dr. Jand-Venes Rolim Medeiros, por ter me orientado no período da graduação e a quem devo muito, pois foi o senhor que me mostrou o caminho da pesquisa e que hoje ainda mantemos parcerias.

Aos meus pais, Socorro e Veridiano, por terem me proporcionado toda a educação digna que tenho hoje e também por todos os conselhos que recebi de vocês para que eu pudesse tornar o homem que sou hoje. Agradeço ainda aos meus irmãos Daylson e Deylson pelo apoio.

A minha noiva, Leiliane, por todos os dias que sempre está ao meu lado, me incentivando e segurando as minhas mãos para que eu nunca viesse a desistir dessa jornada que foi bem árdua e também pelos conselhos que recebo de você todos os dias. Obrigado por sempre ter apoiado as minhas decisões e também por sempre ter entendido os momentos que fiquei um pouco ausente de você.

A Profa. Dra. Renata Ferreira de Carvalho Leitão, pela amizade e exemplo de professora e pesquisadora. Agradeço também aos demais professores do departamento de morfologia (NEMPI) e da Pós-graduação que de alguma forma contribuíram com o meu aprendizado.

As amigas, Josy e Deizeane, pela ajuda dada no laboratório NEMPI em relação aos experimentos e também por compartilharem seus aprendizados sobre a pesquisa. Agradeço ainda aos amigos de laboratório e da sala de cultura Gildênio, Cíntia, Ariel, Mariana, Luciana, Luane, Hildênia, Luana e Fátima, Theídes, Lauro, e a Cecília por ter contribuído com esse trabalho.

O aluno de iniciação científica (IC), Paulo, por ter me ajudado em muitos experimentos nos fins de semana. Agradeço aos ICs do laboratório NEMPI Guilherme e Letícia.

Aos técnicos do NEMPI, Junior e Flávia.

A secretária da Pós-graduação, Laísa, por sempre nos manter informado dos assuntos da Pós.

A Profa. Dra. Eliane de Oliveira Ferreira por ter realizado uma parte desse trabalho no seu laboratório no Rio de Janeiro. Agradeço também aos Drs. Hércules Moura e o Dr. John Barr, *Disease Control and Prevention (CDC)* nos Estados Unidos, por ter realizado a outra parte desse trabalho em seu laboratório.

Aos membros da banca, Profa. Dra. Débora Castelo Branco de Sousa Colares Maia, Dra. Ana Cristina de Oliveira Monteiro Moreira, Dra. Eliane de Oliveira Ferreira e ao Prof. Dr. Renato Azevedo Moreira, por atenderem ao nosso convite para participar da minha qualificação/defesa de mestrado e contribuírem para melhoria dessa dissertação.

A Universidade Federal do Ceará (UFC) pela oportunidade de desenvolver a minha pesquisa nessa instituição.

A Fundação cearense de apoio ao desenvolvimento científico e tecnológico (FUNCAP) pelo apoio financeiro.

RESUMO

C. difficile, bactéria Gram-positiva, anaeróbia formadora de esporos, é considerada a principal causa de diarreia associada ao uso de antibióticos em pacientes hospitalizados no mundo todo, causando a infecção por *C. difficile* (CDI). Os principais fatores de virulência de *C. difficile* são as toxinas A (TcdA) e B (TcdB), porém, outros, como, as proteínas de superfície celular (SLPs) e os flagelos, também já foram descritos e relacionados às respostas inflamatórias observadas na CDI. O objetivo desse estudo foi identificar esses fatores de virulência e outras proteínas envolvidas com a patogênese da CDI por meio da exoproteômica comparativa entre o sobrenadante de uma cepa virulenta, a ICC-45, isolada em um hospital cearense e compará-la com a cepa NAP1/027 (LIBA5756) isolada na Costa Rica. Após o crescimento das cepas de *C. difficile* por um período de incubação de 24 h a 37 °C em jarra de anaerobiose, as proteínas do sobrenadante foram obtidas e analisadas por meio da técnica em solução (*in solution*) que posteriormente foram analisadas pela espectrometria de massas do tipo Nano-LC ESI-MS/MS acoplado a LTQ-Orbitrap. A partir da análise *in solution*, foram identificadas um total de 197 proteínas, sendo que 5 delas pertencem exclusivamente a ICC-45. Por outro lado, as cepas ICC-45 e NAP1/027 (LIBA5756) compartilharam 192 proteínas. Das 192 proteínas compartilhadas, foram selecionadas 26 e categorizadas em seis grupos (patogenicidade, resistência a antimicrobianos, choque térmico e estresse oxidativo, óxido nítrico, metabolismo e outras atividades). Várias proteínas referentes aos fatores de virulência e que contribuem positivamente para a patogenicidade da CDI, foram identificadas em maior quantidade no sobrenadante da cepa ICC-45 quando comparada a cepa NAP1/027 (LIBA5756). Entre essas proteínas, estão as TcdA e TcdB, proteína de camada S (S-layer), proteína de superfície celular (Cwp19), proteínas de parede celular (Cwps), flagelina (FliC) e cisteína protease (Cwp84). Além dessas, outras proteínas referentes as demais categorias (resistência a antimicrobianos, choque térmico e estresse oxidativo, óxido nítrico, metabolismo e outras atividades) também foram encontradas em quantidade maior no sobrenadante da cepa ICC-45. No que tange especificamente as funções moleculares e a localização subcelular, as proteínas exclusivas da cepa ICC-45 estão envolvidas com as funções catalíticas e de ligação. Já com relação às proteínas compartilhadas entre as cepas mencionadas, a maior parte delas estão ligadas com as funções catalíticas, seguida das de ligação e de toxina. Mais da metade das proteínas exclusivas da cepa ICC-45 e das compartilhadas entre elas são de origem citoplasmática. As proteínas das cepas ICC-45 e NAP1/027 (LIBA5756) apresentaram interações físicas e funcionais semelhantes. As proteínas exclusivas da cepa ICC-45 interagem indiretamente com as proteínas de patogenicidade (TcdA, TcdB, FliC, Cwp84 e SlpA) compartilhadas entre as cepas. Esses dados sugerem uma semelhança entre a cepa ICC-45 e a cepa NAP1/027 (LIBA5756), considerada hipervirulenta. Além disso, a detecção de maiores quantidades de determinadas proteínas na cepa ICC-45 pode ter um impacto na patogênese e parâmetros clínicos da doença induzida por essa cepa isolada no Nordeste do Brasil.

Palavras-chave: *Clostridium difficile*. Exoproteômica. Toxinas. Proteínas de superfície. Flagelina.

ABSTRACT

C. difficile, a Gram-positive, spore-forming anaerobic bacterium, is considered the leading cause of antibiotic-associated diarrhea in hospitalized patients worldwide, causing *C. difficile* infection (CDI). The main virulence factors of *C. difficile* are toxins A (TcdA) and B (TcdB), but others, such as surface layer proteins (SLPs) and flagellum, have also been described and related to the observed inflammatory responses in the CDI. The objective of this study was to identify these virulence factors and other proteins involved in the pathogenesis of CDI by comparative exoproteomics between the supernatant of a virulent strain, ICC-45, isolated in a hospital in the state of Ceará and to compare it with the strain NAP1/027 (LIBA5756) isolated in Costa Rica. After the growth of the *C. difficile* strains by an incubation period of 24h at 37°C in the anaerobic jar, the secreted proteins were obtained and analyzed by the in solution technique that were later processed by the type mass spectrometry Nano-LC ESI-MS / MS coupled to LTQ-Orbitrap. From the analysis in solution, a total of 197 proteins were identified, of which 5 belong exclusively the ICC-45 strain. On the other hand, strains ICC-45 and NAP1/027 (LIBA5756) shared 192 proteins. Of the 192 shared proteins, 26 were selected and categorized into six groups (pathogenicity, antimicrobial resistance, heat shock and oxidative stress, nitric oxide, metabolism and other activities). Several proteins related to virulence factors that contribute positively to the pathogenicity of CDI were identified in greater quantity in the ICC-45 strain supernatant when compared to the strain NAP1/027 (LIBA5756). Among these proteins are TcdA and TcdB, S-layer protein, cell surface protein (Cwp19), cell wall proteins (Cwps), flagellin (FliC) and cysteine protease (Cwp84). In addition, other proteins related to the other categories (antimicrobial resistance, heat shock and oxidative stress, nitric oxide, metabolism and other activities) were also found in a larger quantity in the supernatant of the ICC-45 strain. The exclusive proteins of the ICC-45 strain are involved with the catalytic and binding functions. Among the proteins shared, most of them are linked with the catalytic functions, followed by the binding and toxin functions. Both, proteins unique to the ICC-45 strain and those shared between them, more than half are of cytoplasmic origin. Proteins from strains ICC-45 and NAP1 / 027 (LIBA5756) showed similar physical and functional interactions. The ICC-45 exclusive proteins indirectly interacts with the pathogenicity proteins (TcdA, TcdB, FliC, Cwp84 and SlpA) shared among the strains. In addition, the detection of higher amounts of certain proteins in the ICC-45 strain may have an impact on the pathogenesis and clinical parameters of the disease induced by this strain isolated in Northeast of Brazil.

Keywords: *Clostridium difficile*. Exoproteomics. Toxins. Surface proteins. Flagellin.

LISTA DE QUADROS E TABELAS

Quadro 1 – Dados referentes de cada pesquisa realizada com as cepas de <i>C. difficile</i> utilizando as ferramentas da proteômica.....	50
Quadro 2 – Cepas de <i>C. difficile</i> selecionadas para o estudo da exoproteoma.....	54
Tabela 1 – Proteínas secretadas exclusivamente pela cepa ICC-45 de <i>C. difficile</i>	62
Tabela 2 – Quantidade relativa de proteínas secretadas pelas cepas ICC-45 e NAP1/027 (LIBA5756) de <i>C. difficile</i>	63

LISTA DE FIGURAS

Figura 1 – Microbiologia de <i>C. difficile</i> e suas características morfológicas.....	19
Figura 2 – Complexo proteico da camada S de <i>C. difficile</i>	21
Figura 3 – Locus de patogenicidade (PaLoc) das toxinas produzidas por <i>C. difficile</i>	24
Figura 4 – Organização dos quatros domínios funcionais das toxinas de <i>C. difficile</i>	26
Figura 5 – Patogênese da infecção causada por <i>C. difficile</i>	27
Figura 6 – Endoscopia da colite pseudomembranosa causada por <i>C. difficile</i>	28
Figura 7– Interações patógeno-microbiano-hospedeiro por meio de ferramentas proteômica em diferentes níveis (moleculares, unicelulares, de organismos e populacionais).	35
Figura 8 – Abordagens proteômica “ <i>bottom-up</i> ” e “ <i>top-down</i> ” para identificação de proteínas.	37
Figura 9 – Representação esquemática das seis etapas experimentais utilizando abordagem proteômica “ <i>bottom-up</i> ”.	39
Figura 10 – Esquema do princípio das cinco etapas de separação proteica por eletroforese 2D.	40
Figura 11 – Componentes essenciais de um espectrômetro de massa.....	42
Figura 12 – Representação esquemática do processo de formação de íons numa fonte de ESI.	44
Figura 13 – Representação esquemática do espectrômetro de massa híbrido LTQ-Orbitrap. .	45
Figura 14 – Representação esquemática do analisador de massa Orbitrap.	46
Figura 15 – Fluxograma das etapas utilizada na exoproteômica para identificação e análise dos extratos proteicos obtidos das culturas de cepas de <i>C. difficile</i>	58
Figura 16 – Diagrama de Venn das proteínas identificadas e validadas pelo Scaffold.....	59
Figura 17 – Gráfico das proteínas exclusivas da cepa ICC-45 envolvidas nos processos biológicos.	65
Figura 18 – Gráfico das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756) envolvidas nos processos biológicos.	66
Figura 19 – Função molecular e distribuição das proteínas exclusivas da ICC-45.....	67
Figura 20 – Função molecular e distribuição das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756).	68
Figura 21 – Componente celular das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756).	69
Figura 22 – Localização celular das proteínas exclusivas da cepa ICC-45.....	70

Figura 23 – Localização celular das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756).	70
Figura 24 – Diferenças das interações entre as proteínas da cepa ICC-45 (A-1) e NAP1/027 (LIBA5756) (B-1).	72
Figura 25 – Diferenças das interações físicas e funcionais entre as proteínas selecionadas da cepa ICC-45 (incluindo as exclusivas dessa cepa) com as compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756).	73

LISTA DE ABREVIATURAS E SIGLAS

1D-E – Eletroforese unidimensional

2D- DIGE – Eletroforese bidimensional diferencial

2D-E – Eletroforese bidimensional

BI – Tipo por análise de fragmentos gerados por endonucleases de restrição

BHI – do inglês “*Brain Heart Infusion*” (Infusão de cérebro e coração)

CCFA – do inglês “*Cycloserine, Cefoxitin, Fructose Agar*” (Agar Cicloserina, Cefoxitina e Frutose)

CDI – do inglês “*Clostridium difficile infection*” (Infecção por *C. difficile*)

cdtA – Gene que codifica a subunidade enzimática A da toxina binária produzida por *C. difficile*

cdtB – Gene que codifica a subunidade ligante B da toxina binária produzida por *C. difficile*

CdtLoc – Região gênica que codifica a toxina binária produzida por *C. difficile*

cdtR – Gene que codifica o regulador positivo da expressão das subunidades A e B da toxina binária produzida por *C. difficile*

CPD – Domínio da cisteína protease

CTD – Toxina binária produzida por *C. difficile*

CWP – do inglês “*Clostridial Wall Protein*” (Proteínas de parede clostridial)

Cwp84 – cisteína protease ancorada na parede celular de *C. difficile*

ESI – Ionização por eletrospray

FliC – Flagelina

FliD – Proteína Cap Flagelar

fliC – Gene que codifica a flagelina

fliD – Gene que codifica a proteína CAP flagelar

FTICR – Ressonância de ciclotrônica de íons com transformada de Fourier

GTD – Domínio N-terminal da glicosiltransferase

GTPases – Família de enzimas hidrolases que se ligam e hidrolizam o GTP

Hsp – do inglês “*heat shock protein*” (proteína de choque térmico)

Hsp60 – do inglês “*heat shock proteins 60 family*” (família de proteínas de choque térmico)

ICC – Instituto do Câncer do Ceará

IEF – Focalização isoeletrica

IL-1 β – Interleucina-1 β

iNOS – Óxido nítrico sintase induzida

IPG – Gradiente de pH imobilizado

LC – Cromatografia líquida

LIT – Linear Íon Trap

LTQ – Linear Trap Quadropolo

m/z – razão massa/carga

MALDI – Ionização/desorção a laser assistida por matriz

MLST – do inglês “*Multilocus Sequence Typing*” (Tipagem de Sequência Multilocus)

MM – Massa molecular

kDa – KiloDalton

MS – Espectrometria de massas

MS/MS – Espectrometria de massas em tandem

MASCOT – programa de buscas por sequências homólogas de proteínas

NAP – do inglês “*North American Pulsotype*” (Pulsotipo Norte Americano)

NF- κ B – Fator de transcrição nuclear-kb

PaLoc – do inglês “*Pathogenicity locus*” (Lócus de Patogenicidade)

PCR – do inglês “*Polymerase Chain Reaction*” (Reação em cadeia da polimerase)

PFGE – do inglês “*Pulsed-field Gel Electrophoresis*” (Eletroforese de Campo Pulsado)

pI – Ponto isolétrico

PMNL – Leucócitos polimorfonucleares

PTMs – Modificações pós-traducionais

QIT – Quadropolo Íon Trap

RBD – Domínio de ligação ao receptor

SDS-PAGE – Eletroforese em gel de poliacrilamida com dodecil sulfato de sódio

SlpA – do inglês “*Surface layer protein A*” (Proteína de superfície A)

tcdA – Gene que codifica a enterotoxina A produzida por *C. difficile*

TcdA – Proteína correspondente à enterotoxina A produzida por *C. difficile*

tcdB – Gene que codifica a enterotoxina B produzida por *C. difficile*

TcdB – Proteína correspondente à enterotoxina B produzida por *C. difficile*

tcdC – Gene que codifica o regulador negativo da expressão das enterotoxinas A e B produzidas por *C. difficile*

tcdE – Gene que codifica uma proteína responsável pela liberação extracelular das enterotoxinas A e B produzidas por *C. difficile*

tcdR – Gene que codifica o regulador positivo da expressão das enterotoxinas A e B produzidas por *C. difficile*

TGF- β – Fator de crescimento transformador- β

TNF- α – Fator de necrose tumoral- α

TOF – do inglês “*Time-of-flight*” (Tempo de voo)

SUMÁRIO

1 INTRODUÇÃO	17
1.1 <i>Clostridium difficile</i> : abordagem histórica	17
1.2 A biologia e características de <i>Clostridium difficile</i>	18
1.3 Fatores de virulência e as toxinas de <i>Clostridium difficile</i>	20
1.4 Domínios funcionais de TcdA e TcdB	25
1.5 Patogênese da infecção causada por <i>C. difficile</i>	26
1.6 Estudos <i>in vivo</i> e <i>in vitro</i> dos efeitos das TcdA e TcdB de <i>C. difficile</i>	28
1.7 Infecção por <i>C. difficile</i> : epidemiologia global e o impacto econômico na saúde pública	31
1.8 Análise proteômica	34
1.8.1 Abordagem proteômica “ <i>bottom-up</i> ” e “ <i>top-down</i> ”: estratégias para identificação de proteínas	36
1.8.2 Delineamento de um estudo proteômico “ <i>bottom-up</i> ”: análise de peptídeos por MS.....	38
1.8.3 Eletroforese bidimensional	40
1.8.4 Espectrometria de massas	41
1.8.5 Técnica de ionização por electrospray (ESI).....	42
1.8.6 Tipos de analisadores de massas	44
1.9 Análise proteômica aplicada ao <i>C. difficile</i>	46
2 JUSTIFICATIVA E RELEVÂNCIA	51
3. OBJETIVOS	53
3.1 Objetivo geral	53
3.2 Objetivos específicos.....	53
4 MATERIAIS E MÉTODOS	54
4.1 Cepas bacterianas	54
4.2 Cultivo das cepas de <i>C. difficile</i>	54
4.3 Obtenção dos sobrenadantes.....	55
4.4 Precipitação das proteínas obtidas dos sobrenadantes de <i>C. difficile</i>	55
4.5 Tripsinização <i>in solution</i> das proteínas	56
4.6 Identificação das proteínas em MS nLC MS/MS acoplado a LTQ Orbitrap	56
4.7 Análises de bioinformática	57
5 RESULTADOS	59
5.1. Análise das proteínas obtidas dos sobrenadantes das culturas das cepas de <i>C. difficile</i> ...	59
5.2.1 Funções biológicas: programa Blast2Go.....	65
5.2.2 Localização subcelular das proteínas: programa Psortb.....	69

5.2.3 Rede de associação das proteínas funcionais dos sobrenadantes das cepas de <i>C. difficile</i> analisada por STRING	71
6 DISCUSSÃO	74
7 CONCLUSÃO	86
8 REFERÊNCIAS	87
ANEXO I	103

1 INTRODUÇÃO

1.1 *Clostridium difficile*: abordagem histórica

O trato gastrointestinal apresenta uma microbiota extremamente complexa, constituída de pelo menos várias centenas de espécies bacterianas diferentes, que desempenha um importante papel na saúde humana, produzindo nutrientes e atuando como uma barreira natural seletiva contra a colonização intestinal por potenciais micro-organismos patogênicos (GUARNER; MALAGELADA, 2003; LIU *et al.*, 2003; PENDERS *et al.*, 2005). No entanto, se por algum motivo ocorre uma alteração ou desequilíbrio desse microbioma intestinal, condição conhecida como “disbiose”, o indivíduo passa a ficar mais susceptível a infecções por patógenos entéricos (STECHEER; MAIER; HARDT, 2013). Dentre os patógenos entéricos, o *Clostridium difficile* (*C. difficile*), que recentemente foi renomeado como *Clostridioides difficile* (LAWSON *et al.*, 2016), é um bacilo Gram-positivo e anaeróbio esporulado, que vem se destacando como um importante enteropatógeno causador de doenças intestinais associadas principalmente ao uso de antibióticos, podendo levar desde uma leve diarreia a colite pseudomembranosa fulminante e que muitas das vezes o paciente pode vir a óbito (BARTLETT; GERDING, 2008; SANDERS *et al.*, 2013; SILVA *et al.*, 2015).

O primeiro relato de colite pseudomembranosa (condição clínica caracterizada pela formação de pseudomembranas que são compostas por células necróticas, fibrina, muco e leucócitos) foi registrado numa paciente cirúrgica na Universidade Johns Hopkins em 1892 (MARTIN; MONAGHAN; WILCOX, 2016). A paciente era uma mulher de 22 anos que foi submetida a um procedimento cirúrgico pelo Dr. William Osler para ressecção de um tumor no piloro gástrico. No entanto, no início do pós-operatório, desenvolveu diarreia grave e morreu no 15º dia pós-operatório. A autópsia revelou uma pseudomembranosa "membrana diftérica" no intestino delgado (NAPOLITANO; EDMISTON, 2017). Porém, nesta mesma época, o agente causador dessa doença não tinha sido até então elucidado (LYERLY *et al.*, 1988).

Apenas em 1935, Hall e O’Toole, pesquisando a microbiota intestinal de neonatos saudáveis, descreveram um micro-organismo até então desconhecido, que inicialmente recebeu a denominação de *Bacillus difficilis* e, posteriormente, de *C. difficile*, refletindo justamente a dificuldade para isolar e conservar este micro-organismo em cultura pura. Esses mesmos autores revelaram ainda que o *C. difficile* é toxigênico, já que observaram lesões, parada respiratória e morte em animais de laboratório quando submetidos à injeção de

sobrenadante de cultura ou filtrado do meio que continha esse micro-organismo (HALL; O'TOOLE, 1935; LYERLY *et al.*, 1988).

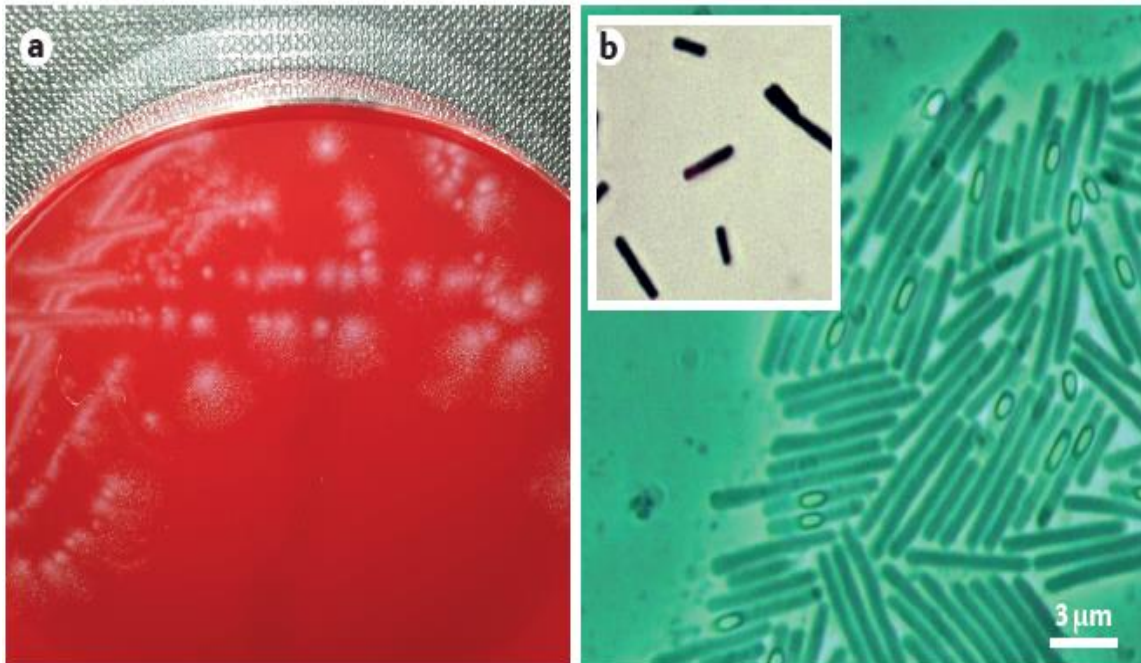
Nos 40 anos que se seguiram a descoberta do *C. difficile*, houve poucos relatos do isolamento do mesmo, haja vista que a presença dessa bactéria era uma condição rara de se observar durante a era pré-antibiótica. Entretanto, por volta de 1975 e início da década de 80, o *C. difficile* começou a ganhar proporções epidêmicas, sendo implicado como o agente responsável pelas complicações comuns ao uso de antibióticos (BARTLETT *et al.*, 1978; GEORGE *et al.*, 1978).

Em 1978, o *C. difficile* foi identificado como a principal causa de colite pseudomembranosa isolado de fezes de pacientes submetidos ao tratamento com o antibiótico clindamicina (BARTLETT *et al.*, 1978; GEORGE *et al.*, 1978). Essas observações foram seguidas por uma série de outros relatos de casos que mostraram uma forte correlação entre a colite pseudomembranosa, uso da antibioticoterapia, colonização por *C. difficile* e produção de toxinas (BARTLETT; TEDESCO, 1980; BURDON *et al.*, 1981; GEORGE *et al.*, 1982; ARONSSON; MÖLLBY; NORD, 1981; DON; DAVIS, 1981). Esses estudos revelaram o *C. difficile* como um patógeno emergente capaz de causar doença gastrointestinal grave em indivíduos submetidos à terapia antibiótica (VOTH; BALLARD, 2005).

1.2 A biologia e características de *Clostridium difficile*

O *C. difficile* é um bacilo formador de esporos subterminais, pertencente ao filo Firmicutes e é encontrado ubiquamente tanto no ambiente quanto no trato gastrointestinal de seres humanos e animais (Figura 1) (RUPNIK; WILCOX; GERDING, 2009). Apesar do *C. difficile* ser responsável por diarreias infecciosas, incluindo a diarreia inflamatória, a colite associada a antibióticos e a colite pseudomembranosa, ele pode ser considerado parte integrante da microbiota intestinal normal de 1-3% de adultos saudáveis e 25% em crianças saudáveis com até 1 ano de vida (BARTLETT, 2006; SCHEEPERS *et al.*, 2014).

Figura 1 – Microbiologia de *C. difficile* e suas características morfológicas.



Fonte: Adaptado de Smits *et al.* (2017). Imagem típica de *C. difficile*. **A)** Colônias birrefringentes de *C. difficile* em uma placa de ágar sangue. **B)** Imagem de microscopia de contraste de fase de uma cultura de *C. difficile* com presença de células vegetativas (hastes alongadas, campo verde), esporos subterminais de fase escura (pontos escuros) e esporos elipsoides de fase brilhante.

A identificação de *C. difficile* antes do início da década de 80 era difícil, conseqüentemente as doenças que estavam relacionadas ao mesmo não eram diagnosticadas adequadamente. Em 1979, George e colaboradores desenvolveram um meio de cultura seletivo e diferencial chamado ágar CCFA (Agar Cicloserina, Cefoxitina e Frutose), meio cuja finalidade é facilitar o isolamento de *C. difficile* a partir de espécimes fecais. Os antibióticos cicloserina (500 µg/mL) e cefoxitina (16 µg/mL) presentes neste meio, inibem o crescimento da maioria das bactérias, mas não o de *C. difficile*. Nesse meio, em temperatura de 37°C por 48 horas de incubação, *C. difficile* forma colônias amareladas com diâmetro de aproximadamente 4 mm, com estrutura interna cristalina e birrefringente (GEORGE; SUTTER; CITRON, 1979). As colônias de *C. difficile* que crescem em meio CCFA apresentam características morfolologicamente diferentes e trazem propriedades fluorescentes suficientes para uma identificação presuntiva, porém, pelo crescimento não específico dessa bactéria neste meio, é necessária a identificação do isolado em cultura pura (LYERLY *et al.*, 1988).

Os métodos laboratoriais para detectar e identificar o *C. difficile* a partir de espécimes fecais evoluíram significativamente, uma vez que a relação etiológica com a diarreia e a colite

associada ao uso de antibióticos já foi estabelecida (SVENSSON; LASALA, 2012). No mercado existem muitos métodos disponíveis para detectar a presença das toxinas A e B produzidas por *C. difficile*, como, por exemplo, os imunoensaio enzimáticos, ensaio de neutralização de citotoxicidade celular, cultura toxigênica, reação em cadeia da polimerase (PCR) e dentre outros (PETERSON *et al.*, 2007; STAMPER *et al.*, 2009; SWINDELLS *et al.*, 2010; NORÉN *et al.*, 2011). No entanto, devido às vantagens e deficiências dos vários ensaios existentes para fornecer um diagnóstico rápido e preciso e por não existir ainda um método considerado padrão ouro no diagnóstico clínico de infecção por *C. difficile* (CDI); as culturas de fezes bacterianas continuam a ser o pilar da avaliação laboratorial, apesar de que o procedimento requerer de 3 a 5 dias para a real confirmação do patógeno entérico (SVENSSON; LASALA, 2012; DUNBAR; ZHANG; TANG, 2013).

1.3 Fatores de virulência e as toxinas de *Clostridium difficile*

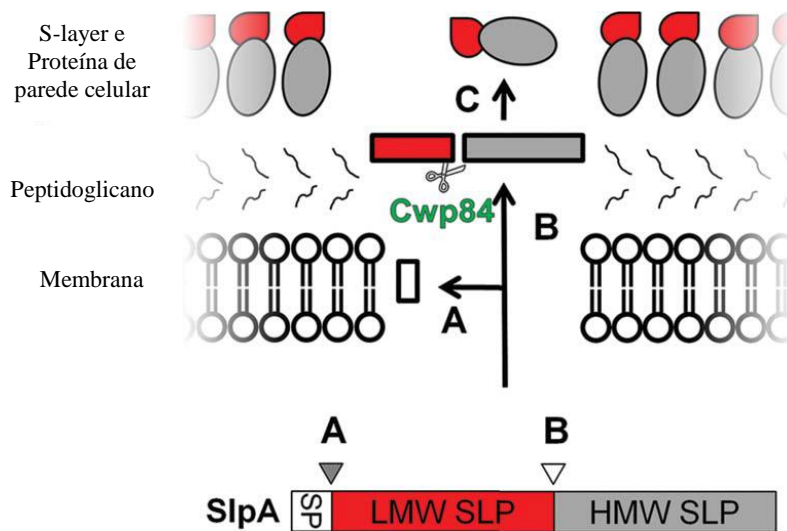
Vários fatores de virulência estão associados ao surgimento da infecção causada por *C. difficile* (CDI), como, por exemplo, as adesinas, os flagelos, as enzimas hidrolíticas, os esporos e as toxinas. Porém, dentre esses os mais importantes são as exotoxinas A e B (MATSUSHITA; OKABE, 2001; RUPNIK; WILCOX; GERDING, 2009; BABAN *et al.*, 2013; PAREDES-SABJA; SHEN; SORG, 2014). Para muitos patógenos entéricos, o primeiro passo fundamental para a colonização e, conseqüentemente, o estabelecimento da infecção, é a adesão às células da mucosa do hospedeiro. Nesse sentido, várias proteínas de superfície celular são responsáveis por coordenarem a adesão da célula bacteriana à parede do intestino e que já foram identificadas em cepas de *C. difficile* (AWAD *et al.*, 2014; CHILTON *et al.*, 2014).

Os principais fatores de virulência de *C. difficile* incluem:

- **Proteínas de superfície celular:** o processo de colonização, como já mencionado, inclui a adesão às células do epitélio intestinal, o qual envolve uma variedade de proteínas associadas à superfície celular do *C. difficile*. Entre essas proteínas está a de camada S (S-layer), que pode conter mais de 30 proteínas. As proteínas de camada S (SLPs) são formadas pelas proteínas de baixo (LMW SLP) e alto peso molecular (HMW SLP), codificadas por um único gene: *slpA*. Essas formas são derivadas quando o seu precursor, SlpA (proteínas A da camada superficial), sofre clivagem proteolítica pela cisteína protease Cwp84 (Figura 2). As SLPs são a família mais bem caracterizada e juntamente com outras proteínas, pertence à família das

proteínas de parede celular (*cell wall proteins*, CWPs). Membros dessa família incluem proteínas, como, a Cwp66 (exibe atividade de adesina), a Cwp84 (atividade proteolítica), Cwp13 (atividade proteolítica), CwpV (confere resistência a fagos) e outras (AWAD *et al.*, 2014; BRADSHAW *et al.*, 2018).

Figura 2 – Complexo proteico da camada S de *C. difficile*.



Fonte: adaptado de Awad *et al.* (2014). A camada S do *C. difficile* é composta por proteínas da camada superficial de alto e baixo peso molecular (HMW e LMW SLPs). A figura mostra os passos envolvidos na maturação do complexo proteico da camada S e que representa três estágios. **A)** a remoção do peptídeo sinal; **B)** a clivagem de SLP pela protease, Cwp84, para gerar os fragmentos HMW e LMW; e **C)** a formação da matriz da camada S pela reassociação dos SLPs LMW e HMW.

- **Flagelos:** importante fator de virulência para muitos patógenos entéricos, conferindo mobilidade e, conseqüentemente, desempenhando papel na invasão e colonização do hospedeiro, por permitir o acesso das bactérias aos tecidos da mucosa. Os flagelos da mesma forma como acontece com outras estruturas da superfície celular, podem atuar como adesinas. As duas proteínas flagelares mais bem caracterizadas e que contribuem para a patogênese da infecção por *C. difficile* (CDI), incluem a FliC (flagelina) e a FliD (proteína Cap flagelar) (AWAD *et al.*, 2014).

- **Esporos:** elemento chave para a disseminação do *C. difficile* sobreviver a longos períodos sob condições de estresse. Os esporos por serem metabolicamente inativos, são intrinsecamente resistentes aos antibióticos e uma vez liberados para o meio ambiente, possuem também resistência à maioria dos desinfetantes sem cloro comumente usados em ambientes hospitalares. Encontram-se ubiquamente no ambiente e são eliminados nas fezes

tanto de seres humanos como de animais, conduzindo as infecções em pacientes hospitalizados e na comunidade. Pelo fato dos esporos apresentarem várias camadas em sua estrutura, eles conseguem sobreviver e persistir em superfícies hospitalares por longos períodos de tempo (AWAD *et al.*, 2014; PAREDES-SABJA; SHEN; SORG, 2014).

- **Toxinas:** vários fatores de virulência contribuem para a colonização e permanência de *C. difficile* no trato gastrointestinal, porém, dentre os já mencionados, os principais e mais estudados são as toxinas A e B. Várias cepas patogênicas de *C. difficile* além de produzirem a toxina A (TcdA – proteína de 308 kDa) e a toxina B (TcdB – proteína de 270 kDa), que são responsáveis por causar diarreia e inflamação na mucosa intestinal, algumas cepas hipervirulentas caracterizada como NAP1 (pulsotipo 1 norte-americano), BI (análise de endonuclease de restrição), toxinotipo III e ribotipo por PCR 027, denominada NAP1/BI/027, produzem uma terceira toxina chamada transferase de *C. difficile* (CDT ou toxina binária) (DI BELLA *et al.*, 2016; QIU *et al.*, 2016; RUPNIK; JANEZIC, 2016).

As TcdA e TcdB pertencem à família de grande toxinas clostridiais (LCTs), que são um grupo de proteínas homólogas de alto peso molecular na qual incluem ainda outras toxinas, como, as toxinas letal e hemorrágica de *C. sordellii* (TcsL e TcsH, respectivamente), α -toxina de *C. novyi* (Tcn α) e uma citotoxina de *C. perfringens* (TpeL). As LCTs são glicosiltransferases que têm por característica glicosilar proteínas da família Rho GTPases promovendo alterações no citoesqueleto e conseqüentemente, interrupção da função celular hospedeira (CHANDRASEKARAN; LACY, 2017).

Entre as proteínas Rho, aquelas glicosiladas pelas toxinas A e B de *C. difficile* são RhoA, RhoB, RhoC, RhoG, Rac1, Rac2, Rac3, Cdc42 e TC10 (CHEN *et al.*, 2015). A sinalização celular dependente de Rho é bloqueada quando ocorre a glicosilação de proteínas Rho, inibindo sua interação com os efetores e, como resultado, leva o bloqueio de todas as vias de sinalização dependentes de Rho (SEHR *et al.*, 1998). Além disso, quando as proteínas Rho são glicosiladas, essas se ligam irreversivelmente às membranas celulares e impedem a translocação de Rho não modificado para as membranas para sinalização (GENTH; AKTORIES; JUST, 1999). Em conjunto, esses processos inibem a interação com os efetores, causando os chamados efeitos citopáticos (ruptura do citoesqueleto, perda de contato célula-célula e aumento da permeabilidade intestinal) e citotóxicos (ativação do inflamassoma, aumento dos níveis de espécies reativas de oxigênio e indução de morte celular programada) (CHAVES-OLARTE *et al.*, 2003)

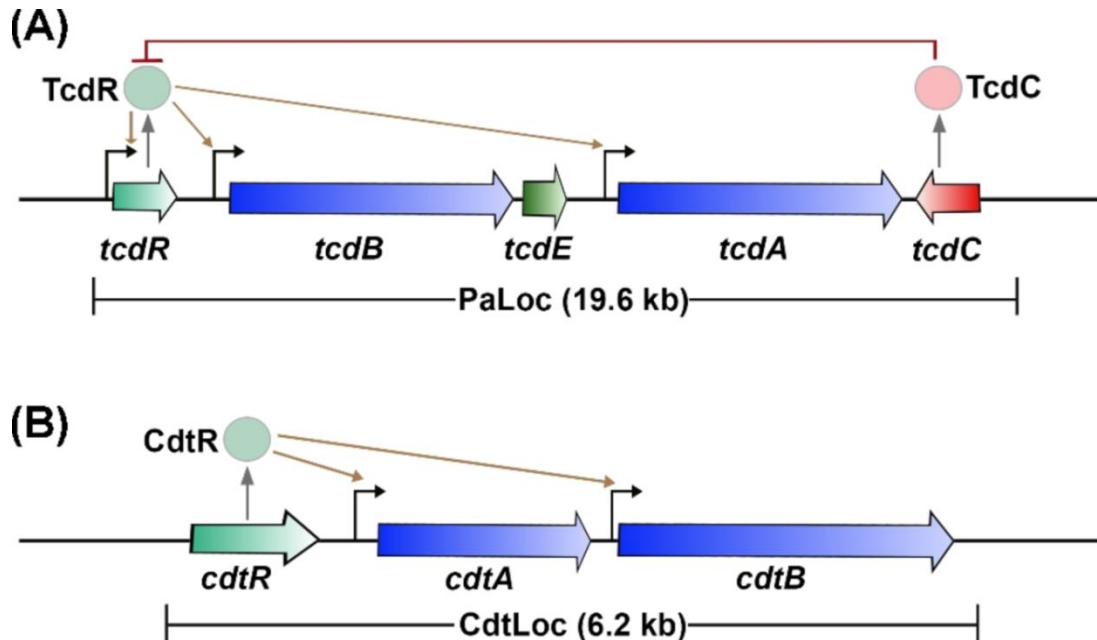
As cepas toxigênicas de *C. difficile*, que codificam os genes da TcdA (*tcdA*) e da TcdB (*tcdB*) estão localizados dentro de uma região cromossômica de 19,6 kb denominada locus de patogenicidade (PaLoc) (Figura 3A) (HAMMOND; JOHNSON, 1995; BRAUN *et al.*, 1996). No entanto, em cepas não toxigênicas o PaLoc é substituído por uma sequência não codificadora de 75-115 nucleotídeos ou uma sequência de 7,2 kb de função desconhecida (BRAUN *et al.*, 1996; DINGLE *et al.*, 2011; MONOT *et al.*, 2015). Por outro lado, cepas não toxigênicas de *C. difficile* podem adquirir a região PaLoc de cepas toxigênicas através de um fenômeno conhecido como transferência horizontal de genes, que irá resultar na conversão de cepas não toxigênicas em cepas produtoras de toxinas (BROUWER *et al.*, 2013). Alterações tais como, polimorfismos de nucleotídeos únicos, inserções ou deleções na região de codificação da toxina dentro do PaLoc, têm sido usadas para classificar ocorrência natural de isolados clínicos de *C. difficile* (RUPNIK *et al.*, 1998). Em 2016, Rupnik e Janezic fazendo uma comparação com uma cepa de referência a VPI10463, definiram 34 toxinotipos de *C. difficile*, mostrando com isso a grande heterogeneidade da região de codificação de toxina entre os isolados clínicos.

Além das toxinas descritas acima, ainda há dentro do locus de patogenicidade (PaLoc) genes para proteínas codificadas por cepas patogênicas (TcdR, TcdC e TcdE) que são envolvidas na regulação da produção e secreção de toxinas (Figura 3A) (BOUILLAUT *et al.*, 2015; MONOT *et al.*, 2015; SMITS *et al.*, 2016). A proteína TcdR é um membro da família da função extracitoplasmática de fatores sigma alternativos e que desempenha um papel crítico na ativação da expressão dos genes *tcdA* e *tcdB* (MONCRIEF; BARROSO; WILKINS, 1997; MANI; DUPUY, 2001). Além disso, o TcdR regula positivamente sua própria expressão (MANI *et al.*, 2002). Já a TcdC funciona como um fator anti-sigma que regula negativamente a expressão de toxinas (MATAMOUROS; ENGLAND; DUPUY, 2007; CARTER *et al.*, 2011). Por fim, a TcdE é uma proteína que compartilha homologia com as proteínas bacterinas do tipo holina e que pode facilitar a liberação das TcdA e TcdB ao meio extracelular (TAN; WEE; SONG, 2001).

A toxina binária (CDT), produzido por algumas cepas de *C. difficile*, é codificada por dois genes, *cdtA* e *cdtB*, que estão localizados em uma região cromossômica de 6,2 kb (distinta do PaLoc) denominada Locus CdtLoc (Figura 3B) (CARTER *et al.*, 2007). As cepas toxigênicas que são negativas para CDT normalmente contêm uma exclusão de 2 kb no CdtLoc (STARE; DELME; RUPNIK, 2007). Além disso, o CdtLoc também contém um terceiro gene, *cdtR*, que codifica um regulador de resposta da família LytTR (CARTER *et al.*, 2007). O CdtR regula positivamente a produção de CDT e, nas cepas epidêmicas do

ribotipo 027, também aumenta a produção de TcdA e TcdB (CARTER *et al.*, 2007; LYON *et al.*, 2016).

Figura 3 – Locus de patogenicidade (PaLoc) das toxinas produzidas por *C. difficile*.



Fonte: Chandrasekaran e Lacy (2017). **A)** Locus de patogenicidade (PaLoc). Os genes que codificam as toxinas, *tcdA* e *tcdB*, estão indicados por setas azuis. Os genes reguladores são mostrados em verde claro (*tcdR* – positivo) ou vermelho (*tcdC* – negativo). Por sua vez, o gene *tcdE* que codifica a holina é mostrado em verde escuro. A direção das setas reflete sentido da transcrição. A proteína TcdR regula positivamente a sua própria expressão, bem como a expressão dos genes *tcdA* e *tcdB* (indicada por setas castanhas). A proteína TcdC é um fator anti-sigma que regula negativamente a expressão das toxinas interferindo na função da TcdR. Já a proteína TcdE está envolvida na secreção de toxinas. **B)** Representação esquemática do locus da toxina binária (CdtLoc). Os genes que codificam a CDT, *cdtA* e *cdtB*, são mostrados em azul. O gene regulador *cdtR* é mostrado em verde-claro. O CdtR regula positivamente a transcrição de *cdtA* e *cdtB*.

As células de *C. difficile* cultivadas em meios ricos expressam tipicamente a TcdA e a TcdB durante a fase estacionária (DARKOH *et al.*, 2015). Vários estímulos ambientais têm sido relatados com a capacidade de influenciar na expressão das toxinas, incluindo a temperatura (KARLSSON *et al.*, 2003), concentrações subinibitórios de certos antibióticos (CHILTON *et al.*, 2012; ALDAPE *et al.*, 2013), sinalização de quórum (DARKOH *et al.*, 2015), a presença de uma fonte de carbono rapidamente metabolizável (DUPUY; SONENSHEIN, 1998) e certos aminoácidos (KARASAWA *et al.*, 1997). Com relação à fonte de carbono rapidamente metabolizável, tais como, a glicose presente no ambiente local da bactéria, esta inibe a produção de toxinas através da proteína de controle de catabólitos de carbono A (CcpA) (ANTUNES *et al.*, 2012). Por sua vez, os aminoácidos de cadeia

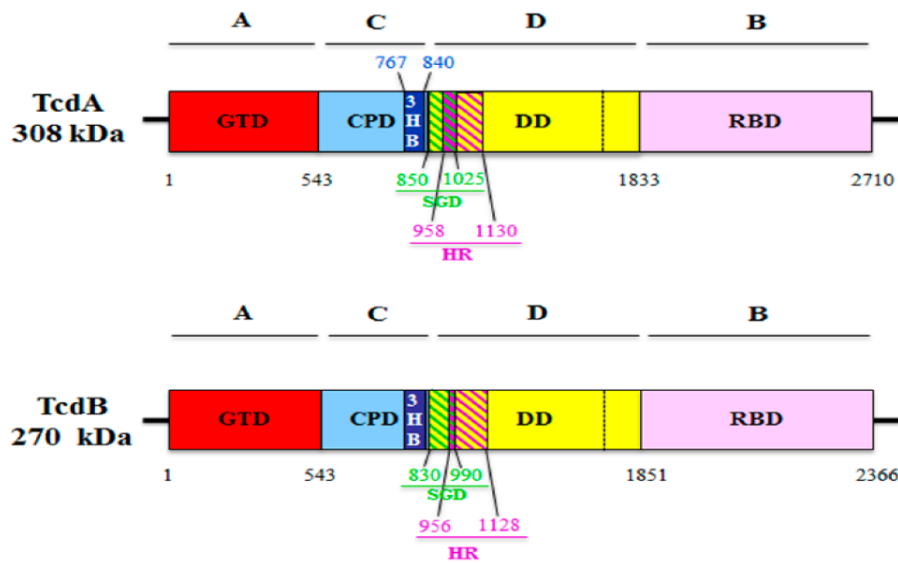
ramificada inibem a produção de toxina através do regulador transcricional global CodY (DINEEN *et al.*, 2007).

1.4 Domínios funcionais de TcdA e TcdB

A TcdA e a TcdB compartilham uma estrutura de domínio comum com 47% de sequência identidade e aproximadamente 68% de similaridade de sequência, sendo que a maior diversidade em sequência é confinada aos seus domínios de ligação C-terminal. A homologia entre as duas toxinas e o modo de entrada similar na célula hospedeira sugerem que essas adotam estruturas 3D semelhantes. Ambas as toxinas possuem um domínio "cabeça" globular e bilobada (correspondente ao domínio de entrega), um domínio de cauda longa (correspondente ao domínio de ligação ao receptor) e um curto domínio da cauda interna (que contém o domínio da glicosiltransferase) (PRUITT *et al.*, 2010; DAVIES *et al.*, 2011).

Em 2008, Jank e Aktories propuseram um modelo chamado ABCD para explicar que ambas as toxinas têm atividade enzimáticas similares e que compartilham uma estrutura de domínio multi-modular (Figura 4) (DI BELLA *et al.*, 2016). Esse modelo compreende os seguintes domínios com suas respectivas funções. O domínio A está localizado na extremidade N-terminal e contém o domínio da glicosiltransferase (GTD), sendo esse responsável por glicosilar as pequenas GTPases, as quais estão envolvidas na regulação do citoesqueleto (DI BELLA *et al.*, 2016; DAVIES *et al.* 2011); o domínio B está localizado na extremidade C-terminal e interage com estruturas de carboidratos como o Gal- α (1,3)-Gal β -(1,4)-GlcNAc sobre as células epiteliais em murinos, formando o domínio de ligação ao receptor (RBD), este forma combinações repetitivas de estruturas oligopeptídicas (CROPs) e o mesmo quando se liga ao íon cálcio aumenta a potência da toxina (JANK; GIESEMANN; AKTORIES, 2007; DI BELLA *et al.*, 2016); o domínio C corresponde ao domínio cisteína protease (CPD) que faz parte da região hidrofóbica e desempenha um papel na clivagem proteolítica da toxina, pois TcdA e TcdB sofrem um processo auto-catalítico de clivagem, na presença de InsP6 (EGERER *et al.*, 2007), e o domínio D (DD), região hidrofóbica, está envolvido na formação de poros antes da translocação do domínio A para o citosol (GENISYUERREK *et al.*, 2011).

Figura 4 – Organização dos quatro domínios funcionais das toxinas de *C. difficile*.



Fonte: Adaptado de Di Bella *et al.*, (2016). As duas toxinas são compostas pelos seguintes domínios: o domínio A, correspondente ao domínio N-terminal da glicosiltransferase (GTD) (vermelho); o domínio C, correspondente ao domínio da cisteína protease (CPD) (ciano) e o domínio do feixe de três hélices (3HB) (azul) identificado tanto na toxina A (TcdA) quanto na toxina B (TcdB) do *C. difficile*; o domínio D, que corresponde ao domínio hidrofóbico de entrega (DD) (amarelo), qual contém o pequeno domínio globular (SGD) em TcdA (linhas diagonais verdes e amarelas) que corresponde à região de formação de poros mínima (MPFR) em TcdB (verde e linhas diagonais amarelas), e sobrepondo-se apenas parcialmente (linhas diagonais roxas e verdes) para a região hidrofóbica (HR) de TcdA e TcdB (linhas diagonais roxas e amarelas); o domínio B, correspondente ao domínio de ligação ao receptor (RBD) (rosa).

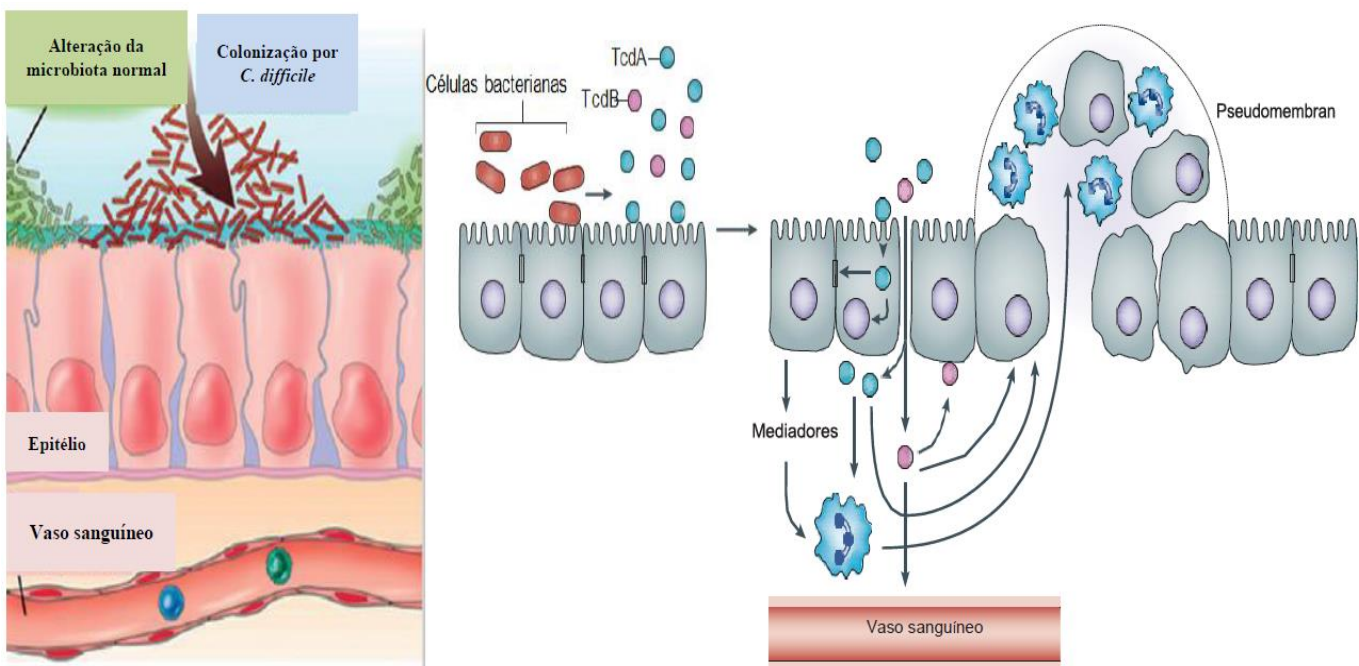
1.5 Patogênese da infecção causada por *C. difficile*

A patogênese da infecção por *C. difficile* (CDI) inicia com a ingestão das formas vegetativas ou esporuladas. Porém, a chave para o sucesso da CDI é sua forma de esporo dormente, que pode sobreviver por longos períodos mesmo em condições adversas (por exemplo, paredes de hospitais, portas, instrumentos cirúrgicos, telefones celulares, etc.) (SCHENCK; BECK; MACDONALD, 2015). É a forma de esporo que é transmitida no ambiente hospitalar, pois as células vegetativas ao chegarem ao estômago são mortas devido à acidez, mas os esporos devido à alta capacidade de sobrevivência, esse sobrevive e germina no intestino delgado por meio da ação dos ácidos biliares. A forma vegetativa metabolicamente ativa multiplica e pode colonizar o cólon humano para causar doença através da produção de toxinas e enzimas hidrolíticas (MCCOLLUM; RODRIGUEZ, 2012).

Vários fatores são necessários para que o *C. difficile* realmente cause doença (Figura 5), porém, o primeiro e mais importante é a disbiose intestinal, ou seja, a alteração de toda a microbiota intestinal normal causada geralmente pelo uso de antibióticos, sendo este

considerado o principal fator de risco para infecção por *C. difficile* (SMITS *et al.*, 2016; McCOLLUM; RODRIGUEZ, 2012). Com o desequilíbrio da homeostase intestinal, inicia a fase de colonização por *C. difficile*, envolvendo a ligação da bactéria ao epitélio intestinal do hospedeiro e a multiplicação tanto na superfície como no lúmen. Após a colonização, as células vegetativas começam a secretar TcdA e TcdB, que causam vários efeitos, como, por exemplo, rompimento das junções oclusivas, morte das células epiteliais e inflamação. Além disso, a lesão no epitélio colônico faz com que as toxinas estimulem este a libertar citocinas pró-inflamatórias e quimioatratores de neutrófilos, o que levam a uma resposta inflamatória aguda com recrutamento de neutrófilos, característica da fisiopatologia clínica da CDI (VOTH; BALLARD, 2005; KELLY; KYNE, 2011; SMITS *et al.*, 2016).

Figura 5 – Patogênese da infecção causada por *C. difficile*.

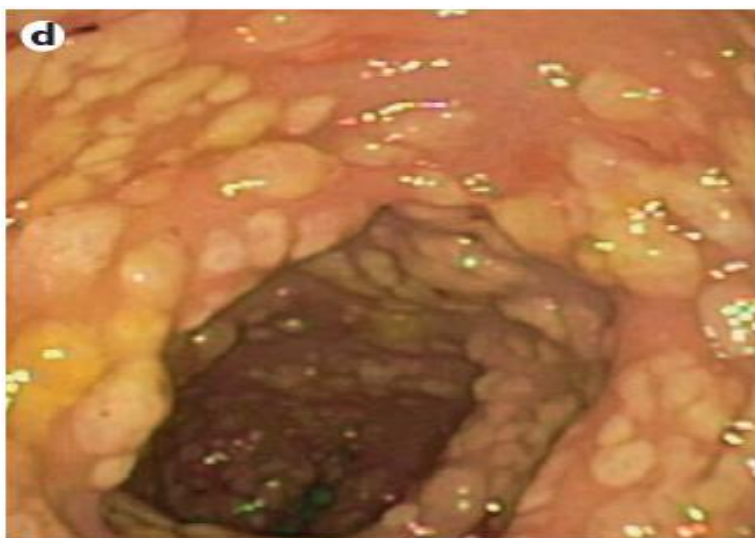


Fonte: Adaptado a partir de Mccollum; Rodriguez (2012) e Rupnik; Wilcox; Gerding (2009). O *C. difficile* coloniza o intestino (cólon) e após a alteração da microbiota intestinal normal, o *C. difficile* começa a secretar as toxinas. A TcdA liga-se ao lado apical da célula e, após a internalização, causa alterações no citoesqueleto de actina que resultam na ruptura das junções oclusivas e no afrouxamento da barreira epitelial, levando à morte celular e produção de mediadores inflamatórios que atraem neutrófilos. O rompimento das junções oclusivas permite que tanto a TcdA quanto a TcdB atravessem o epitélio e, conseqüentemente, induzem liberação de vários mediadores imunomoduladores tanto de células epiteliais, como de fagócitos e mastócitos, resultando em inflamação e acúmulo de neutrófilos. Esses achados contribuem para a formação das pseudomembranas.

Sabe-se que a ruptura da barreira epitelial, com intensa resposta inflamatória e infiltrado neutrofílico no lúmen e danos nos tecidos associados podem contribuir para a formação das pseudomembranas (SMITS *et al.*, 2016). Em 2002b, Brito e colaboradores

demonstraram que a TcdA atua em neutrófilos humanos induzindo mudança tanto na forma quanto na função, bem como também no rearranjo do citoesqueleto de actina e na aderência a substrato de matriz extracelular, com aumento de expressão de Mac-1 de leucócitos polimorfonucleares (PMNL), eventos esses que poderiam estar relacionados com a formação de pseudomembranas, como é observada em casos graves de colite pseudomembranosa (Figura 6).

Figura 6 – Endoscopia da colite pseudomembranosa causada por *C. difficile*.



Fonte: Adaptado de Smits *et al.* (2016). Imagem mostrando as pseudomembranas resultantes da infecção por *C. difficile* formadas por placas amareladas e confluentes.

1.6 Estudos *in vivo* e *in vitro* dos efeitos das TcdA e TcdB de *C. difficile*

Várias pesquisas têm evidenciado os efeitos das toxinas A e B (TcdA e TcdB) de *C. difficile* tanto em modelo animal quanto em cultura de células. Em 1989, Lima e colaboradores, utilizando modelo experimental em coelhos, mostraram que a injeção de TcdA na alça intestinal desses animais causa vários efeitos e, entre esses, foi observado que duas horas após a inoculação a TcdA causou citólise da porção basal das células epiteliais apicais e acúmulo significativo de fluido hemorrágico no lúmen intestinal, efeito este que se correlacionou com a gravidade das lesões histológicas observadas, além daqueles, a TcdA ocasionou ainda ruptura das barreira epitelial e uma intensa resposta inflamatória.

Em 1997, Rocha e colaboradores, estudando a migração de neutrófilos causada pela TcdA de *C. difficile* nas cavidades peritoneais e bolsas de ar subcutâneas de ratos, foi

observado que a TcdA atua como um potente indutor da migração neutrofílica, esta é em parte dependente de citocinas derivadas de macrófagos, como, TNF- α e IL-1 β e leucotrienos (LTB₄). Esses mediadores ajudam a explicar a intensa colite inflamatória observada em modelo animal experimental dessa doença induzida pela TcdA de *C. difficile*. Em 2001, o mesmo grupo de pesquisa, demonstrou que a TcdB, ao estimular macrófagos em experimentos *in vitro*, utilizando câmaras de Üssing, ao contrário da TcdA, não foi capaz de induzir secreção intestinal, nem alteração no transporte eletrogênico em íleo de coelho. Esses dados mostraram ainda que o sobrenadante de macrófagos estimulados com TcdB não estimulou o transporte iônico, porém, ambas as toxinas estimulam a gênese do TNF- α , mas somente a TcdA foi capaz de induzir a síntese de IL-1 β (ROCHA *et al.*, 2001).

Em 2008, o nosso grupo de pesquisa utilizando um modelo de enterite ileal em camundongos induzida pela TcdA, foi mostrado que o polissacarídeo fucoidina, um inibidor de moléculas de adesão celular (L-selectina) essencial para migração leucocitária, diminui a secreção intestinal e infiltração de células inflamatórias e também atue na redução do dano de mucosa intestinal (BARRETO *et al.*, 2008). Em 2006, o mesmo grupo de pesquisa avaliando o efeito de um novo agonista (ATL313) do receptor de adenosina A_{2A} em modelo de enterite, mostrou redução na secreção, lesão tecidual, infiltração de células inflamatórias e preveniu a morte celular induzida pela TcdA, revelando assim, uma outra possível via relevante para a fisiopatologia da colite causada pela TcdA (CAVALCANTE *et al.*, 2006). Em 2011, Junqueira e colaboradores, utilizando o EHNA, um inibidor da enzima adenosina desaminase 1, neste mesmo modelo, demonstrou que o EHNA reduziu a lesão tecidual, a infiltração neutrofílica e os níveis de citocinas pró-inflamatórias (TNF- α e IL-1 β), bem como na expressão de NOS-2 e NF- κ B.

Conforme Lyerly e colaboradores (1988), a lesão intestinal ocasionada pela TcdA de *C. difficile* está associada ao influxo de polimorfonucleares (neutrófilos) para dentro da mucosa colônica, um achado clínico observado na colite pseudomembranosa. Com base nisso, uma pesquisa publicada por Brito *et al.* (2002b) avaliou o efeito da TcdA de *C. difficile* sobre a morfologia e a função dos neutrófilos *in vitro*. Nesta pesquisa, foi mostrado que a TcdA induziu mudanças na forma de PMNL aderente de uma forma esférica ou piramidal compacta para uma forma fina e semelhante a corda. Estas alterações foram acompanhadas pelo rearranjo do citoesqueleto de F-actina em agregados. Os PMNL incubados com TcdA mostraram maior aderência a componentes de matriz extracelular e expressaram menos L-selectina e mais Mac-1, em comparação ao grupo não tratados, sendo estes efeitos ocasionados pela capacidade com que a TcdA tem de glicosilar as Rho GTPases em

neutrófilos humanos. Nesse mesmo ano, o grupo também mostrou mecanismos de como a TcdA induz a apoptose em células epiteliais intestinais do cólon humano (T-84), verificando que o processo de apoptose era dependente do tempo de incubação e da dose de TcdA utilizada, e que o processo de apoptose era completamente inibido pelo bloqueio da atividade enzimática da toxina na Rho GTPases com uridina-5-difosfato-2,3-dialdeído e parcialmente inibida quando se utilizava inibidores de caspases 1, 3, 6, 8 e 9. Portanto, a TcdA induzia a apoptose pelo mecanismo dependente da inativação de Rho, e da ativação das caspases efetoras (3 e 6) e das caspases iniciadoras (8 e 9), além da ativação do Bid e de provocar lesão mitocondrial com a liberação de citocromo c (BRITO *et al.*, 2002a).

Em 2014, Lima e colaboradores verificaram que a TcdA de *C. difficile* inibe de forma dose-dependente a via de sinalização Wnt/ β -catenina em cultura de células epiteliais intestinais de rato (IEC-6) e em células de adenocarcinoma de cólon humano (RKO), através do ensaio repórter de luciferase para fator de célula T (TCF). Nesse estudo, foi demonstrado que a pré-incubação de células RKO com TcdA por 12 h atenuou a ativação da via de Wnt mediada pela Wnt3a, sugerindo que a inativação das Rho GTPases tem um papel significativo nessa inibição. Logo, esses dados sugerem que a inibição dessa via pela TcdA é importante para os efeitos antiproliferativos da TcdA. Recentemente, uma pesquisa realizado por Tao *et al.* (2016) mostrou que as proteínas frizzled (FZD) são receptores epiteliais do cólon para a TcdB de *C. difficile*, onde esta inibiu a via Wnt/ β -catenina por ligar-se ao receptor Frizzled-7 (FZD-7) no epitélio colônico. A TcdB liga-se ao sítio conservado de ligação a Wnt, conhecido como domínio rico em cisteína (CRD), com uma maior afinidade para FZD1,2 e 7. Além de se ligar a esse receptor, impedindo a ligação de ativadores da via, como Wnt-3a, essa toxina mostrou utilizá-lo como meio para ser internalizada dentro dessas células. Esses achados estabeleceu os FZDs como receptores fisiologicamente relevantes para o TcdB no epitélio colônico. Sabendo-se que a TcdB de *C. difficile* tem como alvo os epitélios do cólon humano por ligar justamente as proteínas frizzled dos receptores Wnt, Chen *et al.* (2018) mostrou as bases estruturais para que ocorra o reconhecimento de proteínas frizzled por essa toxina utilizando uma estrutura cristalina de um fragmento de TcdB. Conforme demonstrado, um ácido graxo endógeno é responsável por ligar as FZD atuando como um co-receptor para ligação de TcdB e essa ligação entre o ácido graxo e TcdB pode está envolvido na patogênese da CDI.

Recentemente, o nosso grupo de pesquisa mostrou o envolvimento da via de sinalização TGF- β 1/SMAD, sendo essa ativada pela TcdA tanto *in vitro* (linhagem de células epiteliais intestinais de ratos (IEC-6) quanto *in vivo* (modelo de enterite ileal). Nesta pesquisa,

a TcdA aumentou a expressão de TGF- β 1 e seu receptor, T β RII, *in vitro* e *in vivo*. Além disso, a TcdA induziu a translocação nuclear dos fatores de transcrição SMAD2/3, um marco da ativação da via do TGF- β 1, tanto em células IEC-6 quanto na enterite. Foi mostrado ainda que a adição de TGF- β 1 recombinante (rTGF- β) evitou a apoptose/necrose induzida por TcdA e restaurou a proliferação e a atividade de reparo em células IEC-6 na presença de TcdA. Logo, estes dados sugerem que via do TGF- β 1 pode desempenhar um importante papel protetor contra os danos intestinais causada pela TcdA de *C. difficile* (TINOCO-VERAS *et al.*, 2017).

1.7 Infecção por *C. difficile*: epidemiologia global e o impacto econômico na saúde pública

Desde 2000, surtos de Infecção por *C. difficile* (CDI) tem ocorrido na América do Norte, Europa e Ásia, tornando assim, a principal causa de morbidade, mortalidade e aumento nos custos de saúde pública no mundo todo (ELLIOTT *et al.*, 2017). Estes surtos em escala global foi em grande parte impulsionado pelo surgimento de uma cepa hipervirulenta, a NAP1/BI/027 (pulsotipo 1 norte-americana, BI, ribotipo por PCR 027), que está relacionado ao uso aumentado de fluoroquinolonas (OWENS *et al.*, 2008). Além desse fator, outros podem explicar o aumento da virulência e o surgimento de novos surtos epidêmicos causados por essa cepa como: a capacidade para produzir as três toxinas já mencionadas; maior capacidade de esporulação e uma hiperprodução de toxinas devido a uma deleção de 18pb e uma mutação pontual na posição 117 no gene *tcdC*, gene regulador negativo da toxina (WARNY *et al.*, 2005). Embora a CDI seja considerada a principal causa de infecção em pacientes hospitalizados que recebem tratamento com antibiótico em longo prazo, a CDI também começou a surgir na comunidade, e em indivíduos mais jovens que não possuem os fatores de risco tradicionais (aumento da idade, comorbilidades e a aplicação de fármacos que reduzem o ácido gástrico, por exemplo, inibidores da bomba de prótons (IBP)), muitas vezes com maior incidência em mulheres (SCHÄFFLER; BREITRÜCK, 2018).

O aumento da incidência e gravidade da CDI levou a um grande ônus econômico nos sistemas de saúde devido aos custos associados ao tratamento e permanências prolongadas de pacientes no hospital. Esse ônus econômico é estimado em US \$ 5,4 bilhões em serviços de saúde e US \$ 725 milhões em ambientes comunitários na América do Norte (DESAI *et al.*, 2016). Além disso, estima-se que o *C. difficile* causou aproximadamente 453.000 infecções incidentes e foi associado com aproximadamente 29.000 mortes anualmente nos EUA, com

base em dados de vigilância ativa em termos populacionais e laboratoriais em diversas localizações geográficas dos EUA (LESSA *et al.*, 2015).

A epidemiologia da CDI na Europa é mais variável devido a diferentes sistemas de notificação dentro da União Europeia. Segundo Schäffler e Breitrück (2018), no entanto, por extrapolação dos dados do Reino Unido para a Europa, eles resultam em um total de 172.000 casos de CDI anualmente na União Europeia. Conforme Reichardt e colaboradores (2007), a disseminação epidêmica de cepas hipervirulentas de *C. difficile*, como, por exemplo, o ribotipo 027, levam ao aumento do número de surtos nosocomiais, que conseqüentemente, estão associadas a maior morbidade e mortalidade. Logo, o impacto econômico da CDI é enorme, levando a custos médicos adicionais de mais de um bilhão de dólares por ano nos EUA e três bilhões de euros por ano dentro da União Europeia (KUIJPER; COIGNARD; TU, 2006; DUBBERKE; OLSEN, 2012).

Sabe-se que a CDI é a principal causa de diarreia nosocomial e também um dos maiores problemas de saúde pública no mundo. Em 2005, um estudo realizado por Loo e colaboradores relataram um dos maiores surtos já registrados por CDI que ocorreu na região de Quebec no Canadá com um número 3x maior na incidência em hospitais além de um número mais expressivo de casos complicados na unidade de terapia intensiva com megacólon tóxico, colectomia ou morte. Nesse mesmo trabalho, os autores relataram ainda que dos 157 casos testados e isolados de cepas de *C. difficile*, 82,2% eram do tipo NAP1 levando aos pesquisadores sugerirem-na como uma cepa hipervirulenta.

Na América Latina os estudos sobre a CDI ainda são escassos, o que dificulta análises de parâmetros epidemiológicos. O problema da escassez de informações oficiais sobre a incidência da CDI deve-se em parte ao fato de que a detecção de bactérias anaeróbias não é um procedimento comum de rotina nos laboratórios clínicos em nosso país, principalmente devido à falta de tecnologias e instalações para o seu cultivo, contribuindo, portanto, para a subnotificação de CDI (BALASSIANO *et al.*, 2009). Entretanto, Quesada-Gómez *et al.* (2010) descreveu de forma pioneira, a CDI por NAP1/027 na América Latina em 2009 ao avaliar sua presença em amostras de fezes de 37 pacientes internados no hospital da Costa Rica.

Atualmente, no Brasil poucos grupos de pesquisa dedicam-se a investigar micro-organismos anaeróbios no país. Entre esses, a maioria estão localizados no Rio de Janeiro, São Paulo, Belo Horizonte, Porto Alegre e Fortaleza (BALASSIANO *et al.*, 2012; PIRES *et al.*, 2014; SILVA *et al.*, 2015; SILVA *et al.*, 2017; COSTA *et al.*, 2016). Até 2012, a literatura científica continha somente um artigo sobre CDI no Nordeste do Brasil e, este citava

um paciente assintomático positivo para as toxinas de *C. difficile* no Hospital Walter Cantídio (MESQUITA *et al.*, 2011).

A partir de 2013 nosso grupo de pesquisa deu início aos trabalhos na investigação da incidência e caracterização fenotípica e genotípica de cepas de *C. difficile* isoladas de pacientes com diarreia em hospitais cearenses. Em 2014, o nosso grupo de pesquisa publicou o primeiro relato de caso documentado na América Latina de uma paciente com infecção por *C. difficile* adquirida na comunidade, sendo a mesma portadora de HIV e câncer, ressaltando a importância dessa infecção no nordeste do Brasil (COSTA *et al.*, 2014).

A continuidade desse estudo proporcionou o isolamento de uma cepa de *C. difficile* que até então era desconhecida pela comunidade científica, denominada de ICC-45. Por meio de uma abordagem polifásica foi atribuída a cepa ICC-45 um novo padrão de macrorestrição de ribotipagem de PCR e PFGE (Eletroforese de Campo Pulsado), que além de ser toxigênica ($tcdA^+$, $tcdB^+$ e $ctdB^+$), foi classificada como MLST Clade 2 (tipagem de sequência multilocular) a qual pertence a linhagem dos Clades das outras cepas consideradas hipervirulentas, como, a NAP1/027. Além disso, nessa mesma pesquisa, foi mostrado que a cepa ICC-45 produziu uma toxina B variante. Costa *et al.* (2016) mostrou ainda que a cepa ICC-45 codifica para uma variante de TcdB a qual induziu em cultura de células HeLa um efeito citopático distinto de acordo com o seu toxinotipo. Dentro da mesma pesquisa, um dado que chamou muita atenção foi que, mesmo a cepa ICC-45 produzindo menos toxina que a NAP1, os sobrenadantes das cepas NAP1 e ICC-45 mostraram potencial virulento semelhante, já que elas produziram uma expressão similar de citocinas pró-inflamatórias, danos epiteliais e resposta ao estresse oxidativo no modelo de alça ileal em camundongos, mostrando assim as características comuns na patogênese da CDI (COSTA *et al.*, 2016).

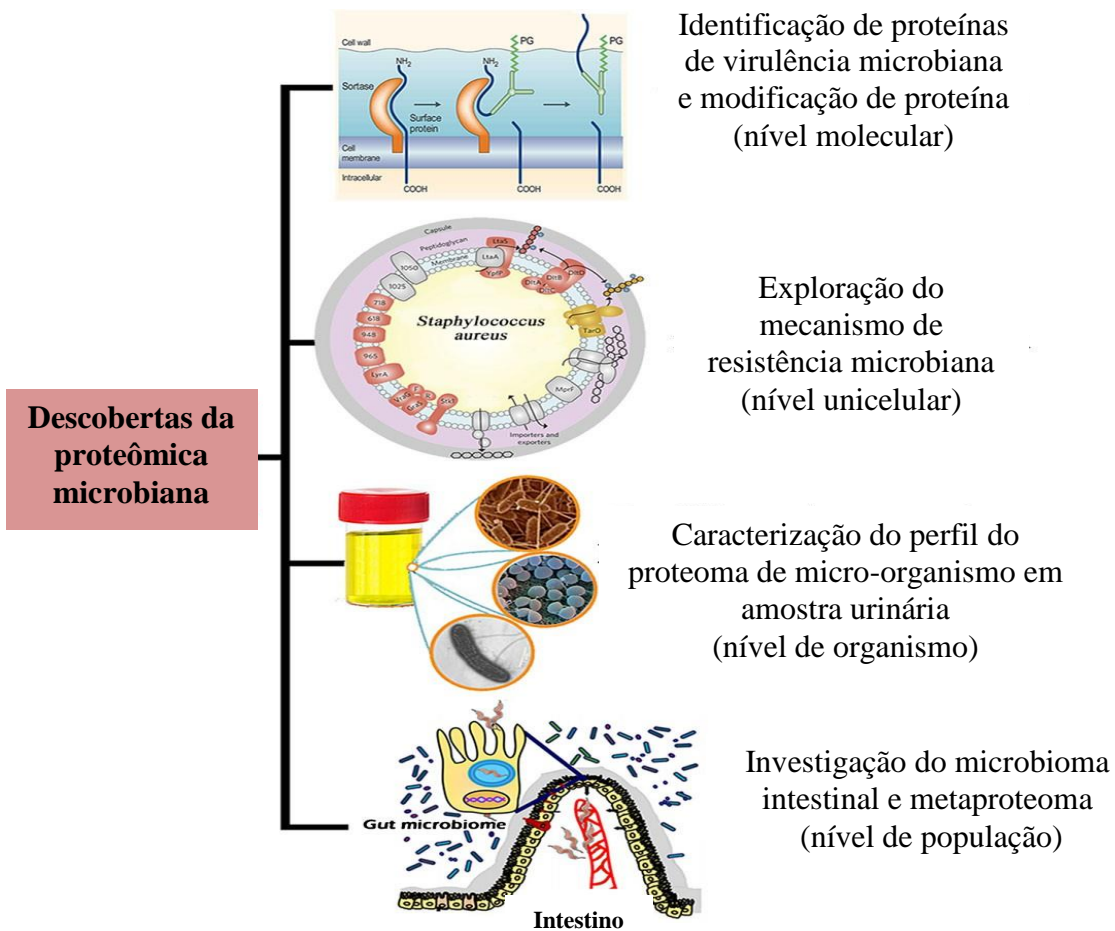
Em 2017, um estudo realizado em um hospital oncológico no Nordeste do Brasil (Hospital Haroldo Juaçaba do Instituto do Câncer do Ceará (ICC)), Costa e colaboradores demonstraram alto nível de positividade para CDI (48%, 23/48) em pacientes com câncer e diarreia. Essa elevada incidência de CDI observado no ICC, teve como cepa predominantemente identificada a NAP4, encontrada também em outros países, como, EUA e Canadá, presentes em hospitais e em casos na comunidade. Logo, todos esses dados enfatizam a importância de uma melhor compreensão da situação epidemiológica dessa infecção tanto em hospitais brasileiros e comunidades quanto nos outros países (COSTA *et al.*, 2017).

1.8 Análise proteômica

A proteômica (proteoma), diferentemente da genômica consiste no conjunto de proteínas codificadas, seja por uma célula, tecido ou mesmo por um organismo, e que são expressas por um genoma, sob determinadas condições. Por sua vez, a análise proteômica irá consistir em um estudo detalhado das proteínas utilizando técnicas de separação e identificação, tais como, eletroforese uni (1D) e bidimensional (2D), cromatografia líquida, espectrometria de massas (MS) e a bioinformática (HEIN *et al.*, 2013). Sabendo disso, essa ferramenta tem ganhado um importante papel na era pós-genômica, pois vem sendo utilizada em várias pesquisas e aplicações, desde identificação bacteriana rápida em surtos de doenças transmitidas por alimentos, diagnóstico rápido de doenças infecciosas, descobertas de novos biomarcadores que podem ajudar na detecção precoce de doenças, como também monitorar com precisão a evolução da mesma frente a um determinado tratamento e até mesmo nas interações entre patógeno-microbiano-hospedeiro (Figura 7) (CHENG *et al.*, 2016; CHEN *et al.*, 2017).

Com o avanço da proteômica, diversas terminologias ômicas surgiram e que atualmente vem sendo utilizadas na pesquisa, tais como, a secretoma (análise das proteínas secretadas), a metaboloma (análise de perfil dos metabólitos), a transcriptoma (análise de transcritos de mRNA) e a exoproteoma, termo genérico utilizado para descrever o “proteoma externo”, ou seja, o subconjunto de proteínas presentes no meio extracelular (ARMENGAUD; DUPORT, 2017; VINUSHA *et al.*, 2018). As proteínas obtidas do exoproteoma podem ser resultantes da secreção celular, de mecanismos de exportação ou até mesmo da lise celular. Logo, essa abordagem tem como objetivo descrever e quantificar as proteínas localizadas fora das células, utilizando as metodologias empregadas na proteômica (ARMENGAUD; DUPORT, 2017).

Figura 7– Interações patógeno-microbiano-hospedeiro por meio de ferramentas proteômica em diferentes níveis (moleculares, unicelulares, de organismos e populacionais).



Fonte: Adaptado de Chen *et al.* (2017). A imagem mostra o perfil proteômico de proteínas de virulência microbiana e suas modificações, como, os proteoglicanos (BARTLETT; PARK, 2010); a identificação de proteínas bacterianas modificadas de cepa *S. aureus* em resposta ao tratamento com o antibiótico daptomicina (MA *et al.*, 2017); a identificação de biomarcadores para diagnósticos e prognósticos de doenças humanas encontrados em fluidos corporais, como a urina (YU *et al.*, 2015) e abordagens metaproteômicas combinadas com a metagenômica para identificar potenciais assinaturas funcionais na doença de Crohn em amostras de fezes, revelando proteínas e vias que diferenciaram principalmente indivíduos com doença de Crohn no íleo de pacientes saudáveis (ERICKSON *et al.*, 2012).

A análise proteômica nos permite obter informações qualitativas sobre proteínas, incluindo a identificação, distribuição, modificações pós-traducionais (PTMs), interações, estrutura e função, além de informações quantitativas, como abundância, sua distribuição em diferentes localizações e mudanças temporais na síntese e degradação (OTTO *et al.*, 2012). Isso só é permitido devido ao uso de técnicas analíticas extremamente sensíveis e precisas, como, a espectrometria de massas (MS), ferramenta ideal que vem sendo aplicada dentro da proteômica, devido sua capacidade de analisar uma gama de proteínas e peptídeos ao mesmo tempo com alta especificidade, precisão, velocidade e sensibilidade (ZUBAREV;

MAKAROV, 2013). Apesar da MS ser uma técnica centenária, a autoria dessa pode ser atribuída à introdução das técnicas de ionização/dessorção a laser assistida por matriz (MALDI) e a ionização por eletrospray (ESI) ocorridas na década de 1980, para os quais Koichi Tanaka e John Fenn, respectivamente, foram agraciados com o Prêmio Nobel de Química em 2002. O desenvolvimento dessas técnicas de ionização possibilitou a transferência eficiente de grandes biomoléculas intactas para a fase gasosa como íons carregados – um pré-requisito essencial para a análise de MS, como, por exemplo, peptídeos e proteínas (LÖSSL; VAN DE WATERBEEMD; HECK, 2016). Logo, independentemente da abordagem a ser empregada na análise proteômica, esta em uma de suas etapas do processo de análise irá empregar sempre a MS (AEBERSOLD; MANN, 2003).

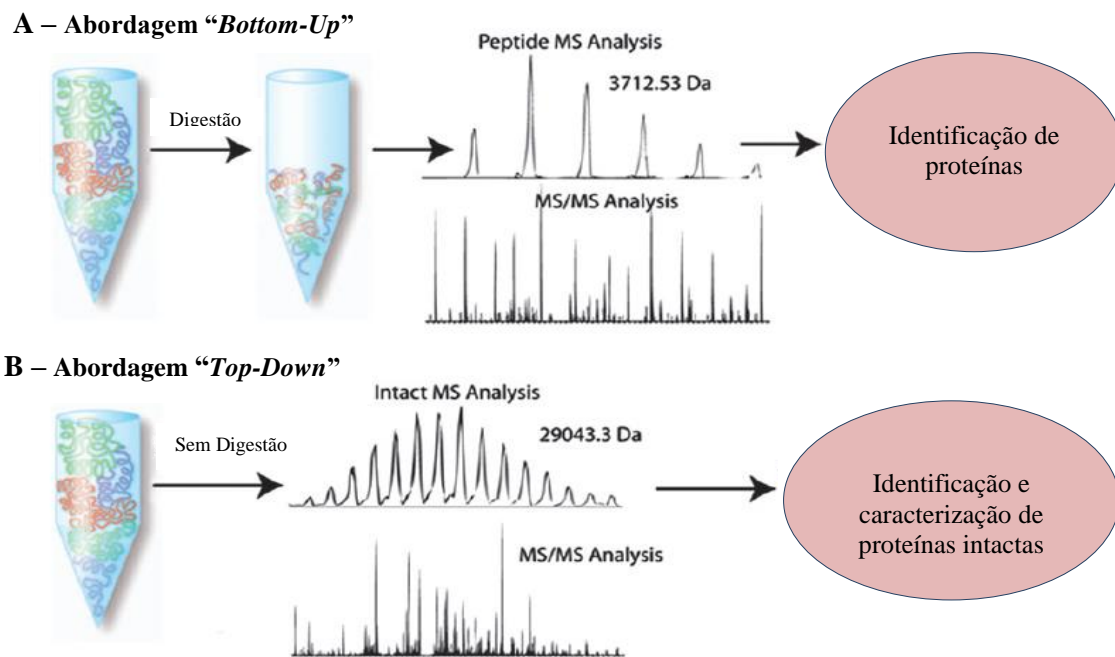
1.8.1 Abordagem proteômica “*bottom-up*” e “*top-down*”: estratégias para identificação de proteínas

A análise proteica baseada em MS e a proteômica tornaram-se ferramentas essenciais para identificar, quantificar e caracterizar as moléculas envolvidas em determinado processo. Logo, em se tratando de estudos proteômicos existem basicamente duas maneiras para que uma amostra seja analisada por MS, sendo essas classificadas em “*bottom-up*” e “*top-down*” (Figura 8) (KELLIE *et al.*, 2010). A proteômica “*bottom-up*” tem sido tradicionalmente realizada e essa implica na digestão química ou enzimática das proteínas antes da sua introdução no MS, ou seja, representa uma forma indireta de analisar o analito original (a proteína intacta), que será primeiramente hidrolisada para produzir pequenos peptídeos (KELLIE *et al.*, 2010; THOLEY; BECKER, 2017). Desta forma, a detecção e identificação tipicamente baseada em pequenos fragmentos dos peptídeos permite inferir a identificação da proteína original. Porém, imediatamente, tornam-se evidentes várias desvantagens com relação a essa abordagem: um peptídeo ou até mesmo vários peptídeos podem não ser específicos para uma determinada proteína ou proteína individual; grandes regiões da proteína podem não ser identificadas, o que pode deixar informações importantes sobre as PTMs ou variantes de sequência e modificações ou podem ainda ocorrer variações de sequência em peptídeos diferentes, fazendo com que a sua relação entre si se perca após a digestão (SMITH; KELLEHER, 2013).

A fim de contornar os inconvenientes da proteômica “*bottom-up*”, surge então a “*top-down*” que consiste em uma análise direta das proteínas intactas no próprio MS (sem digestão enzimática). Porém, para que isso aconteça as amostras precisam ser fracionadas antes da

introdução ao MS e um dos métodos empregados para a separação de proteínas intactas, peptídeos e moléculas pequenas é a cromatografia líquida (LC). A abordagem “*top-down*” por manter as proteoformas intactas e as suas massas, esta descreve a sequência completa de aminoácidos incluindo todas as PTMs que as proteínas podem sofrer. Assim, nessa abordagem há uma cobertura de 100% da sequência, no entanto, esse método requer o uso de MS de altíssima resolução e exatidão de massa, o que torna altíssimo o custo dos equipamentos que adotam como estratégia a “*top-down*”, isso faz com que a abordagem tradicional seja ainda a mais empregada em estudos proteômicos, apesar de suas limitações (KELLIE *et al.*, 2010; CAPRIOTTI *et al.*, 2011; SMITH; KELLEHER, 2013).

Figura 8 – Abordagens proteômica “*bottom-up*” e “*top-down*” para identificação de proteínas.

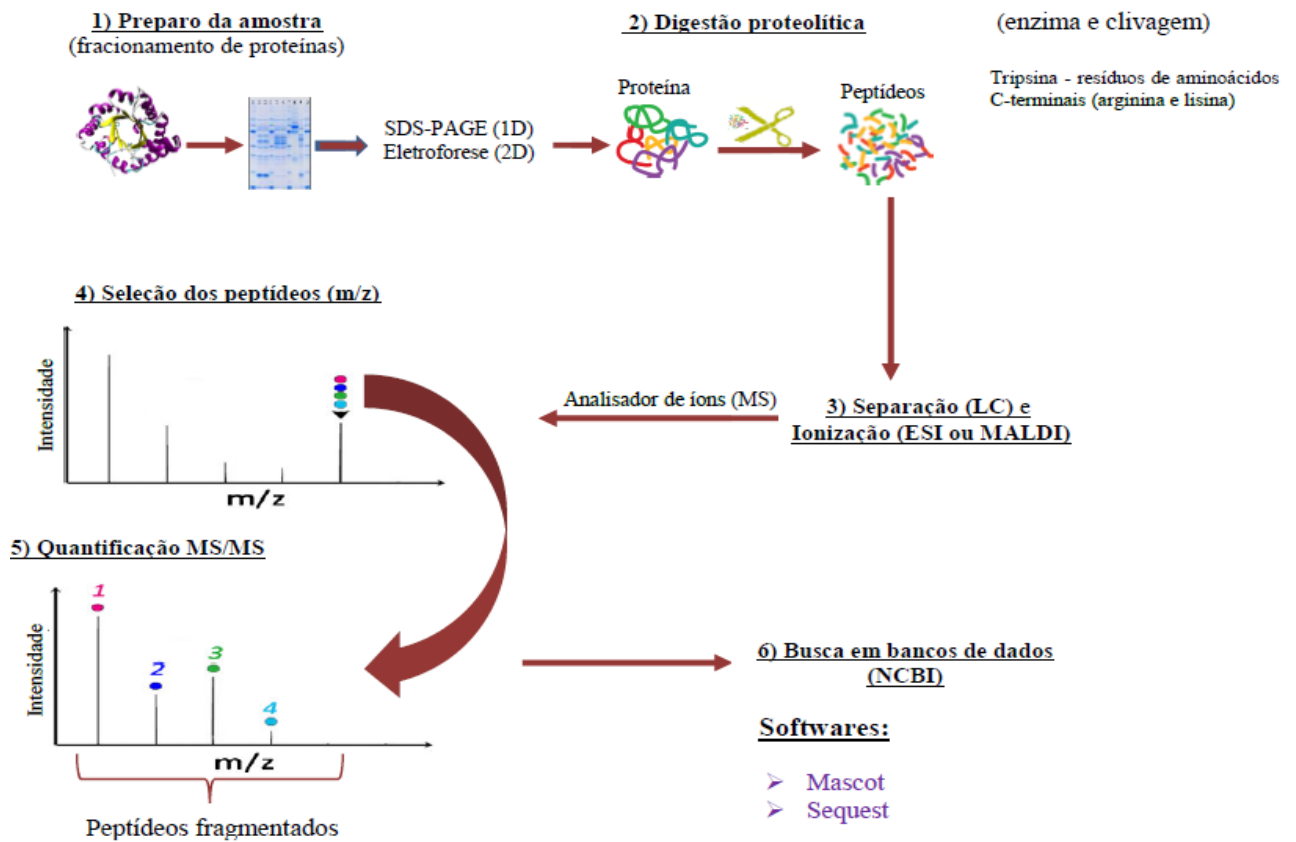


Fonte: Adaptado de Kellie *et al.* (2010). Na abordagem “*bottom-up*” (A), utiliza-se a digestão enzimática para clivar proteínas intactas em peptídeos. Os peptídeos clivados por sua vez, são analisados através de espectrometria de massa em tandem, e a identificação de proteínas pode ocorrer através da identificação de peptídeos emparelhados *in silico* a uma base de dados de proteínas por um algoritmo de pesquisa. Embora coberturas de baixa sequência sejam típicas, coberturas de proteoma muito altas são comuns. Por outro lado, na abordagem “*top-down*” (B), as proteínas intactas são diretamente analisadas no espectrômetro de massa (sem digestão enzimática). O precursor resultante e as massas dos fragmentos são então adaptados às sequências candidatas a partir de uma base de dados de proteínas. Isso permite o potencial de localização e a caracterização completa das modificações pós-traducionais que as proteínas sofrem.

1.8.2 Delineamento de um estudo proteômico “*bottom-up*”: análise de peptídeos por MS

Conforme já mencionado, a abordagem proteômica “*bottom-up*” que consiste na caracterização de peptídeos resultantes da digestão de proteínas em solução, também pode ser chamada de proteômica “*shotgun*” (TUBAON; HADDAD; QUIRINO, 2017). Dessa forma, a estratégia proteômica “*bottom-up*” pode ser descrita resumidamente em seis etapas (Figura 9), sendo que a primeira é facultativa, pois depende do tipo de abordagem a ser utilizada. Entre as etapas estão: **1)** A mistura de proteínas extraídas (células, lisado tecidual ou fracionamento subcelular) será submetida a uma etapa de fracionamento de proteínas por eletroforese 1D ou 2D e LC; **2)** Digestão proteolítica – protease mais comumente usada é a tripsina, uma serina protease que cliva os resíduos de aminoácidos C-terminais (arginina e lisina) resultando em peptídeos multivariados; **3)** Os peptídeos agora obtidos podem ser separados de três formas (LC – troca iônica/fase reversa; ESI ou MALDI); **4)** Após a separação dos peptídeos, o espectro de massas dos mesmos provenientes da digestão proteolítica são adquiridos e estes indicam a razão massa/carga (m/z) dos peptídeos analisados; **5)** Software de aquisição de MS são utilizados para selecionar um número pré-definido de peptídeos nos espectros de massa (m/z) e proceder o isolamento de cada um deles. Nessa mesma etapa, os peptídeos são fragmentados no MS para medir os m/z dos fragmentos – procedimento chamado de espectrometria de massa em tandem (MS/MS); **6)** Cada peptídeo fragmentado é seguido por 5 a 20 espectros de massa em tandem. Por sua vez, a massa peptídica é então utilizada em conjunto com as massas dos seus fragmentos para identificação do peptídeos. Ao final do processo, esses dados são rastreados por softwares de busca através de um bando de dados de sequências de aminoácidos dos peptídeos por algoritmos que calculam o espectro previsto para cada possível sequência peptídica linear identificada e devolvem a que mais provavelmente deu origem ao espectro de massa e aos dados do espectro de massa em tandem. Os principais softwares utilizados para localizar as proteínas em bancos de dados, como, por exemplo, o NCBI são o Mascot e o Sequest (COX; MANN, 2011; ZHANG; GE, 2011; ZHANG *et al.*, 2013).

Figura 9 – Representação esquemática das seis etapas experimentais utilizando abordagem proteômica “bottom-up”.



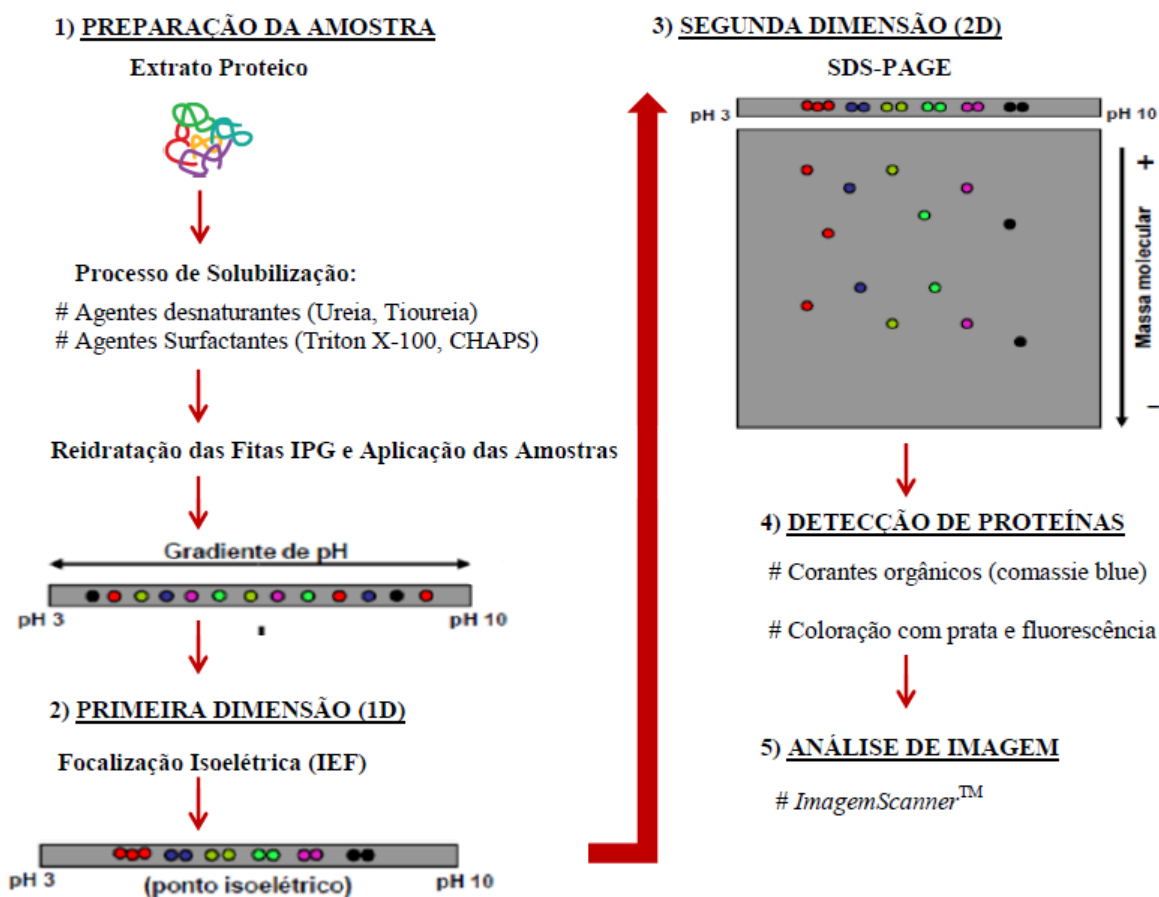
Fonte: Elaborado pelo autor.

A abordagem proteômica “bottom-up”, apesar de suas limitações como já mencionadas, continua sendo bastante empregada no campo da análise proteômica, pois essa estratégia engloba ferramentas analíticas, como, a eletroforese 2D e a MS que permitem alta resolução na separação e caracterização de proteínas, colocando-a como uma tecnologia central para análises e identificação de milhares de proteínas a mesmo tempo (CATHERMAN; SKINNER; KELLEHER, 2014; CHEN *et al.*, 2017). Além disso, outra ferramenta empregada na proteômica, a eletroforese em gel diferencial bidimensional (2D-DIGE), permite realizar uma análise comparativa entre duas ou três amostras de proteínas no mesmo gel, podendo estabelecer assim níveis de expressão proteica diferente entre grupos de amostras de tecido patológico normais e doentes utilizando uma marcação com diferentes corantes fluorescentes (HOLLAND, 2018).

1.8.3 Eletroforese bidimensional

A eletroforese bidimensional em gel de poliacrilamida (2D-E), ferramenta de separação mais utilizada na proteômica, foi desenvolvida inicialmente por O'Farrell em 1975 e atualmente é considerada um método poderoso para separar misturas complexas de proteínas (células, tecidos ou outras amostra biológicas) para análise proteômica (O'FARRELL, 1975; ZHANG *et al.*, 2013). Essa técnica permite a separação das proteínas em duas dimensões, sendo que, na primeira dimensão (1D) consiste numa focalização isoeétrica (IEF), ou seja, as proteínas são separadas de acordo com a sua carga elétrica em um gel de poliacrilamida com um gradiente de pH – ponto isoetrico (pI) da proteína. Já na segunda dimensão (2D), as proteínas que foram previamente separadas pela IEF são submetidas a uma eletroforese em gel de poliacrilamida com dodecil sulfato de sódio (SDS-PAGE), no qual agora são separadas de acordo com suas massas moleculares (MM). Portanto, a 2D-E consiste em cinco etapas principais que se encontra resumidamente abaixo (Figura 10) (RABILLOUD; LELONG, 2011).

Figura 10 – Esquema do princípio das cinco etapas de separação proteica por eletroforese 2D.



Fonte: Elaborado pelo autor. **1.** O processo começa com a extração de proteínas de uma amostra biológica que em seguida deve passar por um processo de solubilização (quebra das interações macromoleculares) para obter uma amostra compatível com a IEF. Porém, antes da IEF as fitas IPG (immobilized pH gradiente - IPG) têm que ser reidratada; **2.** A amostra é então carregada em um gradiente de pH orientado com o lado ácido no ânodo e o lado básico no cátodo. Após a IEF (1D), as proteínas atingiram seu ponto isoelétrico (pI) e, portanto, não possuem carga elétrica remanescente. A tira é então equilibrada num tampão contendo SDS, de modo que todas as proteínas se tornam fortemente carregadas negativamente; **3.** O gel da IEF é então carregado em cima de um gel SDS-PAGE, e as proteínas são separadas agora de acordo com suas massas moleculares; **4.** Após este passo, as proteínas são detectadas diretamente no gel podendo ser utilizado três tipos principais de coloração: corantes orgânicos (comassie blue), coloração com prata e fluorescência; **5.** Por último, os géis 2D são analisados por um scanner específico, como, por exemplo, o ImageScannerTM.

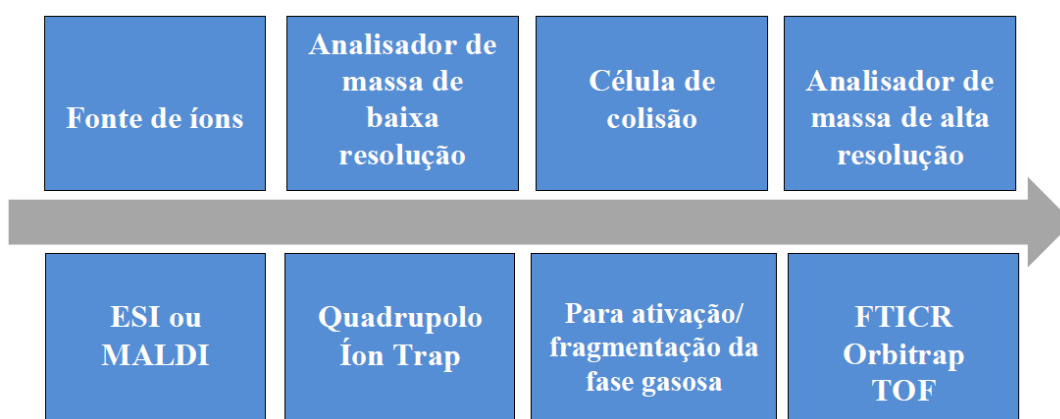
1.8.4 Espectrometria de massas

O advento da espectrometria de massas (MS) como uma tecnologia analítica, remonta ao final do século XIX, com o trabalho inovador sobre os raios catódicos do físico J. J. Thomson em 1897 e que tem sido amplamente utilizada até hoje para analisar vários compostos orgânicos de baixa massa molecular (LÖSSL; VAN DE WATERBEEMD; HECK, 2016). O surgimento da proteômica possibilitou grandes avanços tecnológicos na MS, esses beneficiaram não só a proteômica, como também impactou positivamente a MS no campo da biologia estrutural e molecular por ser considerada uma técnica altamente sensível e precisa. Logo, isso permitiu com que a MS entrasse no campo da biologia celular, molecular e estrutural possibilitando essa técnica analisar grandes biomoléculas e biopartículas numa faixa de 1 nm a 100 μ m, como, por exemplo, complexos proteicos, vírus, bactérias e células (GESSEL; NORRIS; CAPRIOLI, 2014; PENG; CHOU; PATIL, 2014).

A MS é uma técnica que se baseia numa relação massa e carga (m/z) de moléculas que são ionizadas em fase gasosa. Os espectrômetros de massa são constituídos por três componentes principais: uma fonte de ionização, um analisador de massas e um detector. Inicialmente, os analitos são separados de seu meio de transporte (fase aquosa ou orgânica), sendo então transformados em íons gasosos na fonte de íons. Após a ionização, os analitos ficam carregados com uma carga, o qual por sua vez é transmitido para as regiões de alto vácuo do espectrômetro de massas, incluindo geralmente um analisador de massa de baixa resolução e uma célula de colisão. A caminho do detector, os analitos ao passarem por vários campos elétricos e/ou magnéticos, são separados dos demais que possuem ainda carga neutra. Na última fase, as relações m/z dos analitos são medidas de forma precisa por um analisador de massa de alta resolução (Figura 11) (HAAG, 2016; LÖSSL; VAN DE WATERBEEMD; HECK, 2016).

Atualmente, a habilidade com que a MS tem de analisar proteínas e amostras biológicas, advém do surgimento das técnicas de ionização branda, como, a técnica de ionização por electrospray (ESI-MS) e a ionização/dessorção a laser assistida por matriz (MALDI) - técnicas capazes de transformar macromoléculas em íons. O termo ionização branda significa que uma quantidade mínima de energia é transmitida para o analito durante a etapa de ionização. Portanto, independentemente da técnica de ionização utilizada, a sensibilidade e acurácia do espectrômetro de massas está diretamente relacionada com o tipo de analisador de massa empregado na separação dos íons obtidos. Por esse motivo, a escolha do analisador é de fundamental importância, pois depende das propriedades do analito após a ionização e dos requisitos do experimento que será realizado (KARAS; HILLENKAMP, 1988; FENN *et al.*, 1989).

Figura 11 – Componentes essenciais de um espectrômetro de massa.



Fonte: Adaptado de Lössl; Van De Waterbeemd; Heck (2016). A fonte de íons facilita a transferência do analito para a fase gasosa. Por sua vez, o analisador de massa de baixa resolução permite a seleção em massa de espécies específicas de analito e, no caso de Íon Trap, também pode ser utilizado para ativação/fragmentação de íons. Caso contrário, a fragmentação da fase gasosa, por exemplo, usando CID (dissociação induzida por colisão), HCD (dissociação colisional de alta energia) ou ETD (dissociação de transferência de elétrons), ocorre na célula de colisão. Por último, um analisador de massa FTICR (Ressonância de ciclotrão de íons com transformada de Fourier), TOF (Tempo de voo) ou Orbitrap de alta resolução facilita medições de massa precisas.

1.8.5 Técnica de ionização por electrospray (ESI)

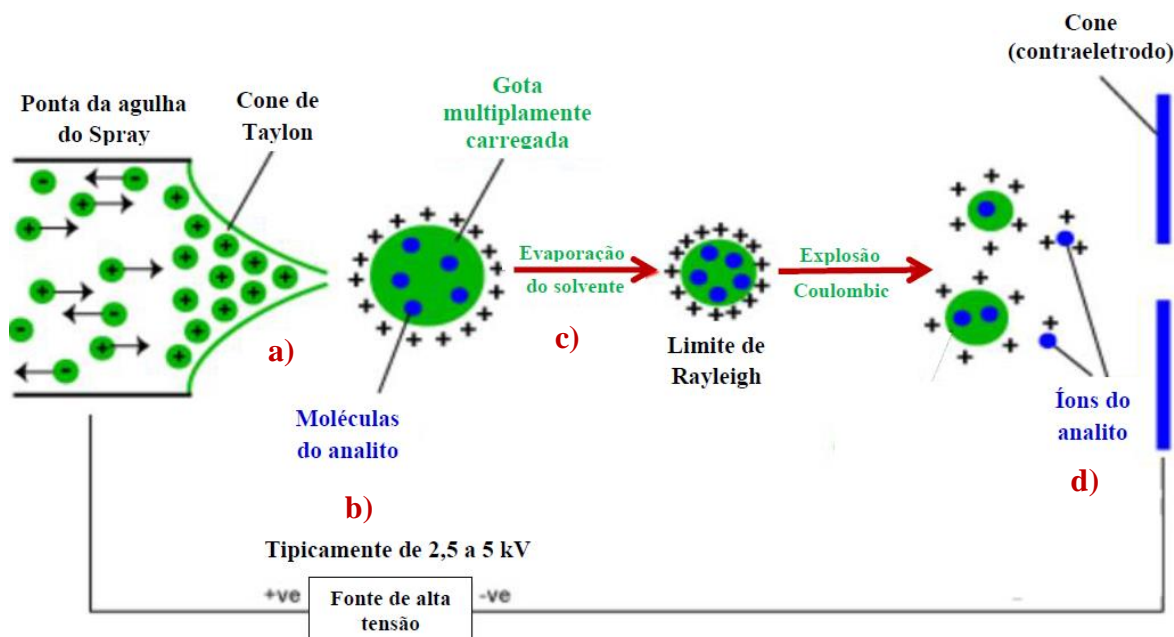
A ionização por electrospray (ESI), atualmente é a técnica preferida de ionização quando se trata de uma amostra que se contrata na sua forma líquida. Por ionizar moléculas diretamente da fase líquida, é compatível com técnicas tradicionais de separação, como, por

exemplo, a cromatografia, que é amplamente utilizada em química analítica. A técnica de espectrometria de massas com ionização por electrospray (ESI-MS) foi desenvolvido por John Fenn, o qual recebeu o prêmio Nobel de química em 2002 para análise de macromoléculas biológicas (FENN *et al.*, 1989; BANERJEE; MAZUMDAR, 2012).

A ESI-MS tornou-se a técnica mais utilizada em análises químicas e bioquímicas, dentre as suas vantagens incluem a alta sensibilidade, consumo reduzido de amostra e análise de biomoléculas de elevado peso molecular (BANERJEE; MAZUMDAR, 2012). Além disso, há outras características essenciais da ESI que faz com que essa apresente um diferencial em relação às demais ferramentas de ionização. A primeira diz a respeito da capacidade com que ela tem de produzir íons multiplamente carregados, reduzindo assim a relação m/z . Logo, isso permite com que compostos de elevada massa molecular possa ser analisado. Outra característica é que as amostras ao serem analisadas devem ser introduzidas em solução, permitindo que seja possível o acoplamento com outras técnicas de separação. Por último, a ESI é uma técnica de ionização suave, provocando pouca ou nenhuma fragmentação dos analitos estudados (VAN DE WATERBEEMD; HECK, 2016).

Na ESI o processo pode ser resumidamente da seguinte maneira (Figura 12). O analito é introduzido na fonte em solução a partir de uma bomba de seringa ou como o fluxo de eluente da cromatografia líquida com uma taxa de fluxo na ordem de $1\mu\text{l}/\text{min}$. O fluxo da solução do analito passa pela agulha de eletropulverização que tem uma alta diferença de potencial em relação ao contra-eletrodo aplicado a ele (tipicamente na faixa de 2,5 a 5 kV). Logo, isso força a pulverização de gotículas carregadas da agulha com uma carga superficial da mesma polaridade para a carga na agulha. Por sua vez, as gotículas são repelidas da agulha em direção ao cone de amostragem da fonte no eletrodo contador (mostrado em azul). Quando as gotículas atravessam o espaço entre a ponta da agulha e o cone, ocorre a evaporação do solvente. Conforme a gota se move em direção ao cone do eletrodo contador (que passa ao analisador), a evaporação do solvente ocorre e, conseqüentemente, a gotícula encolhe até atingir o ponto em que a tensão superficial não consegue mais sustentar a carga, processo é conhecido, como, limite de Rayleigh, e nesse ponto ocorre uma "explosão Coulombic" e a gota é rasgada. Essa fase permite com que sejam produzidas gotículas menores e o processo é repetido. Finalmente, após todo o solvente ter evaporado, a carga é passada para o analito. Portanto, esse é um método muito suave de ionização, uma vez que muito pouca energia residual é retida pelo analito após a ionização. Isso faz com que a ESI-MS seja considerada uma técnica tão importante em estudos biológicos (WILM, 2011; CHALLAMALLA *et al.*, 2012).

Figura 12 – Representação esquemática do processo de formação de íons numa fonte de ESI.



Fonte: Adaptado de Challamalla *et al.* (2012). Etapa (A): Uma solução contendo íons é passada por um capilar onde é aplicada uma alta tensão (positiva ou negativa), ocorrendo então redução/oxidação de alguns íons. Etapa (B): Nesta etapa é formado um spray eletrolítico. Etapa (C): Após o spray eletrolítico, as gotículas carregadas começam a diminuir de tamanho devido à evaporação do solvente, até o momento em que são obtidos os íons isolados (Etapa D).

1.8.6 Tipos de analisadores de massas

Em relação aos analisadores de massas, parte do instrumento na qual realiza a separação dos íons gerados através da sua razão m/z , existe atualmente no mercado diversos modelos diferentes, como, por exemplo, Tempo de voo (TOF), Íon Trap e Orbitrap, que são utilizados na proteômica. Cada um desses apresenta suas respectivas vantagens e desvantagens de acordo a análise pretendida (EL-ANEED; COHEN; BANOUB, 2009). Um dos principais critérios para distinguir os analisadores de massa é justamente o seu poder de resolução em massa. Com base nisso, esse texto focará somente no analisador de massa do tipo Orbitrap, empregado nesse estudo.

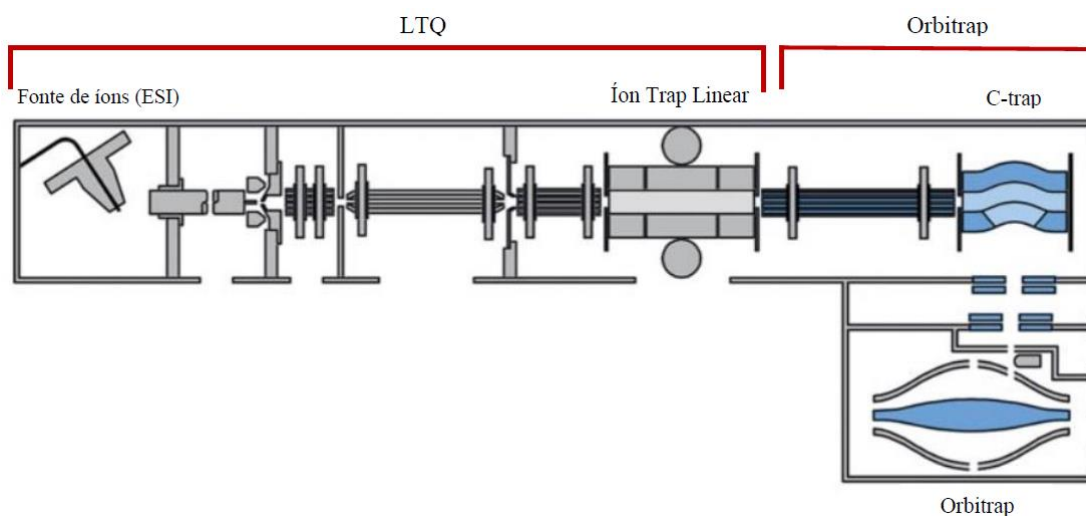
1.8.6.1 Espectrômetro de massas orbitrap

O novo espectrômetro de massas híbrido, o orbitrap, faz parte de um instrumento que foi desenvolvido e construído pela Thermo Masslab Ltd (Manchester, Reino Unido) e

atualmente é comercializado com o nome Linear Trap Quadrupole (LTQ)-Orbitrap[®] (Figura 13). O LTQ-Orbitrap é um sistema de alto desempenho, que combina a aquisição rápida de dados de íon trap linear com alta precisão de massa e análise de massa Orbitrap, permitindo com isso, múltiplos níveis de fragmentação (MS^n) para elucidar a estrutura do analito e propiciar o acoplamento com fontes de ionização contínuas, tais como, electrospray (ESI) (HU *et al.*, 2005; SCIGELOVA *et al.*, 2011).

O princípio do sistema híbrido do orbitrap foi descrito pela primeira vez em 2000 por Makarov. O espectrômetro de massas Orbitrap consiste de dois eletrodos, sendo um localizado na parte central (interno) tipo fuso alongado ao longo de um eixo que é mantido sob alta tensão quando os íons estão sendo retidos em um movimento orbital ao redor do fuso, esse é cercado pelo segundo eletrodo tipo barril externo mantido no potencial da terra. O eletrodo externo é dividido em duas metades para formar eletrodos de captação para detecção de corrente de imagem (Figura 14). Cabe ressaltar ainda que devido às forças eletromagnéticas que os íons sofrem no momento que a distribuição do potencial quadro-logarítmico é gerada pelos eletrodos Orbitrap, os íons oscilam na dimensão axial (representado por z) enquanto orbitam em torno de um eletrodo central. Como os íons oscilam na dimensão axial, eles movimentam-se da região de um eletrodo para o outro com uma frequência dependente de suas relações m/z . Essas oscilações produzidas são convertidas em espectro de massas (m/z) usando a transformada de Fourier (Figura 14) (MAKAROV, 2000; HU *et al.*, 2005; SCIGELOVA *et al.*, 2011).

Figura 13 – Representação esquemática do espectrômetro de massa híbrido LTQ-Orbitrap.

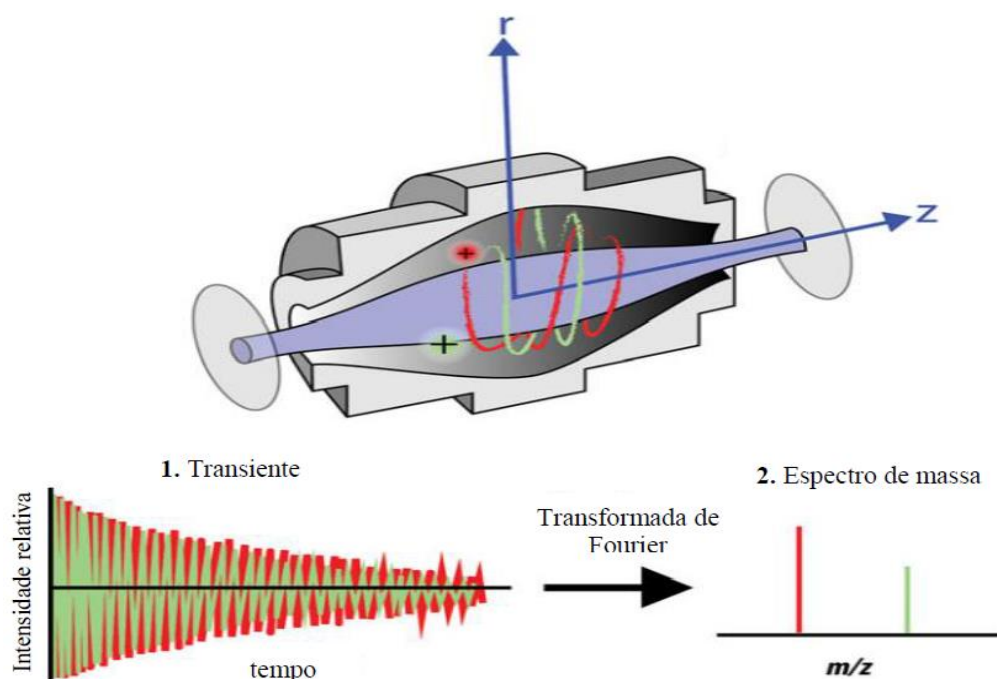


Fonte: Adaptado de Scigelova *et al.* (2011). Nesse sistema híbrido, os íons são transmitidos através do íon trap linear que por sua vez, será repassado para o C-trap e, conseqüentemente, para o orbitrap. O C-trap tem como

finalidade focalizar os feixes de íons em “pacotes” estreitos e transferí-los de maneira rápida para dentro do orbitrap por meio de pulsos.

O desenvolvimento do espectrômetro de massas híbrido, o LTQ-Orbitrap, deve-se ao acoplamento do íon trap linear com o orbitrap devido precisamente ao interfaceamento pelo C-trap. Logo, a combinação desses analisadores de massa permitiu com que o orbitrap realizasse experimentos em sequência com alta resolução e precisão, capaz de alcançar com isso, potência de resolução acima de 100.000 e precisão de massa ≤ 5 ppm (MAKAROV *et al.*, 2006).

Figura 14 – Representação esquemática do analisador de massas Orbitrap.



Fonte: Adaptado de Scigelova *et al.* (2011). No orbitrap, os íons que são capturados em um campo eletrostático oscilam em torno de um eletrodo central tipo fusão alongado enquanto também oscilam na dimensão axial. A oscilação axial é detectada por eletrodos para produzir um transiente, que é transformado no espectro de massa resultante pela transformação de Fourier. O transiente é a frequência de oscilação iônica na dimensão axial e no Orbitrap, todos os espectros de massa (m/z) são detectados simultaneamente.

1.9 Análise proteômica aplicada ao *C. difficile*

Atualmente, análise proteômica descrita como uma tecnologia-chave na era pós-genômica tem fornecido informações essenciais complementares às fornecidas pela genômica. Por meio da análise proteômica, as proteínas identificadas podem ser analisadas rapidamente,

com precisão e elevada sensibilidade utilizando, por exemplo, a técnica de espectrometria de massa (MS) e a eletroforese (MOURA *et al.*, 2013). Com base nisso, um estudo de análise proteômica comparativa entre três isolados clínicos de *C. difficile* (a cepa B-1, a Tra 5/5 e a 027 SM) com propriedades biológicas e virulência variadas foi realizado por Chilton e colaboradores em 2014. Nessa pesquisa, um total de 888 proteínas diferentes foram detectadas e dentre essas, apenas 38% foram compartilhadas entre os isolados, sendo a maioria delas enzimas metabólicas. Entre as três linhagens de cepas, 136 proteínas (15%) foram detectadas apenas na linhagem 027 SM, 162 (18%) somente na linhagem B-1 e 49 (6%) apenas na linhagem Tra 5/5. Notadamente, entre todas as proteínas identificadas incluem, a TcdA, TcdB, proteínas relacionadas a flagelos, proteínas de superfície e dentre outras. Utilizando LC-MS/MS, a TcdA foi detectada em todas as três cepas, enquanto que a TcdB foi detectada nas cepas 027 SM e B-1, mas não na cepa Tra 5/5. Em relação às proteínas flagelares, a flagelina (FliC) foi identificada em todas as três linhagens, enquanto que a proteínas-cap flagelar (FliD) foi identificada na linhagem B-1 e na linhagem 027 SM, mas não foi detectada na linhagem Tra 5/5. Porém, os genes correspondentes à *fliC* e à *fliD* foram encontrados nos genomas de todas as três linhagens por análise blástica. Logo, isso indica que a falta de detecção da FliD na linhagem Tra 5/5 não é devida à falta do gene correspondente. Na mesma pesquisa, ainda identificaram-se 20 proteínas de superfície celular ou de camada S e entre essas, incluíam SlpA, Cwp66 e Cwp84, detectadas nas três cepas analisadas (CHILTON *et al.*, 2014).

No mesmo ano, Chong e colaboradores realizaram uma análise proteômica abrangente de um isolado clínico de NAP1 de *C. difficile* resistente ao metronidazol. Nesse estudo, os autores apresentaram uma caracterização proteômica dos isolados de NAP1/027 e determinaram ainda o efeito do metronidazol no proteoma dos mesmos, focando na expressão de proteínas associado à resistência ao metronidazol. A primeira cepa isolada em 2009 foi denominada de CD26A54_S (sensibilidade reduzida), esta por sua vez na presença de concentrações subletais de metronidazol, tornou-se estavelmente resistente ao mesmo fármaco, esta foi denominada agora de CD26A54_R (resistentes a Met). Na análise proteômica, na ausência de metronidazol foi observado em ambos os isolado (CD26A54_S e CD26A54_R) em comparação com VLOO13 (cepa Met^S NAP1 de *C. difficile* usada como controle) uma expressão diferencial de transporte de elétrons e proteínas redox-ativas. Além disso, uma expressão diferencial de ferredoxina, um intermediário na redução da forma pró-fármaco do metronidazol para a sua forma tóxica, foi observada nas cepas CD26A54_S/R, na ausência de metronidazol. Dois resultados importantes e que chamaram atenção e que podem

estar associada à resistência ao metronidazol em *C. difficile*, diz a respeito aos mecanismos de reparo de DNA e ao metabolismo de ferro. Nessa pesquisa, na ausência de metronidazol os autores identificaram várias proteínas putativas de reparo de DNA com expressão elevada até ± 2 vezes nas cepas CD26A54_S e CD26A54_R em comparação com a cepa controle VLOO13. Entre essas, está a proteína excinuclease ABC, conhecida como sistema de reparo UvrABC e que está envolvida no reparo de lesões de DNA. Em relação à expressão do regulador de absorção férrico (Fur), na ausência de metronidazol foi observado um aumento de $\pm 2,6$ vezes na expressão de Fur nas cepas CD26A54_R e CD26A54_S em relação a VLOO13. Estes resultados sugerem uma possível resposta multifatorial que pode estar justamente associada à resistência ao metronidazol em *C. difficile*, incluindo os possíveis papéis do metabolismo alterado do ferro e/ou reparo do DNA (CHONG *et al.*, 2014).

Em 2010, Jain *et al.* ao realizar uma análise do subproteoma da cepa do *C. difficile* 630 identificou um total de 107 proteínas, entre elas estavam as proteínas a SlpA, cisteína protease (Cwp84), proteína da camada S, proteínas de ligação ao substrato de transportadores ABC, pili, proteínas ribossomais e outras. Nesse mesmo estudo, das 107 proteínas identificadas a maioria estavam envolvidas com síntese proteica (45,8%) (JAIN *et al.*, 2010). Em 2011, o mesmo grupo de pesquisa ao realizar uma análise proteômica com a mesma cepa em resposta ao estresse térmico, foram identificadas 447 proteínas. Dessas, 49 proteínas foram moduladas sob estresse térmico e dentre essas estavam: as proteínas clássicas de choque térmico (chaperonas) GroEL, GroES, DnaK, proteases de Clp e HtpG, que foram reguladas positivamente, além de várias rubreritinas induzidas por estresse e proteínas associadas à modificação de proteínas, tais como prolil isomerases e a prolina racemase (JAIN *et al.*, 2011).

Em 2012, Boetzkes e colaboradores realizaram uma análise de secretomas de três cepas de *C. difficile*, sendo elas os isolados hipervirulentos CD196 e CDR20291 pertencente a NAP1/027 e a de referência CD630. Nessa análise, utilizando a LC-MS foram identificadas 158 proteínas diferentes nos sobrenadantes de *C. difficile*, sendo a maioria delas de origem citoplasmática. Entre essas, 32 proteínas em CDR20291, 36 em CD196 e 26 em CD630 foram identificadas como sendo secretadas por cepas de *C. difficile*. As proteínas identificadas eram principalmente proteínas da camada S, proteínas de ligação ao substrato de transportadores ABC, hidrolases da parede celular, pili e proteínas hipotéticas desconhecidas. Com relação às TcdA e TcdB, os três isolados expressam esses principais fatores de virulência, porém, ambas as toxinas foram apenas fracamente detectada pela LC-MS. Entretanto, as mesmas foram identificadas utilizando a técnica de Western blot. Um dado importante foi a identificação de

cinco proteínas secretadas exclusivamente nos sobrenadantes das cepas hipervirulentas e entre essas, estava a proteína componente de ligação a ADP-ribosiltransferase, que é uma parte da toxina binária (CDT) identificada apenas em cepas hipervirulentas NAP1/027. Logo, esses dados demonstram a grande variação genômica e proteômica nas diferentes cepas de *C. difficile* (BOETZKES *et al.*, 2012).

Numa pesquisa recente realizada por Dresler *et al.* (2017) de análise proteômica comparativa usando a técnica de quantificação por label-free (LFQ), quantificou um total de 662 proteínas de oito ribotipos de *C. difficile* (001, 005, 010, 012, 014, 027, 078, e 176) cultivadas *in vitro*. Entre todas as proteínas quantificadas, vários fatores de virulência foram identificados, como a TcdA, TcdB, CDT, proteínas flagelares e proteínas associadas ao complexo funcional Pro-Pro endopeptidase (PPEP-1), e dentre outras. A análise proteômica comparativa mostrou maior expressão protéica ou expressão única de proteínas ligadas à patogenicidade ou metabolismo de ferro nos ribotipos 027 e 176. Esses resultados são essenciais, pois mostra a relação genética e a importância clínica dessas cepas a nível proteômico. Em outro estudo utilizando a mesma abordagem proteômica a LFQ, Moura e colaboradores identificaram a partir de um filtrado de cultura de *C. difficile* (ATCC 43255) um total de 101 proteínas, entre elas estavam as proteínas TcdA, TcdB e a S-layer. Nessa mesma pesquisa, os autores realizaram ainda uma quantificação precisa de TcdA e TcdB em três lotes de toxinas purificados a partir das cepas de *C. difficile* (VPI 10463-ATCC 43255). Na mesma pesquisa, foi realizada ainda uma ampla cobertura de sequência de aminoácidos das toxinas (91% para TcdA e 95% para TcdB) utilizando a análise de MS (MOURA *et al.*, 2013).

O Quadro 1, resume os dados referentes as pesquisas descritas acima utilizando diferentes cepas de *C. difficile* na área da proteômica, com as seguintes informações: tipos de cepas utilizadas; condições de cultivo; fonte das proteínas; quantidade de proteínas identificadas; e o espectrômetro de massa utilizado (MS).

Quadro 1 – Dados referentes de pesquisas realizadas com diferentes cepas de *C. difficile* utilizando as ferramentas da proteômica.

Cepas utilizadas	Condições de cultivo	Fonte das proteínas	Quantidade de proteínas identificadas	MS	Referência
B-1, Tra 5/5 e 027 SM	Ágar sangue Columbia (CBA) e cultivadas durante 64 h a 37 °C sob condições anaeróbicas	Lisado celular	888	1D Nano-LC/MS/MS acoplado ao Orbitrap e 2D MALDI-TOF MS	CHILTON <i>et al.</i> , 2014
VLOO13, CD26A54_S ou CD26A54_R	Ágar Brucella suplementado (10 µg/ml de vitamina K, 5 µg/ml de hemina e 5% de sangue de carneiro) e cultivadas durante 48 h a 37 °C sob condições anaeróbicas	Lisado celular	1942	Nano-LC/MS/MS acoplado ao Orbitrap	CHONG <i>et al.</i> , 2014
Cepa 630	BHI (Oxoid) e cultivadas a 37 °C sob condições anaeróbicas	Sobrenadante	107	ESI-Q-q-Quadrupole linear íon trap-MS/MS	JAIN <i>et al.</i> , 2010
Cepa 630	BHI (Oxoid) e cultivadas a 37 °C sob condições anaeróbicas	Lisado celular	447	ESI-Q-q-Quadrupole linear íon trap-MS/MS	JAIN <i>et al.</i> , 2011
CD630, CD196 e CDR20291	BHI e cultivadas durante 24 h a 37 °C sob condições anaeróbicas	Sobrenadante	158	Nano-LC/MS/MS acoplado ao Orbitrap	BOETZKES <i>et al.</i> , 2012
RTs 001, RTs 005, RTs 010, RTs 012, RTs 014, RTs 027, RTs 078, e RTs 176	Ágar Anaeróbio Schaedler CM0437 (Oxoid) e cultivadas durante 48 h a 37 °C sob condições anaeróbicas	Sobrenadante	662	LC-MS/MS envolvendo espectrometria de massa direcionada e LFAQ	DRESLER <i>et al.</i> , 2017
ATCC 43255	-	Filtrado comercial de cultura liofilizada	101	Nano-LC/MS/MS acoplado ao Orbitrap	MOURA <i>et al.</i> , 2013

2 JUSTIFICATIVA E RELEVÂNCIA

C. difficile é considerado uma das mais importantes causas de diarreia infecciosa associada à assistência médica em unidades de saúde de cuidados agudos e crônicos, sendo responsável por 15-20% dos casos de diarreia relacionada ao uso de antibióticos (DUBBERKE; WERTHEIMER, 2009). Atualmente, a CDI vem sendo considerada uma importante causa de mortalidade e morbidade, tornando-se uma grande preocupação nos centros de serviços de saúde no mundo todo, além de ocasionar grandes impactos nos custos de saúde (DINGLE *et al.*, 2011). Entre as possíveis causas do aumento da morbimortalidade é o surgimento de cepas consideradas mais virulentas, como, por exemplo, a NAP1/BI/027. Essa cepa produz quantidades maiores de toxinas do que as cepas de referência e são altamente transmissíveis, devido à sua maior capacidade de esporulação (MERRIGAN *et al.*, 2010).

O aumento da incidência e gravidade da CDI levou a um grande ônus econômico nos sistemas de saúde devido aos custos associados ao tratamento e permanências prolongadas de pacientes no hospital. Esse ônus econômico é estimado em US \$ 5,4 bilhões em serviços de saúde e US \$ 725 milhões em ambientes comunitários na América do Norte (DESAI *et al.*, 2016). Além disso, estima que 29.000 mortes anualmente nos EUA sejam atribuídas à CDI (LESSA *et al.*, 2015). Desde 2001, a prevalência e gravidade da CDI vêm aumentando significativamente, o que levou ao aumento do interesse em pesquisas e descoberta de novos fatores de virulência, como, por exemplo, as adesinas (proteínas de superfície – Cwp66 e Cwp84), as proteínas flagelares (flagelina (FliC) e a proteína cap flagelar (FliD), as proteínas de esporulação e as toxinas, que são consideradas como o principal fator de virulência (PECHINE; JANOIR; COLLIGNON, 2005; WRIGHT *et al.*, 2005; RUPNIK; WILCOX; GERDING, 2009).

Com base nisso, o nosso grupo de pesquisa de forma pioneira no Nordeste, em especial no Estado do Ceará, vem pesquisando o *C. difficile* em hospitais cearenses desde o ano de 2013. Uma pesquisa realizada no hospital oncológico, Hospital Haroldo Juaçaba do Instituto do Câncer do Ceará (ICC), Costa e colaboradores detectaram uma alta incidência de CDI (48%) em pacientes com câncer e diarreia. Esse mesmo estudo teve como cepa de *C. difficile* predominantemente isolada e identificada a NAP4, que também é encontrada em outros países, como, EUA e Canadá. Nessa pesquisa foi isolada uma cepa de pulstipo e ribotipo de *C. difficile* que até então era desconhecida pela comunidade científica e passou a

ser denominada até o momento de ICC-45, que juntamente com a cepa testada por nós, NAP1/027 (LIBA5756), coloca-se na classe MLST Clade 2 das cepas consideradas hipervirulentas (COSTA *et al.*, 2016; COSTA *et al.*, 2017).

Em 2016, um estudo realizado por nosso grupo de pesquisa mostrou que em modelos animais utilizando sobrenadantes, a cepa ICC-45 é tão virulenta quanto a NAP1/027 isolada na Costa Rica, mesmo a ICC-45 produzindo menos toxinas que a NAP1 (COSTA *et al.*, 2016). Por meio da proteômica, tem sido possível identificar outros fatores de virulência de *C. difficile* envolvidos em sua patogênese, tornando assim um passo importante para uma melhor compreensão da biologia desse patógeno (BOETZKES *et al.*, 2012).

Tendo em vista a importância epidemiológica da CDI, a análise das proteínas do sobrenadante de culturas de cepas de *C. difficile* por meio de técnicas proteômicas torna-se relevante, já que na literatura existem diferentes genomas de cepas de *C. difficile* já sequenciadas (SEBAIHIA *et al.*, 2006). No entanto, os dados com a caracterização biológica das proteínas identificadas e analisadas, raramente estão disponíveis em conjunto com a análise genômica (CHILTON *et al.*, 2014).

Dessa forma, utilizou-se nesse trabalho técnicas de proteômica para responder questionamentos que a genômica por si só não consegue responder, pois aquela permite investigar com maior profundidade outros potenciais fatores de virulência de *C. difficile*, o que pode ajudar a elucidar a razão molecular para a variação da virulência de *C. difficile* que vem nos últimos anos sofrendo uma rápida evolução e emergindo como um sério patógeno humano. Logo, por meio da análise proteômica as proteínas detectadas podem ser analisadas em conjunto de modo a compreender melhor os mecanismos moleculares da doença associadas ao *C. difficile* e, conseqüentemente, levar à rápida identificação de alvos celulares para fins terapêuticos e assim, poder trazer um retorno positivo para a sociedade com a melhoria da qualidade de vida dos indivíduos com CDI (CHEN *et al.*, 2013).

3. OBJETIVOS

3.1 Objetivo geral

- Identificar as proteínas do sobrenadante de cultura de cepas de *C. difficile* utilizando uma análise exoproteômica comparativa entre sobrenadantes da cepa ICC-45 isolada no Hospital Haroldo Juaçaba e NAP1/027 (LIBA5756) da Costa Rica.

3.2 Objetivos específicos

- Detectar as proteínas de superfície (Cwp66, Cwp84 e SlpA), proteínas flagelares (FliC e FliD), toxinas A e B e outras envolvidas na virulência de *C. difficile* presentes nos sobrenadante das cepas analisadas por meio da exoproteômica.
- Analisar as proteínas identificadas quanto à participação nos processos biológicos, função molecular e componente celular utilizando o programa Blast2Go.
- Identificar a localização celular das proteínas presentes nos sobrenadantes das cepas em análise utilizando o programa Psortb.
- Identificar e comparar as interações físicas e funcionais entre as proteínas identificadas na cepa ICC-45 e NAP1/027 (LIBA5656), utilizando como ferramenta o banco de dados STRING.
- Identificar as interações das proteínas exclusivas da ICC-45 com as compartilhadas entre as cepas por meio do banco de dados STRING.

4 MATERIAIS E MÉTODOS

4.1 Cepas bacterianas

Para o estudo da exoproteômica, foram utilizadas duas cepas de *C. difficile*, a ICC-45 (cepa toxigênica, novo ribotipo/genótipo) isolada de paciente internado no Hospital Haroldo Juaçaba, Fortaleza-CE, em 2014 por nosso grupo de pesquisa (COSTA *et al.*, 2016) e a NAP1/027 (LIBA5756 – cepa hipervirulenta isolada na Costa Rica) (QUESADA-GÓMEZ *et al.*, 2010). Com relação a cepa NAP1 (LIBA5756), essa foi cedida pelo Prof. Dr. Carlos Quesada do Laboratório de Pesquisa em Bacteriologia Anaeróbia da Universidade da Costa Rica. As cepas selecionadas para esse estudo fazem parte da Coleção de Culturas do Laboratório de Bacteriologia, do Departamento de Biomedicina, da Universidade Federal do Ceará (UFC). Os dados gerais sobre as cepas utilizadas nesta pesquisa estão listadas no Quadro 2.

Quadro 2 – Cepas de *C. difficile* selecionadas para o estudo da exoproteoma.

Cepas	Paciente	Condição Clínica	Perfil Toxigênico	Pulsotipo	Ribotipo	Período de Isolamento
ICC-45	Feminino, 34 anos	CA de mama	<i>tcdA+</i> , <i>tcdB+</i> , <i>cdtB+</i> , sem deleções <i>tcdC</i>	Novo	Novo	Agosto de 2014
NAP1 (LIBA5756)	-	-	<i>tcdA+</i> , <i>tcdB+</i> , <i>cdtB+</i> , com deleções <i>tcdC</i>	NAP1	027	2009

CA – câncer; *tcdA+* – cepa positiva para gene da toxina A (TcdA); *tcdB+* - cepa positiva para gene da toxina B (TcdB); *cdtB+* – cepa positiva para gene da toxina binária; e *tcdC* - gene regulador negativo que controla a produção das toxinas (TcdA e TcdB).

4.2 Cultivo das cepas de *C. difficile*

As cepas de *C. difficile* foram reativadas a partir de um estoque mantido em criotubos a -20°C contendo uma solução de BHI-PRAS (do inglês *Brain Heart Infusion – pre-reduced anaerobically and sterilized*) suplementado com 0,5% de extrato de levedura, 0,05% de cisteína e 1% de menadione, acrescida de 20% de glicerol (COSTA *et al.*, 2016). Desse

estoque, uma alíquota foi esgotada em placa de Petri contendo o meio Ágar Brucella suplementado com 5% de sangue de carneiro e menadione (vitamina K, 5 mg/mL) (Sigma®) e incubadas sob condições anaeróbicas (jarra de anaerobiose - 90% N₂, 10% CO₂, 5% H₂), por aproximadamente 72 horas a 37 °C para obter o crescimento das colônias de *C. difficile*.

4.3 Obtenção dos sobrenadantes

Após o crescimento das cepas de *C. difficile* (ICC-45 e a NAP1/027 (LIBA5756)), foi realizado o repique 2 a 4 colônias de cada cepa em 40 mL de caldo BHIS. As cepas foram incubadas pelo período de 24 horas a 37 °C em jarra de anaerobiose (90% N₂, 10% CO₂, 5% H₂). Após o tempo de incubação, centrifugado 2x (4000 rpm, 8 min, 4 °C), os sobrenadantes resultantes foram passados através de filtros de poro de tamanho de 0,2 µm, para eliminar as células vegetativas. Os sobrenadantes foram congelados a -80 °C até o momento do uso (COSTA *et al.*, 2016).

Como controle, utilizamos um caldo de BHIS não inoculado e submetido às mesmas condições do cultivo. Além disso, foram realizadas três réplicas biológicas desse experimento para obter os extratos proteicos para a análise exoproteômica.

4.4 Precipitação das proteínas obtidas dos sobrenadantes de *C. difficile*

Após a obtenção dos sobrenadantes das cepas ICC-45 e NAP1/027 em BHIS, foi realizada uma precipitação com clorofórmio PA e metanol PA, na proporção de 1 volume do sobrenadante: 1 volume de clorofórmio: 4 volumes de metanol. Após agitação, por cerca de 2 minutos com o auxílio do vórtex e centrifugação a 2.000 x g, duas fases de separação foram obtidas. A fase superior foi descartada e mais 3 volumes de metanol foram acrescentados. Em seguida, o sobrenadante foi submetido a novas etapas de agitação e centrifugação a 2000 x g. O sobrenadante foi descartado e o sedimento liofilizado. Após secagem, o sobrenadante seco foi armazenado a - 80 °C (WESSEL; FLÜGGE 1984).

O sobrenadante seco obtido foi enviado ao Laboratório de Biologia de Anaeróbios, Departamento de Microbiologia Médica, do Instituto de Microbiologia Paulo de Góes (IMPPG) da Universidade Federal do Rio de Janeiro (UFRJ), da Prof^a Dr^a Eliane de Oliveira Ferreira. No laboratório, o material foi hidratado com 500 µL de tampão de bicarbonato de amônio (Fluka) 25 mM (pH 8,0) com cloreto de cálcio a 1 mM (AMBIC), compatível com a espectrometria de massas, e as proteínas dosadas no Qubit® 2.0 Fluorometer (Invitrogen,

California, USA), conforme as instruções do fabricante. Em seguida, o material foi preparado para a proteômica *in solution*.

4.5 Tripsinização *in solution* das proteínas

Dos sobrenadantes já em tampão AMBIC foram retirados 100 μL e a eles misturados 300 μL de AMBIC (volume final: 400 μL). Esse volume foi adicionado nas colunas do tipo Spinfilter de 30 kDa (Millipore-Amicon ULTRA 0,5 mL) para digestão das proteínas obtidas. Para tal, foram acrescentados 400 μL de água HPLC (alta pureza) às colunas, sendo estas centrifugadas a 8000 $\times g$ por 5 min a 8 °C. Após a centrifugação, foram adicionados 400 μL de Ditioneitol (DTT; Sigma-Aldrich) 10 mM, sendo incubado por 30 min a 37 °C. Após a centrifugação e descarte do sobrenadante, foram adicionados 350 μL de iodoacetamida 50 mM, sendo o material incubado no escuro a temperatura ambiente por 30 min. Uma centrifugação a 8000 $\times g$ foi realizada por 10 min e a coluna lavada com 400 μL de AMBIC. Os tubos de coleta foram trocados por um novo e a cada coluna foram acrescentados 400 μL de tampão AMBIC contendo tripsina (*Sequence Grade modified Trypsin*; Promega, Wisconsin, USA) (5 μg /amostra). Os tubos ficaram incubados em banho-maria *overnight* a 37 °C, e após uma centrifugação a 8000 $\times g$, os peptídeos coletados foram secos e armazenadas a -20 °C. Uma parte desse produto seco já tripsinizado foi colocado em uma placa de 96 poços cônica (Orbitrap) e enviado para os Estados Unidos (*National Center for Environmental Health, Centers for Disease Control and Prevention – CDC, Clinical Chemistry Branch Division of Laboratory Science*, do Dr. Hércules Moura e Dr. John Barr) para as análises de espectrometria de massas. O material da placa foi resuspenso em ácido fórmico 0,1% e agitado com o auxílio de um vórtex, antes da injeção no aparelho. Todas as análises foram realizadas em um MS Nano-LC ESI-MS/MS (Waters Corporation, Milford, EUA) acoplado a LTQ Orbitrap (Thermo Fischer) (MOURA *et al.*, 2013).

4.6 Identificação das proteínas em MS nLC MS/MS acoplado a LTQ Orbitrap

Os produtos da digestão, obtidos *in solution*, foram submetidos à cromatografia líquida acoplada à MS utilizando-se o equipamento nano-LC MS/MS acoplado a LTQ Orbitrap Velos. Foram aplicados 3 μL de cada produto digerido de proteínas em uma microcoluna (300 μm x 5 nm) empacotada com partículas C18 PepMap™ (Waters), com uma vazão de 5.000 $\mu\text{L}/\text{min}$. Os peptídeos foram eluídos em coluna analítica C18 PepMap™ (75 μm id x 15cm) (Waters) usando o seguinte gradiente de fase móvel da *High Performance*

Liquid Chromatography (HPLC): solvente A (água grau HPLC com ácido fórmico a 0,1%) e solvente B (ácido fórmico a 0,1% em acetonitrila). Após 10 min de 5% do solvente B, o gradiente foi aumentado para 40% do solvente B em 10 min e 95% do solvente B em 2 min. Após 15 min de 95% do solvente B, o sistema foi levado a 95% do solvente A e 5% do solvente B em 1 min. A coluna foi então reequilibrada com 95% do solvente A e 5% do solvente B, por 15 min. O MS foi operado em modo positivo com uma voltagem de 35 kV para a aquisição de dados baseados na leitura dos valores de 400 a 1.400 m/z, em uma resolução de massa nominal de 60.000 para a aquisição de íons. Para a análise no MS o aparelho foi programado para selecionar os 15 íons mais intensos com carga duas ou mais cargas. O tempo de aquisição do MS/MS foi de 120 min. Para calibração do método quantitativo e controle interno, uma solução de peptídeos digeridos de álcool desidrogenase de levedura (Water) foi injetados na coluna na concentração de 100 fMol e com picos de 30.030 Da (Adaptado de MOURA *et al.*, 2013).

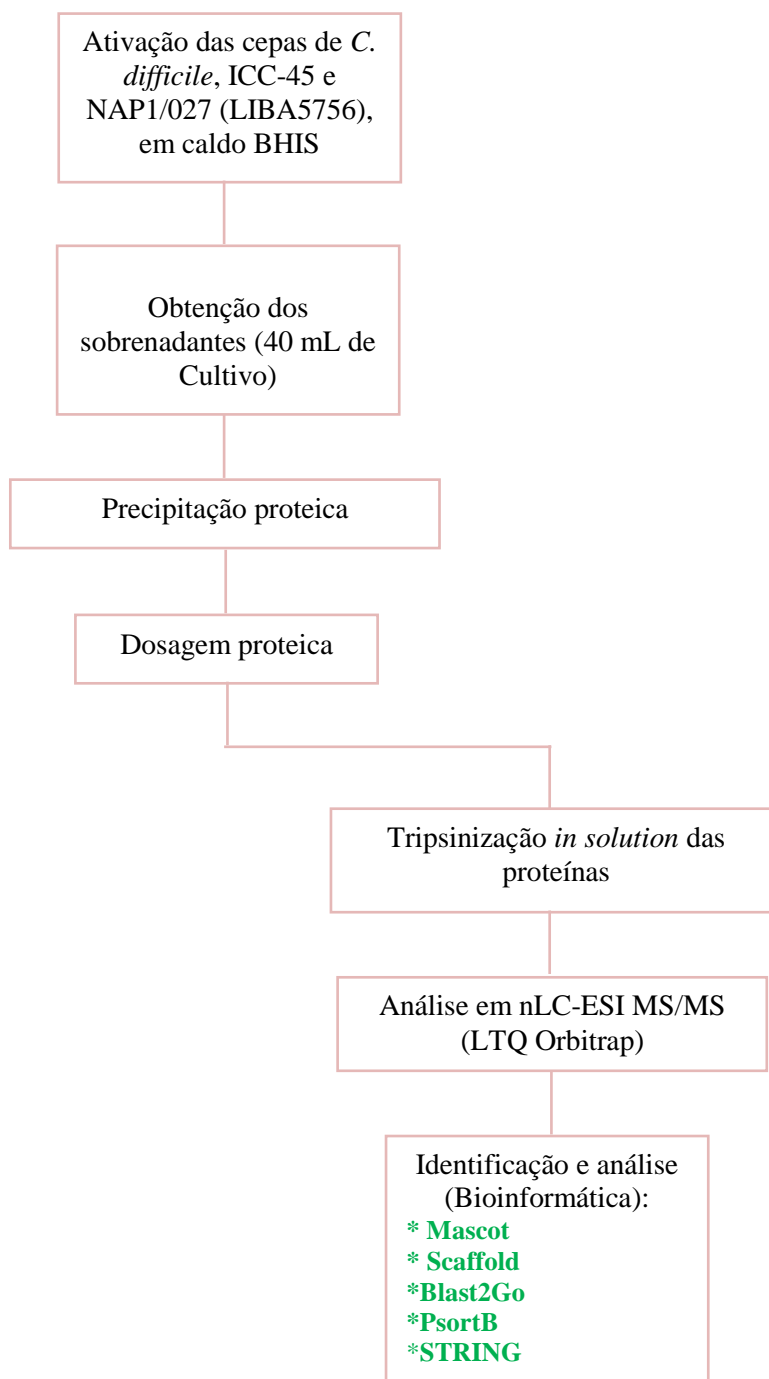
4.7 Análises de bioinformática

Com relação aos peptídeos submetidos à análise no espectrômetro de massas Nano-LC ESI-MS/MS acoplado a LTQ Orbitrap, geraram espectros de massa em *tandem* que foram analisados pelo programa *Mascot Distiller* (Matrix Science, London, UK; versão 2.5.1). A identificação das proteínas foi realizada através da busca por sequências homólogas pelo programa Mascot. Esse foi programado para pesquisar e reconhecer proteínas com base no banco de dados do NCBI modificado e criado com o termo “Cdiff-nr-630-base-Sep2015-Vs4”, no qual a sequência de aminoácidos da ADH 30030 (controle) e das cepas ATCC 43255 e R20291 foram depositadas, sendo a tripsina utilizada como o agente da digestão. Uma tolerância de massa péptica de 50 p.p.m. para o íon parental e 0,60 Da para os íons fragmentos foi especificado. A desamidação de asparagina e glutamina, oxidação da metionina e carboximetil de cisteína foram também especificados no Mascot como modificações variáveis.

Os arquivos gerados pelo Mascot foram devidamente convertidos para serem analisados no Scaffold (V 4.8.6.0, *Proteome Software Inc.*, Portland, OR), programa utilizado para identificar, validar e organizar proteínas de amostras biológicas. As identificações de proteínas foram aceitas se pudessem ser estabelecidas com probabilidade maior que 99,0% e contivessem pelo menos 2 peptídeos identificados. No trabalho atual, o FDR (taxas de falsas descobertas) foi estabelecido em 1%; assim, 99% das proteínas identificadas nessa pesquisa

devem estar corretas. Posteriormente, foram utilizados os programas Blast2Go (<https://www.blast2go.com/b2ghome>) utilizado para prever as funções biológicas das proteínas identificadas. O PsortB v. 3.0 (<http://www.psort.org/psortb/>) foi utilizado para prever a localização celular das proteínas identificadas. O banco de dados STRING v. 10.5 (<https://string-db.org/>), foi utilizado para prever as interações físicas e funcionais entre as proteínas identificadas na cepa ICC-45 e NAP1/027 (LIBA5756) e também as interações das proteínas exclusivas da ICC-45 com as 26 compartilhadas entre as cepas. Na Figura 15, ilustra o fluxograma que representa todas as etapas do estudo.

Figura 15 – Fluxograma das etapas utilizada na exoproteômica para identificação e análise dos extratos proteicos obtidos das culturas de cepas de *C. difficile*.



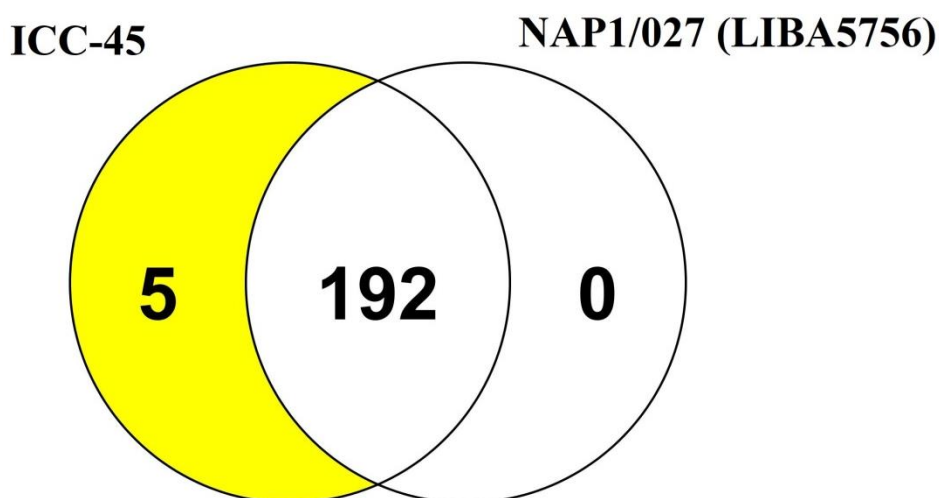
5 RESULTADOS

5.1. Análise das proteínas obtidas dos sobrenadantes das culturas das cepas de *C. difficile*

Na análise exoproteômica utilizando a metodologia *in solution* ou *gel-free*, os arquivos dos peptídeos gerados após o processamento pelo espectrômetro de massas (Orbitrap), os dados brutos (*raw data*), foram inicialmente processados com o software Mascot para identificação e quantificação de proteínas no qual as sequências de aminoácidos dos peptídeos trípticos obtidos foram comparadas com sequências de proteínas depositadas em um banco de dados de *C. difficile*. Em seguida, as proteínas identificadas foram analisadas no Scaffold com o intuito de validar as proteínas reconhecidas pelo Mascot. A partir das análises obtidas na *in solution*, foram identificadas um total de 197 proteínas das cepas de *C. difficile* analisadas nesse estudo. Todas as proteínas identificadas, bem como o número de acesso e a probabilidade de identificação das proteínas das cepas em estudo, estão listadas no Anexo 1.

Dentre as 197 proteínas identificadas, foi possível destacar proteínas exclusivas e compartilhadas entre as cepas de *C. difficile*, conforme o diagrama de Venn construído pelo próprio Scaffold (Figura 16). Como pode ser observado no diagrama de Venn, 5 proteínas identificadas pertencem exclusivamente a cepa ICC-45. Por outro lado, a cepa ICC-45 e a NAP1/027 (LIBA5756) compartilharam 192 proteínas. Já a cepa NAP1/027 (LIBA5756) não apresentou nenhuma proteína exclusiva sua.

Figura 16 – Diagrama de Venn das proteínas identificadas e validadas pelo Scaffold.



Legenda: Diagrama de Venn representando as 5 proteínas exclusivas da cepa ICC-45 e as 192 proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756).

Entre as proteínas identificadas exclusivas da cepa ICC-45, estão àquelas envolvidas com a via glicolítica (fosfogliceromutase), biossíntese de porfirinas e ligante de metal (ácido delta-aminolevulínico desidratase), ligante de DNA (regulador transcricional) e metabolismo de carboidratos (fosfato de ribulose), conforme o Tabela 1.

Em relação às 192 proteínas compartilhadas entre as cepas de *C. difficile*, foram selecionadas 26 e categorizadas em seis grupos das quais estão as proteínas envolvidas com: a patogenicidade (toxinas, proteínas de superfície celular, proteínas flagelares, proteínas de parede celular, hidrolases e proteases); resistência a antimicrobianos (beta-lactamases e piruvato-ferredoxina); choque térmico e estresse (chaperonas); óxido nítrico (flavorubredoxina redutase do óxido nítrico); metabolismo e atividade catalítica (hidrolase de trealose-6-fosfato e cisteína dessulfurase); e outras atividades (fator de alongamento de transcrição). Cabe ressaltar ainda que, ao analisar as proteínas secretadas pelas cepas ICC-45 e NAP1/027 (LIBA5756), foi possível realizar uma análise comparativa da quantidade relativa de proteínas (quantidade de espectro total) que cada cepa produz e isso só foi permitido porque o espectrômetro de massas (Orbitrap) possibilita essa análise (Tabela 2).

Com relação a essa abordagem, podemos observar na Tabela 2, no que se refere especificamente ao grupo das proteínas envolvidas com a patogenicidade da infecção causada por *C. difficile* (CDI), que a cepa ICC-45 apresenta quantidades maiores de espectro total de proteína da camada S (S-layer), toxina A (TcdA) e B (TcdB), proteína de superfície celular (Cwp19), proteína de parede celular (Cwp22), hidrolase, proteína de ligação à parede celular (Cwp28), flagelina (FliC) e cisteína protease (Cwp84) quando comparada a quantidade de espectro produzida pela cepa NAP1/027 (LIBA5756).

No grupo das proteínas que estão relacionadas com resistência a antimicrobianos, pode ser observado que a cepa ICC-45 produz bem mais proteínas da família das beta-lactamases, oxidoreductase de piruvato-ferredoxina e nitroreductase do que a NAP1/027 (LIBA5756). Já em relação ao grupo que engloba proteínas que estão envolvidas com estresse oxidativo (cisteína dessulfurase e rubreritina) e proteínas de choque térmico (chaperonas), pode-se perceber que a cepa ICC-45 produz em torno de quase 3x mais a quantidade de cisteína dessulfurase em relação à NAP/027 (LIBA5756) e também secreta bem mais chaperonas e, por sua vez, essa cepa secreta mais rubreritina, proteína associada ao estresse oxidativo, do que ICC-45.

No que tange ao grupo das proteínas que estão associadas aos efeitos de proteção do óxido nítrico, como pode ser averiguado na Tabela 2, a cepa ICC-45 produz quase 3x mais a flavorubredoxina redutase do óxido nítrico quando comparada a NAP1/027 (LIBA5756) e da mesma forma, a ICC-45 também produz quantidades de flavoproteína putativa de óxido

nítrico redutase a mais do que a NAP1/027 (LIBA5657). Agora quanto ao grupo das proteínas que estão envolvidas com metabolismo e atividade catalítica, podemos observar que a ICC-45 produz mais hidrolase de trealose-6-fosfato, fosfotransferase proteica de fosfoenolpiruvato e alanina racemase quando comparada a NAP1/027 (LIBA5756). Enfim, o grupo das proteínas que apresentam outras atividades pode-se observar que a ICC-45 secreta mais D-prolina redutase, fator de alongamento de transcrição e pilin do que a NAP1/027 (LIBA5756), porém, essa secreta bem mais proteína de choque frio (CspB) quando comparada a ICC-45. Logo, com a finalidade de identificar quais funções, processos biológicos, componentes celulares e a localização celular que todas essas proteínas estão envolvidas incluindo a proteínas compartilhadas entre as cepas e as exclusivas da ICC-45 foram realizadas análises em bancos de dados para responder esses questionamentos.

Tabela 1 – Proteínas secretadas exclusivamente pela cepa ICC-45 de *C. difficile*.

Proteínas Identificadas	Número de Acesso	Nome do Gene^a	Massa Molecular^a
Phosphoglyceromutase	gi 255308206	gpmI	56 kDa
hypothetical protein	gi 126699562	CD630_1953	44 kDa
delta-aminolevulinic acid dehydratase	gi 255308448	hemB	36 kDa
Rrf2 family transcriptional regulator	gi 126698875	iscR	16 kDa
ribulose-phosphate 3-epimerase	gi 126700194	rpe1	24 kDa

^aAnotações funcionais foram retiradas do banco de dados de proteína (www.uniprot.org).

Tabela 2 – Quantidade relativa de proteínas secretadas pelas cepas ICC-45 e NAP1/027 (LIBA5756) de *C. difficile*.

Proteínas Identificadas	Número de Acesso	Nome do Gene^a	Massa Molecular^a	Espectro Total^b (ICC-45)	Espectro Total^b (NAP1)
Proteínas envolvidas com a patogenicidade					
cell surface protein (S-layer precursor protein)	gi 255307831	slpA	76 kDa	5359	4001
toxin A	gi 255305655	tcdA	308 kDa	3399	2202
toxin B	gi 126698238	tcdB	269 kDa	2633	1501
cell surface protein	gi 1001999562	cwp19	77 kDa	45	37
cell wall protein	gi 400927472	cwp22	71 kDa	18	4
cell-wall hydrolase	gi 126700384	CD630_2768	25 kDa	30	11
cell wall binding repeat 2 family protein	gi 531118578	cwp28	51 kDa	13	2
flagellin subunit	gi 260685615	fliC	34 kDa	38	0
cell surface protein (putative cell surface-associated cysteine protease)	gi 260210525	cwp84	87 kDa	62	14
Proteínas relacionadas com resistência a antimicrobianos					
beta-lactamase family protein	gi 531115368	QEW_1874	112 kDa	49	14
pyruvate-ferredoxin oxidoreductase	gi 126700296	Pfo	128 kDa	895	650
Nitroreductase	gi 126700191	CD630_2572	24 kDa	12	8
Proteínas relacionadas ao estresse oxidativo e choque térmico					
Rubrerythrin	gi 126699128	CD630_15240	20 kDa	515	573
cysteine desulfurase	gi 126698876	iscS2	43 kDa	115	52
chaperone DnaK	gi 126700078	dnaK	66 kDa	527	335
60 kDa chaperonin	gi 255305190	groL	57 kDa	312	167
heat shock protein 90	gi 126697845	htpG	75 kDa	19	3

Proteínas associadas aos efeitos do óxido nítrico

anaerobic nitric oxide reductase flavorubredoxin	gi 126698752	norV	44 kDa	171	58
putative nitric oxide reductase flavoprotein	gi 255306635	CDR20291_1521	94 kDa	87	84

Proteínas envolvidas com metabolismo e atividade catalítica

trehalose-6-phosphate hydrolase	gi 126700708	treA	66 kDa	266	156
phosphoenolpyruvate-protein phosphotransferase Chain D, Alanine Racemase	gi 260687988	ptsI	63 kDa	114	75
	gi 645985739	alr2	43 kDa	43	9

Proteínas envolvidas em outras atividades

D-proline reductase PrdA	gi 126700863	prdA	67 kDa	39	22
transcription elongation factor GreA	gi 126701179	greA	18 kDa	21	1
cold shock protein CspB	gi 126698954	cspB	7 kDa	663	670
Pilin	gi 255308544	CD630_3513	18 kDa	32	26

^aAnotações funcionais foram retiradas do banco de dados Swiss Proteina (www.uniprot.org)

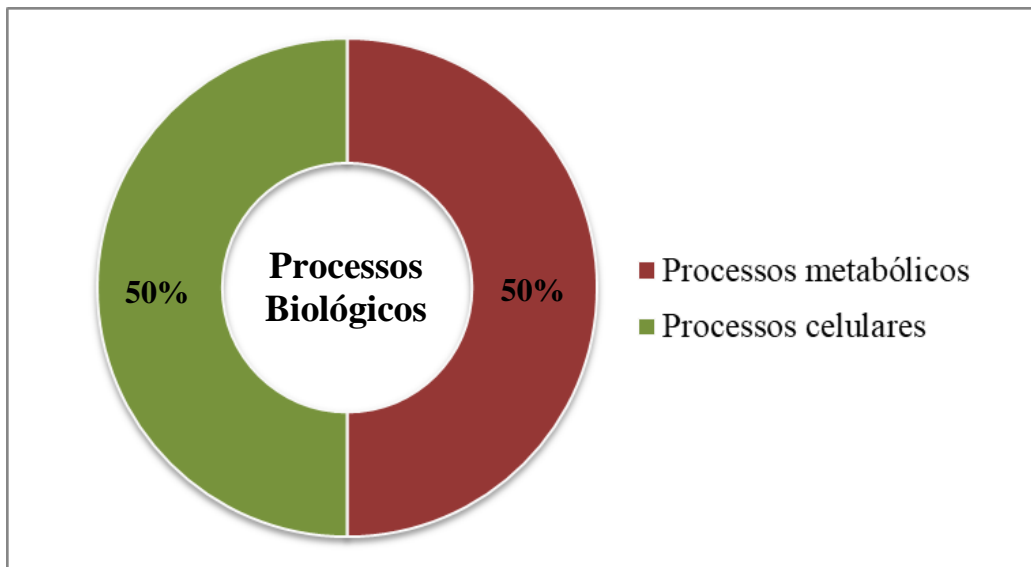
^bConcentração baseada na quantidade de espectro total dado pelo programa Scaffold.

5.2 Análise das proteínas de *C. difficile* por meio de programas de bioinformática

5.2.1 Funções biológicas: programa Blast2Go

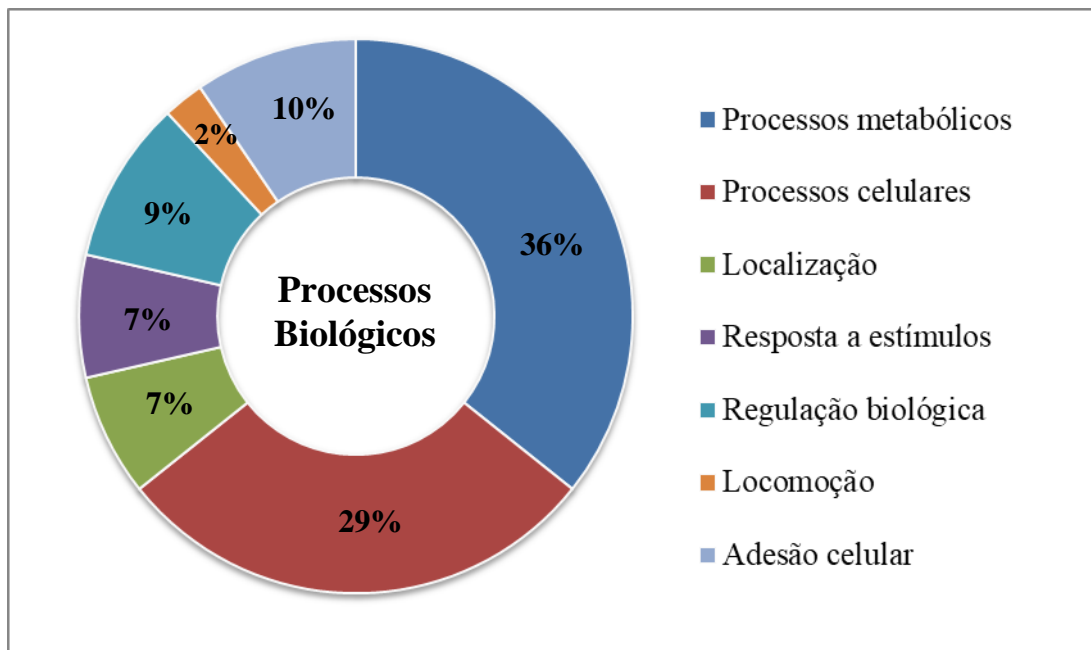
O programa Blast2Go consiste numa ferramenta de bioinformática que tem como objetivo realizar uma análise funcional com base nas anotações do *gene ontology* (GO *annotations*), ou seja, analisar as proteínas quanto sua participação em processos biológicos, função molecular e componente celular. Nesse caso, no que se refere aos processos biológicos das proteínas exclusivas da cepa ICC-45, podemos observar que há um padrão semelhante no que diz respeito às proteínas envolvidas nesse processo, ou seja, 50% das proteínas estão relacionadas com processos metabólicos e 50% com processos celulares (Figura 17). Já com relação às proteínas que são compartilhadas entre ICC-45 e NAP1/027 (LIBA5756), pode-se perceber que essas estão envolvidas em vários processos biológicos, sendo que a maioria delas estão ligadas a processos metabólicos (36%) e as demais estão subdivididas em processos celulares (29%), localização (7%), resposta a estímulos (7%), regulação biológica (9%), locomoção (2%) e adesão celular (10%) (Figura 18).

Figura 17 – Gráfico das proteínas exclusivas da cepa ICC-45 envolvidas nos processos biológicos.



Legenda: Proteínas identificadas da cepa ICC-45 de *C. difficile* com base nas anotações do *gene ontology* (GO *annotations*). O gráfico representa o percentual de proteínas envolvidas nos processos biológicos.

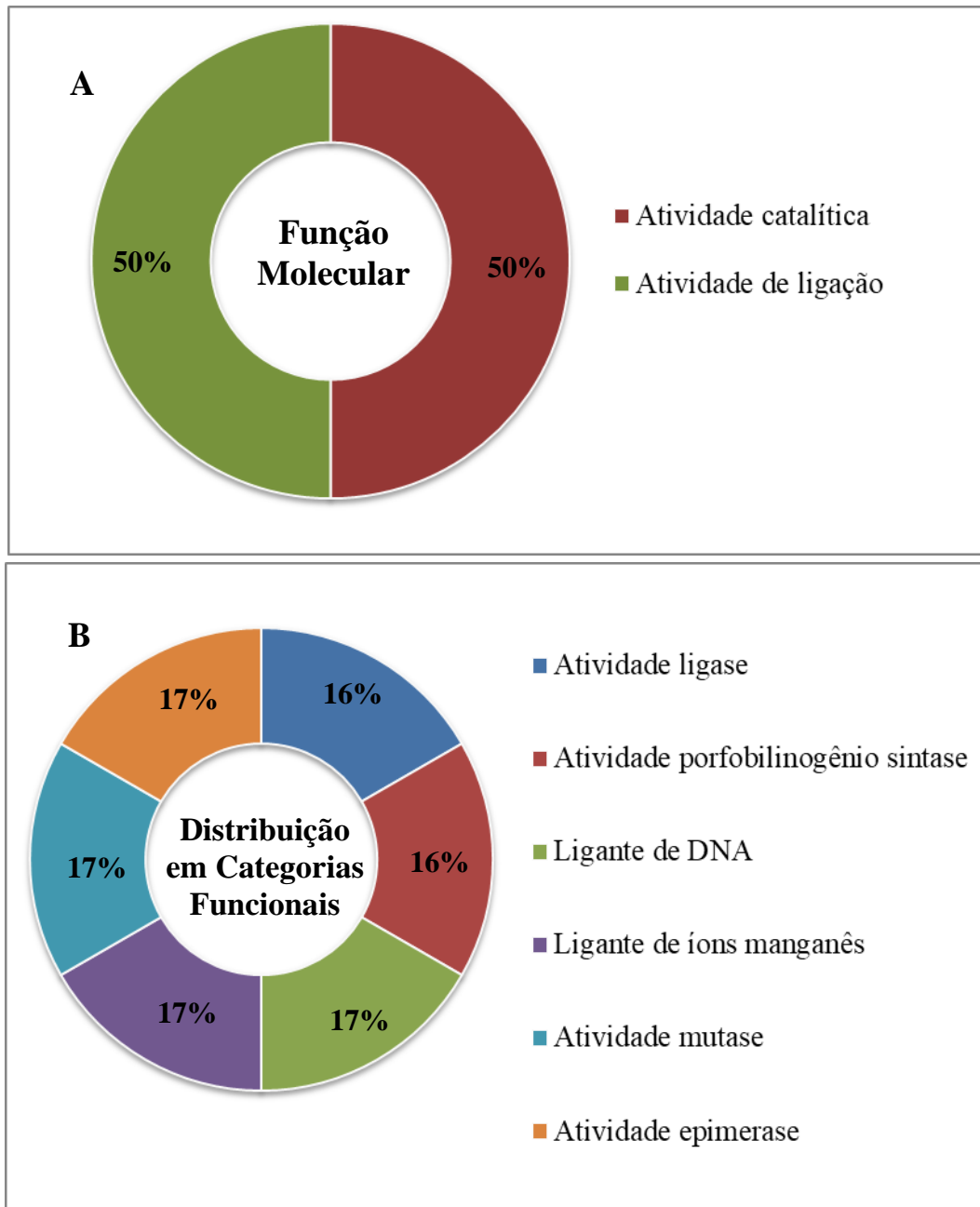
Figura 18 – Gráfico das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756) envolvidas nos processos biológicos.



Legenda: Percentual de proteínas identificadas e compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756) e que estão envolvidas com processos biológicos, sendo a maioria relacionada a processos metabólicos. Todas essas proteínas de *C. difficile* foram identificadas com base nas anotações do *gene ontology* (GO annotations).

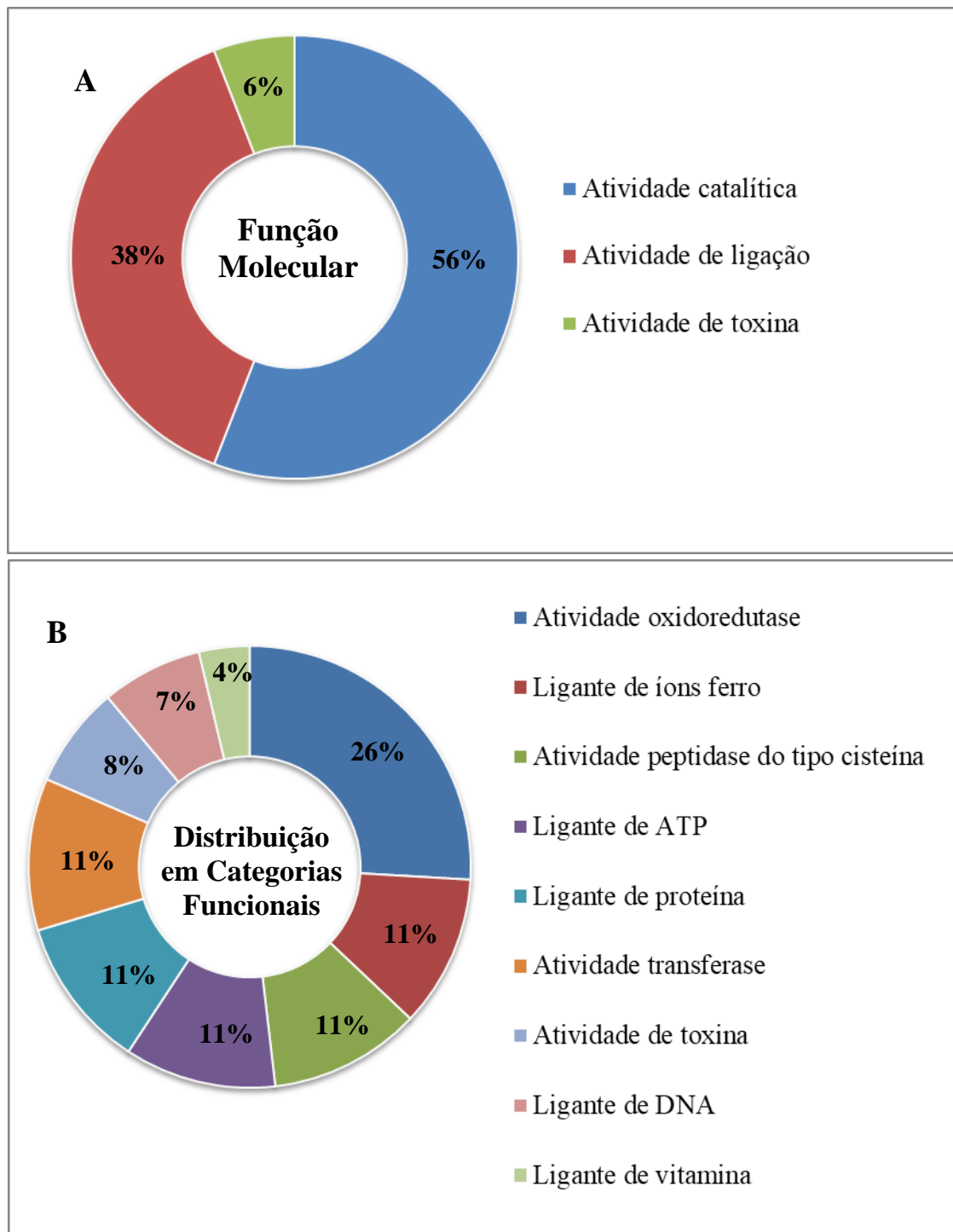
Com relação às funções moleculares, no que tange as 5 proteínas exclusivas da cepa ICC-45, percebe-se que houve uma similaridade de 50% para atividade catalítica e 50% para atividade de ligação (Figura 19A). Dentre essas porcentagens, as proteínas exclusivas da ICC-45 foram distribuídas em outras categorias funcionais desempenhando agora outras atividades, como, ligase (16%), porfobilinogênio sintase (16%), ligante de DNA (17%), ligante de íons manganês (17%), mutase (17%) e epimerase (17%) (Figura 19B). Já no que se refere às funções moleculares das 26 proteínas compartilhadas entre a cepa ICC-45 e NAP1/027 (LIBA5756), podemos perceber que mais da metade exercem atividade catalítica (56%) e as demais, apresentam atividade de ligação (38%) e atividade de toxina (6%) (Figura 20A). Agora, no que diz respeito à distribuição dessas porcentagens das proteínas compartilhadas entre a cepa ICC-45 e NAP1/027 (LIBA5756), as proteínas foram distribuídas em outras categorias funcionais apresentando atividades das quais podemos observar que a grande parte delas estão envolvidas com oxidoreductase (26%) e as demais, estão relacionadas com funções de ligante de íons ferro (11%), atividade peptidase do tipo cisteína (11%), ligante de ATP (11%), ligante de proteína (11%), atividade transferase (11%), atividade de toxina (8%), ligante de DNA (7%) e ligante de vitamina (4%) (Figura 20B).

Figura 19 – Função molecular e distribuição das proteínas exclusivas da ICC-45.



Legenda: Função molecular e distribuição as quais pertencem as proteínas identificadas exclusivamente no sobrenadante da cepa ICC-45 com base nas anotações do *gene ontology* (*GO annotations*). O gráfico A representa o percentual de proteínas envolvidas com a função molecular, no qual apresentam um padrão de similaridade entre funções. Já no gráfico B representa o percentual de proteínas no que diz respeito sua distribuição em categorias funcionais.

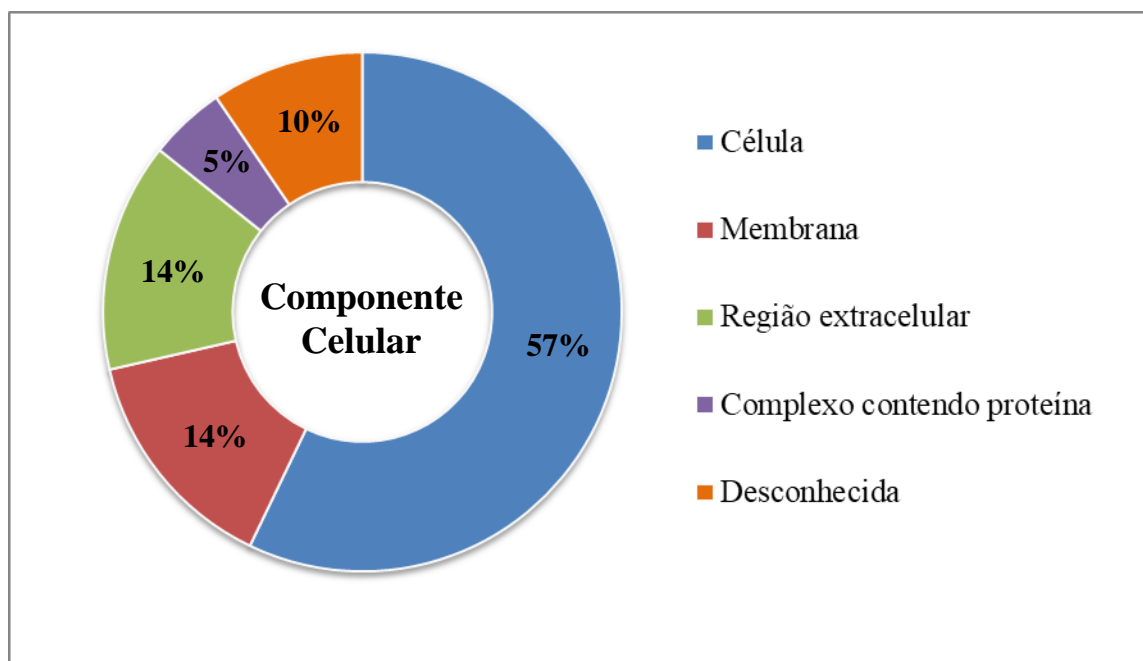
Figura 20 – Função molecular e distribuição das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756).



Legenda: Função molecular e distribuição das proteínas compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756) identificadas baseadas nas anotações do *gene ontology* (*GO annotations*). O gráfico A representa o percentual de proteínas envolvidas com a função molecular, sendo a maioria relacionada com atividade catalítica. Já no gráfico B representa o percentual de proteínas no que diz respeito sua distribuição em categorias funcionais, no qual a grande parte estão envolvidas com atividade oxidoreductase.

Os resultados pertencentes às proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756) que foram categorizadas em componentes celulares (Figura 21), pode-se perceber que mais da metade pertencem às proteínas como componente celular (57%) e as demais estão subdivididas em: 14% como membranares; 14% como região extracelular; 5% como complexo contendo proteína; e 10% como componentes desconhecidos. No que tange aos componentes celulares das 5 proteínas exclusivas da cepa ICC-45, todas pertencem como componente celular (dados não mostrados).

Figura 21 – Componente celular das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756).



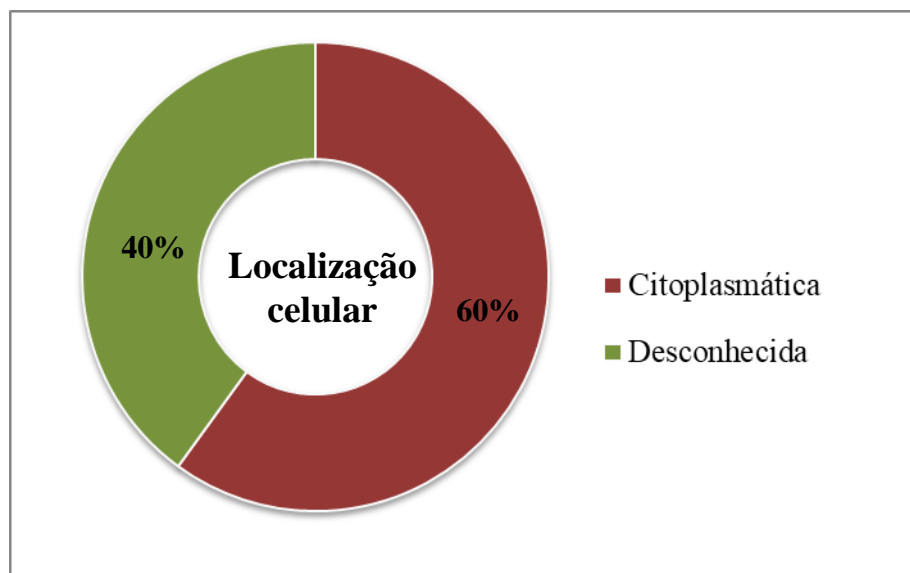
Legenda: Gráfico dos componentes celulares das proteínas compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756) identificadas com base nas anotações do *gene ontology* (*GO annotations*). O gráfico representa o percentual de proteínas que fazem parte dos componentes da célula, sendo que a maioria pertence às proteínas como componente celular.

5.2.2 Localização celular das proteínas: programa Psortb

De acordo com a localização celular realizada por meio do programa Psortb no que se refere às proteínas que pertencem exclusivamente a cepa ICC-45, a maior parte das proteínas identificadas são de origem citoplasmática (60%), enquanto que as outras são de origem desconhecida (40%) (Figura 22). Já com relação à localização celular das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756), mais da metade são também de origem citoplasmática (58%), seguida daquelas relacionadas à parede celular (19%),

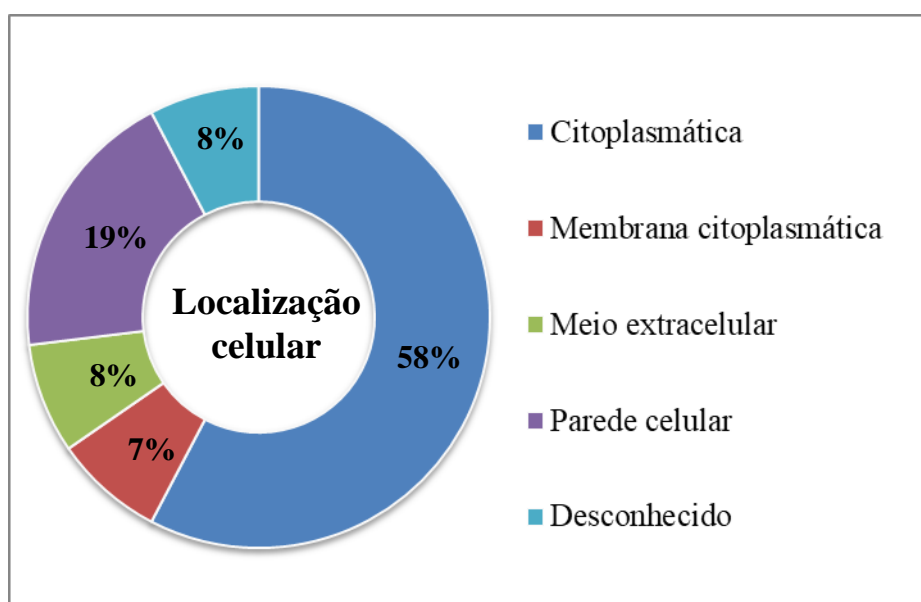
membrana citoplasmática (7%), meio extracelular (8%) e de origem desconhecida (8%) (Figura 23).

Figura 22 – Localização celular das proteínas exclusivas da cepa ICC-45.



Legenda: Localização celular das proteínas identificadas exclusivamente no sobrenadante da cepa ICC-45. O gráfico representa o percentual de proteínas identificadas conforme sua localização celular, sendo que 60% das proteínas exclusivas da ICC-45 são de origem citoplasmática. O programa PsortB foi utilizado para a construir os gráficos.

Figura 23 – Localização celular das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756).



Legenda: Localização celular das proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756). O gráfico representa o percentual de proteínas compartilhadas entre as cepas e identificadas conforme sua localização

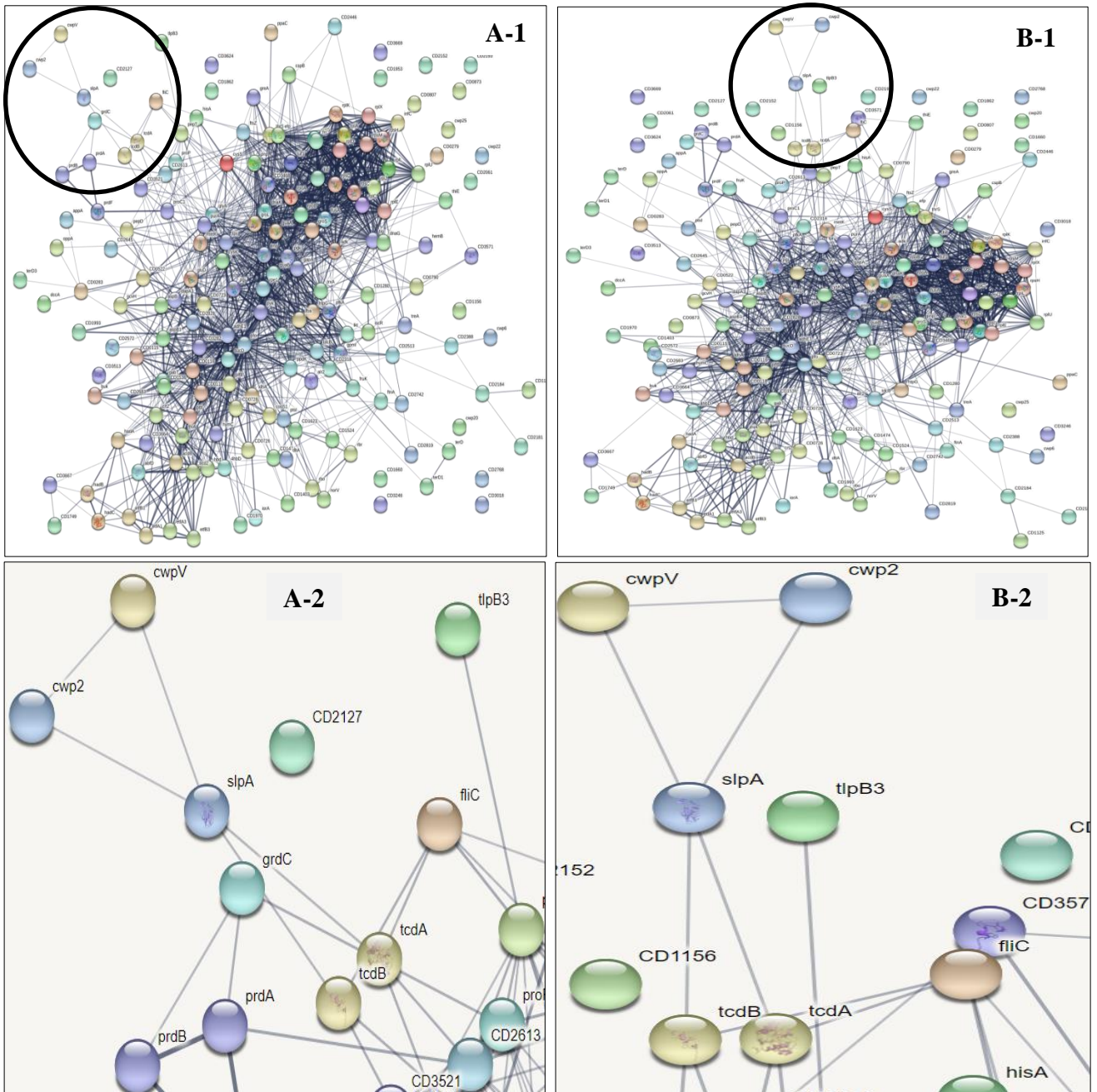
celular, sendo a maioria delas de origem citoplasmática (58%). O programa PsortB foi utilizado para a construir os gráficos.

5.2.3 Rede de associação das proteínas funcionais dos sobrenadantes das cepas de *C. difficile* analisada por STRING

O banco de dados STRING consiste numa ferramenta para predizer as associações de proteína-proteína, incluindo as interações diretas (físicas) e as indiretas (funcionais), desde que ambas sejam específicas e biologicamente significativas. Com base nesse banco de dados, podemos perceber que a cepa ICC-45 e a NAP1/027 (LIBA5756) apresentam um padrão semelhante no que diz respeito às interações entre as proteínas, ou seja, quando a mesma área específica de interações proteína-proteína é selecionada e ampliada nas duas cepas, conforme os círculos pretos, percebe-se que a proteína A de camada superficial (SlpA) interage de forma semelhante com a TcdA e a TcdB em ambas as cepas (Figura 24).

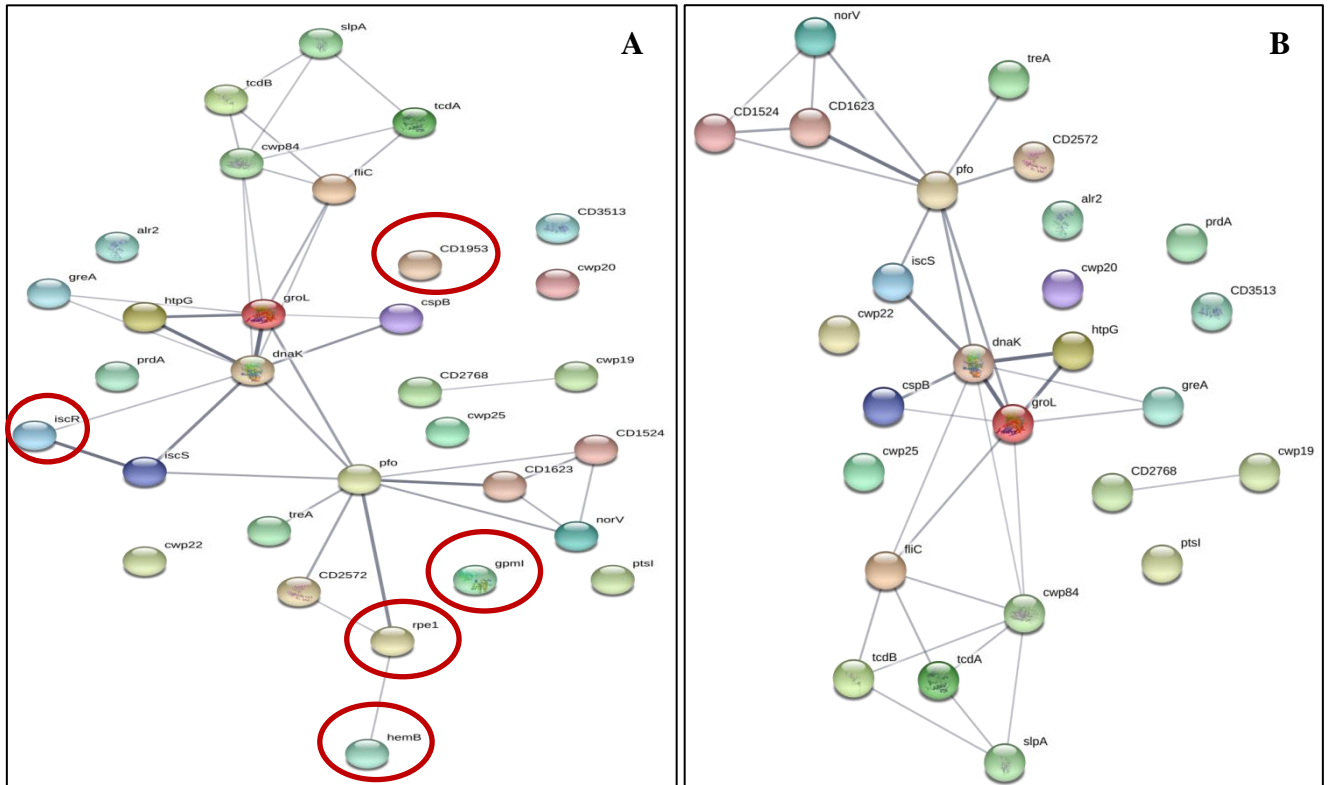
Quando se observa as interações proteína-proteína exclusivas da ICC-45 com as 26 compartilhadas entre ICC-45 e NAP1/027 (LIBA5756), percebe-se que dentre as 5 proteínas exclusivas da cepa ICC-45 destacadas em círculos vermelhos, 3 dessas interagem com proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756) e duas não apresentam nenhuma interação com as demais. No que diz respeito às 3 interações, a proteína IscR (regulador transcricional) interage diretamente com a proteína chaperona DnaK, essa interage com a chaperona GroL, que conseqüentemente, irá interagir com proteínas de patogenicidade (TcdA, TcdB, FliC, Cwp84 e SlpA), por sua vez, a IscR também interage com cisteína desulfurase (IscS). O ácido desamino-aminolevulínico desidratase (HemB) interage com a ribulose-fosfato 3-epimerase (Rpe1). Por outro lado, a Rpe1 apresenta-se interagindo com outras duas proteínas, sendo elas a nitroreductase (CD2572) e a oxidoreductase de piruvato-ferredoxina (Pfo) (Figura 25A). No que tange a Figura 25B, as 26 proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756), podemos perceber que a forma como as proteínas interagem umas com as outras é diferente, ou seja, as associações entre as proteínas das duas cepas estão relacionadas, mas apresentam um padrão diferente.

Figura 24 – Semelhanças entre as interações das proteínas da cepa ICC-45 (A-1) e NAP1/027 (LIBA5756) (B-1).



Legenda: A ampliação de uma área específica das interações entre as proteínas da cepa ICC-45 (A-2), mostra que a SlpA apresenta interações semelhantes as proteínas da NAP1/027 (LIBA5756) (B-2). A SlpA interage diretamente com a TcdA e a TcdB em ambas as cepas.

Figura 25 – Diferenças das interações físicas e funcionais entre as proteínas selecionadas da cepa ICC-45 (incluindo as exclusivas dessa cepa) com as compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756).



Legenda: Interações entre as 31 proteínas das cepas ICC-45, isto é, 5 proteínas exclusivas da ICC-45 (círculos vermelhos) mais 26 compartilhadas e selecionadas das duas cepas (ICC-45 e NAP1/027) quando comparada com as interações das 26 proteínas compartilhadas entre as cepas. Dentre as cinco proteínas da ICC-45, apenas 3 (*iscR*, *hemB* e *rpe1*) apresentam interações com as demais, inclusive essas influenciam nas interações com as proteínas de patogenicidade (*TcdA*, *TcdB*, *FliC*, *Cwp84* e *SlpA*), enquanto que duas proteínas (*gpml* e *CD1953*) não interagem umas com as outras, conforme destacado em círculos vermelhos (A). O interatoma das 26 proteínas compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756), é diferente quando comparada com as interações das 31 das cepas ICC-45 (B).

6 DISCUSSÃO

Em nossa pesquisa, utilizamos os sobrenadantes das cepas ICC-45 e NAP1/027 (LIBA5756) do *C. difficile* e identificamos 192 proteínas na cepa NAP1/027 (LIBA5756) e 197 proteínas na ICC-45. De forma semelhante, a literatura refere identificação que varia de 107 a 158 proteínas utilizando sobrenadante ou filtrado de cultura, sugerindo que o resultado da nossa pesquisa é compatível com os resultados de outros autores (JAIN *et al.*, 2010; BOETZKES *et al.*, 2012; MOURA *et al.*, 2013). No entanto, Dresler *et al.* (2017) teve um rendimento maior, obtendo 662 proteínas identificadas. Essa maior quantidade de proteínas identificadas por esse autor, pode estar relacionada ao tempo de cultivo da cepa (48 h), enquanto que em nosso trabalho foi de 24 h e esse período de cultivo foi semelhante ao de Boetzkes *et al.* (2012) que identificou 158 proteínas.

Das 197 proteínas identificadas na cepa ICC-45, cinco proteínas foram detectadas exclusivamente nessa cepa. Por outro lado, a cepa ICC-45 e a NAP1/027 (LIBA5756) compartilharam 192 proteínas. Dentre as cinco proteínas que foram detectadas apenas na cepa ICC-45, quatro estão envolvidas com processos importantes, como, por exemplo, via glicolítica (fosfogliceromutase), biossíntese de porfirinas e ligante de metal (ácido delta-aminolevulínico desidratase), ligante de DNA (regulador transcricional) e metabolismo de carboidratos (fosfato de ribulose).

A fosfogliceromutase (PGM) desempenha um papel essencial nas vias tanto da glicólise quanto da gliconeogênese, onde essa enzima interconverte o ácido 3-fosfoglicérico (3PGA) e o ácido 2-fosfoglicérico (2PGA). A glicólise ou via glicolítica é uma via metabólica central que exerce importantes funções, como, a geração de precursores de reações anabólicas e a conservação da energia necessária para abastecer todos os outros processos celulares (COMMICHAU *et al.*, 2009). Existem dois tipos diferentes de PGM, os PGMs dependentes de cofator (dPGMs) que requerem o ácido 2,3-difosfoglicérico (DPG) para catálise (essa molécula participa como um doador de fosfato) e os PGMs independentes de cofator (iPGMs) que não exigem o DPG. Os dPGMs são os únicos PGM presentes em vertebrados (incluindo humanos), certos invertebrados e algumas bactérias, principalmente Gram-negativas. Já os iPGMs são os únicos ou predominantes PGM presentes em muitas outras bactérias, em particular Gram-positivas, incluindo membros das espécies do gênero *Bacillus* e *Clostridium* formadores de esporos (JEDRZEJAS *et al.*, 2000). Um estudo realizado por Nukui *et al.* (2007) mostrou que a fosfogliceromutase independente de cofator (iPGMs) é uma enzima crucial para o crescimento de células e esporos em espécies de *Bacillus*. Os autores mostraram

ainda nessa mesma pesquisa, que dois íons Mn^{2+} na catálise da iPGM levam a uma extrema sensibilidade do pH dessa reação, o que é crucial tanto para o acúmulo de 3PGA na esporulação quanto para o catabolismo de 3PGA durante a germinação de esporos (NUKUI *et al.*, 2007).

Corroborando com o achado da presença de uma enzima essencial para germinação de esporos em membros das espécies do gênero *Bacillus* e *Clostridium*, Costa e colaboradores (2017 - dados não publicados), mostraram que a cepa ICC-45 apresenta maior expressão gênica relativa de *spo0A*, fator de transcrição ativo nos estágios iniciais da esporulação em *C. difficile*, em comparação com NAP1/027 (LIBA5756). Entretanto, na proteômica não foi detectada a proteína Spo0A em nenhuma das cepas estudadas, o que pode ser explicado pelo fato de essa proteína ser expressa na fase estacionária de crescimento e, conseqüentemente, o período de 24 horas de incubação das cepas não tenha sido suficiente para sua expressão (PETTIT *et al.*, 2014).

O ácido delta-aminolevulínico desidratase (ALAD), é uma enzima contendo sulfidril citosólico que catalisa a condensação de duas moléculas de ácido aminolevulínico (ALA) para formar o porfobilinogênio, que é o precursor da heme, citocromos, cobalaminas e vitamina B₁₂. Um estudo realizado por Bardag-Gorce e French (2011), mostrou que quando proteossomas purificados foram incubados com ALAD, verificou-se que ALAD causou um aumento na atividade do proteassoma (complexos de proteases chave responsáveis pela degradação de proteínas dependentes de ubiquitina e que desempenham papéis cruciais em processos celulares). Os autores mostraram ainda que a ALAD não é um substrato para os proteossomas, mas sim um complexo regulador de oito subunidades de 240 kDa do proteassoma 20S e que ALAD esteja envolvido na modulação da atividade do proteassomo (BARDAG-GORCE; FRENCH, 2012). Esses resultados podem sugerir a hipótese de que a ALAD detectada apenas na ICC-45 pode desempenhar um papel significativo na regulação da atividade do proteassomo. Porém, estudos são necessários para esclarecer essa hipótese.

Os membros da família Rrf2 (regulador transcricional) são proteínas relativamente pequenas (12-18 kDa) formada por uma estrutura do tipo *helix-turn-helix* (HTH) na sua porção N-terminal. A família Rrf2 é representada por quatro diferentes reguladores (CymR, NsrR, RirA e IscR) (SANTOS; PEREIRA; MACEDO-RIBEIRO, 2015). Estudos mostram que o regulador de cisteína global (CymR) controla a utilização de fonte de enxofre no hospedeiro e desempenha papel na formação de biofilme, além de estar também envolvido com resistência ao estresse e resposta oxidativa em *Staphylococcus aureus* (SHEPARD *et al.*, 2011; JI *et al.*, 2012). O NsrR é um repressor de transcrição sensível ao óxido nítrico (NO) no

qual controla o metabolismo do NO em várias bactérias, como, por exemplo, *Escherichia coli* e *Bacillus subtilis*. Em *Bacillus subtilis*, o NsrR reprime a transcrição dos genes da nitrito redutase (*nasDEF*) que estão sob controle positivo do sistema de transdução de sinal de dois componentes ResD-ResE (KOMMINENI *et al.*, 2010). O RirA (regulador de ferro rizobial A) controla a regulação de genes responsivos na manutenção da homeostase do ferro em rhizobia e também na regulação do ferro, resposta ao estresse oxidativo e virulência em *Agrobacterium tumefaciens* (NGOK-NGAM *et al.*, 2009). Dentre esses, temos ainda o regulador da biossíntese de ferro-enxofre (IscR) que foi detectado apenas na cepa ICC-45. O IscR abriga um clusters [2Fe-2S] no qual permite a interferência entre as vias biossintéticas ISC (cluster de ferro-enxofre) e de SUF (assimilação de enxofre) para coordenar a utilização de ferro e cisteína para a montagem do cluster Fe/S (SANTOS *et al.*, 2014). Em *E. coli* e várias outras bactérias, os genes envolvidos nesse processo são regulados em resposta à disponibilidade de [Fe-S] através da proteína [Fe-S] IscR e, conseqüentemente, são induzidos durante a privação de ferro e estresse oxidativo. Na mesma bactéria, a ISC é considerada o sistema de manutenção responsável pela maturação de uma grande variedade de proteínas Fe/S, enquanto o sistema SUF é acionado principalmente sob condições de estresse, como estresse oxidativo ou falta de ferro (SANTOS *et al.*, 2014; SANTOS; PEREIRA; MACEDO-RIBEIRO, 2015).

A Ribulose-fosfato 3-epimerase (Rpe) também conhecida como pentose-5-fosfato 3-epimerase (Ppe) identificada na cepa ICC-45, é uma enzima que converte a D-ribulose 5-fosfato (Ru5P) em D-xilulose 5-fosfato no ciclo de fosfato pentose redutor de Calvin (AKANA *et al.*, 2006). Um estudo realizado por Sobota e Imlay (2011), mostrou que a Rpe em *E. coli* é rapidamente danificada por peróxido de hidrogênio (H₂O₂), mas pode ser protegido por manganês (Mn). Além disso, os autores mostraram ainda que em estudos *in vitro* com a Rpe, essa emprega um átomo de ferro ferroso como um cofator exposto a solvente e que o H₂O₂ oxida rapidamente esse metal em uma reação de Fenton e, por sua vez, o ferro oxidado é liberado imediatamente, causando uma perda de atividade. Esses resultados sugerem que quando o Mn é fornecido nas células, a enzima é protegida e a função da via das pentoses é restaurada (SOBOTA; IMLAY, 2011).

Além dessas já mencionadas na cepa ICC-45, foi identificada uma proteína hipotética conservada de peso molecular de 44 kDa e, que após comparação em bancos de dados genômico, foi verificada que ela apresenta 100% de identidade com a coenzima F₄₂₀: γ -glutamil ligase (FbiB) em *C. difficile*. De acordo com a literatura, a coenzima F₄₂₀ é um nome dado a um grupo de cofatores ativo redox e dentre esse grupo está a (FbiB), no qual é encontrada principalmente em archaea e actinobactérias (incluindo micobactérias). Durante a

biossíntese do cofator F₄₂₀, a enzima F₄₂₀- γ -glutamil-ligase (FbiB) catalisa os passos finais da via de biossíntese do F₄₂₀ por adições sucessivas de resíduos de L-glutamato a F₄₂₀-0, produzindo uma cauda de poli- γ -glutamato (SELENGUT; HAFT, 2010; BASHIRI *et al.*, 2016).

Em *Mycobacterium tuberculosis*, a (FbiB) é uma proteína que consiste em dois domínios: um domínio N-terminal da superfamília da F₄₂₀-ligase e um domínio C-terminal com similaridade de sequência com as proteínas da superfamília nitro-mononucleótido de flavina (FMN) redutase. Já a enzima responsável por essa reação em archaea (CofE) compreende um único domínio e esse apresenta homologia com o domínio N-terminal da FbiB e produz o F₄₂₀-2 (atividade γ -glutamil ligase) como a espécie principal (BASHIRI *et al.*, 2016). O domínio C-terminal da FbiB possui 27% de identidade de sequência com uma nitroredutase de *Mycobacterium smegmatis* (NfnB), no qual tem como função conferir resistência às drogas antituberculares benzotiazinona (BTZ) e dinitrobenzamida (DNB) (MANINA *et al.*, 2010; REHAN *et al.*, 2011). Estudos têm sugerido que a coenzima F₄₂₀ protege o *Mycobacterium tuberculosis*, contra o estresse oxidativo e nitrosativo durante a patogênese (DARWIN *et al.*, 2003; PURWANTINI; MUKHOPADHYAY, 2009). É possível que essa proteína esteja também ligada a proteção do *C. difficile* ao óxido nítrico produzido no intestino do hospedeiro.

Das 192 proteínas compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756), foram selecionadas 26 e categorizadas em seis grupos das quais estão as proteínas envolvidas com: a patogenicidade; resistência a antimicrobianos; estresse oxidativo e choque térmico; resistência ao óxido nítrico; metabolismo e atividade catalítica; e outras atividades.

No grupo das proteínas envolvidas com a patogenicidade da infecção causada por *C. difficile* (CDI), a cepa ICC-45 apresenta quantidades maiores de espectro total de proteína da camada S (S-layer), toxina A (TcdA) e B (TcdB), proteína de superfície celular (Cwp19), proteína de parede celular (Cwp22), hidrolase, proteína de ligação à parede celular (Cwp28), flagelina (FliC) e cisteína protease (Cwp84) quando comparada a quantidade de espectro (proteína) produzida pela cepa NAP1/027 (LIBA5756). Esses fatores de virulência aumentados na ICC-45 quando comparada a NAP1/027 (LIBA5756) pode explicar o motivo dessa cepa ser tão virulenta quanto a NAP1 em modelo animal, causando um intenso processo inflamatório e lesão tecidual semelhante a NAP1 (COSTA *et al.*, 2016).

Em 2015, Sun e Hirota mostraram que a TcdA e TcdB juntamente com outros fatores de virulência do *C. difficile*, como, as proteínas da camada superficial (SLPs) e a FliC, contribuem para as respostas inflamatórias observadas na patogênese da infecção por causada

por *C. difficile* (CDI) (SUN; HIROTA, 2015). Além disso, Ryan *et al.* (2011) mostrou que as SLPs interagem com o receptor toll-like 4 (TLR4) induzindo uma resposta imune pró-inflamatória inata do hospedeiro (RYAN *et al.*, 2011). A FliC, um componente estrutural dos flagelos do *C. difficile*, pode desencadear respostas inflamatórias ativando o TLR5 nas células epiteliais intestinais e essa ativação desempenha um papel crucial nas infecções bacterianas do intestino grosso (YOSHINO *et al.*, 2013). Em 2013, Chen e colaboradores, mostraram que o tipo de condições de crescimento do *C. difficile*, como, por exemplo, mudanças de nutrientes ou choque osmótico, condições essas que quando presentes no meio de crescimento podem suprimir a produção de proteínas importantes que estão envolvidas diretamente com a virulência de cepas do *C. difficile*, como, as toxinas A (TcdA) e B (TcdB) (CHEN *et al.*, 2013). As mudanças nas condições de crescimento do *C. difficile*, seja a retirada de algum nutriente essencial ou submetê-lo algum tipo de estresse, podem fazer com que cepas de *C. difficile* deixem de secretar para o meio extracelular proteínas importantes e, essas podem contribuir para uma maior ou menor virulência do *C. difficile* e influenciar na sua patogenicidade.

O nosso grupo de pesquisa mostrou que a cepa ICC-45 produz uma TcdB variante que induz um efeito citopático distinto de acordo com o seu fenótipo de virulência aumentado (COSTA *et al.*, 2016). Nesse mesmo estudo, foi mostrado que a cepa ICC-45, conforme a análise de tipagem de sequência de Multilocus (MLST), pertence ao grupo de cepas filogeneticamente relacionado com cepas NAP1/027 e outras cepas pertencentes ao MLST do Clade 2 das cepas consideradas hipervirulentas. Isso pode explicar a semelhança entre as proteínas da ICC-45 e da NAP1/027 (LIBA5756), já que as duas cepas estão localizadas dentro do mesmo Clade 2, no qual estão as cepas hipervirulentas.

Sabendo-se que a adesão celular é uma etapa essencial para que ocorra a colonização no hospedeiro, várias proteínas de superfície celular, como, por exemplo, a Cwp66, proteínas da camada superficial (camada S), proteínas flagelares e proteínas de parede celular, já foram descritas e coordenam a adesão da célula bacteriana à parede do intestino (WRIGHT *et al.*, 2005). Em 2015, Pantaléon *et al.* descreveu a cisteína protease (Cwp84) do *C. difficile* como moduladora da formação de biofilme e as propriedades da superfície celular (PANTALÉON *et al.*, 2015). Nesse sentido, o nosso grupo de pesquisa vem desenvolvendo trabalhos com as cepas ICC-45 e NAP1/027 (LIBA5756) para avaliar a capacidade de formar biofilme e quais possíveis proteínas estão envolvidas com a formação desses biofilmes.

No grupo das proteínas que estão relacionadas com resistência a antimicrobianos, a cepa ICC-45 produz bem mais proteínas da família das beta-lactamases, oxidoreductase de

piruvato-ferredoxina e nitroredutase do que a NAP1/027 (LIBA5756). A resistência aos antibióticos desempenha um papel importante na condução das atuais mudanças epidemiológicas da CDI e, conseqüentemente, no surgimento de novos ribotipos do *C. difficile*. Certos antibióticos, como, por exemplo, cefalosporinas (CFs), clindamicina (CLI) e, mais recentemente, fluoroquinolonas (FQs), são conhecidos por apresentar um risco maior de CDI do que outros (SLIMINGS; RILEY, 2014). Um estudo sobre avanços recentes de resistência a antibióticos em *C. difficile* realizado por Spigaglia em 2016, foi mostrado que a resistência à CLI e CFs é muito comum em isolados clínicos de *C. difficile* (55% e 51%, respectivamente), como também a resistência à eritromicina (ERY) e FQs (47%). Nesse mesmo estudo, foi mostrando ainda que maioria das cepas testadas quanto à suscetibilidade a FCs de segunda geração [cefotetam (CTT) e cefoxitina (FOX)] foram resistentes (79%), enquanto a resistência aos FCs de terceira geração [ceftriaxona (CRO) e cefotaxima (CTX)] está presente em menor número de isolados (38%). Da mesma forma, a resistência à ciprofloxacina (CIP), uma FQ de segunda geração, é muito comum em *C. difficile* (99% das cepas testadas), enquanto resistência a moxifloxacina (MXF) e gatifloxacina (GAT), FQs de quarta geração, foram detectadas em 34% das cepas analisadas para esses antibióticos. A resistência aos antibióticos β -lactâmicos deve-se principalmente por dois mecanismos, sendo um deles relacionada a enzimas que degradam antibióticos (β -lactamases) e o outro devido à modificação de sítios alvo (proteínas de ligação à penicilina (PBPs)). Logo, a resistência que o *C. difficile* tem a múltiplos agentes, pode ser uma vantagem seletiva no que diz respeito a sua disseminação a partir do uso desses antibióticos (SPIGAGLIA, 2016).

Em 2017, Costa e colaboradores mostraram que a cepa ICC-45 foi resistente a CRO e a CLI, mas sensível a metronidazol, vancomicina, rifampicina, moxifloxacina e levofloxacina. O uso de cefalosporinas de terceira geração, como, por exemplo, a CRO tem sido associado a um alto risco de desenvolvimento de CDI (OWENS *et al.*, 2008; BLONDEAU, 2009). A resistência da cepa ICC-45 ao antimicrobiano CLI está de acordo com outros estudos que também relatam esse mesmo achado (OWENS *et al.*, 2008; SECCO *et al.*, 2014)

Um estudo realizado com o protozoário *G. lamblia*, demonstrou que o piruvato de ferredoxina oxidoreductase (Pfo) está envolvido com a resistência ao metronidazol (MTZ) nesse protozoário (LEITSCH *et al.*, 2011). Em 2013, Kondadi *et al.* descreveram que *H. pylori* contem enzimas nitroredutase que podem contribuir para a resistência ao metronidazol nessa bactéria (KONDADI *et al.*, 2013). Um estudo proteômico de isolado clínico de cepa NAP1 de *Clostridium difficile* realizado por Chong *et al.* (2014), mostrou que a redução da expressão do Pfo pode levar à redução da atividade do metronidazol, e possivelmente resultar

em um fenótipo de suscetibilidade reduzida ou resistente (CHONG *et al.*, 2014). Outro estudo de proteômica realizado por Dresler *et al.* (2017), mostrou que não foi observada a nitroreductase nos ribotipo (RTs) 027, 176 e 078 do *C. difficile*. No entanto, todos os isolados envolvidos nesse último estudo eram suscetíveis ao MTZ. Logo, esses autores sugeriram que a ausência de expressão dessa proteína provavelmente não desempenha um papel crucial na resistência ao MTZ (DRESLER *et al.*, 2017).

No grupo das proteínas que estão envolvidas com estresse oxidativo (cisteína dessulfurase e rubreritina) e proteínas de choque térmico (chaperonas), a cepa ICC-45 produziu em torno de quase 3x mais a quantidade de cisteína dessulfurase em relação à NAP/027 (LIBA5756) e também produziu bem mais chaperonas (DnaK, Hsp60 e Hsp90), por sua vez, a cepa NAP/027 (LIBA5756) produziu mais rubreritina, proteína associada ao estresse oxidativo, do que ICC-45. As ruberitinas são proteínas responsivas ao estresse oxidativo descritas inicialmente por seu papel na proteção de bactérias estritamente anaeróbicas. Um estudo de análise proteômica com cepa de *C. difficile* 630 em resposta ao estresse térmico, foi observado um aumento na abundância da rubreritina em resposta ao estresse térmico (JAIN *et al.*, 2011).

Uma pesquisa realizada por Giordano *et al.* (2018), mostrou que a cisteína dessulfurase (IscS2) desempenha um importante papel na resistência ao oxigênio em *C. difficile*. Nesse mesmo estudo, os autores mostraram que a deleção da suposta cisteína dessulfurase (IscS2) envolvida no sistema cluster de ferro-enzofre (Isc) causou um grave defeito de crescimento do *C. difficile* na presença de 2% de oxigênio. Além disso, esse mutante atrasou a colonização em um modelo convencional de CDI em camundongos (GIORDANO *et al.*, 2018). Sabe-se que baixos níveis de oxigênio estão presentes no trato gastrointestinal, sendo que as maiores concentrações estão associadas à superfície das células epiteliais. Nesse sentido, o mesmo grupo de pesquisa demonstraram que o *C. difficile* sob condições microaerófilas sobrevive em até 3% de oxigênio. Esses dados fornecem informações sobre possíveis mecanismos de adaptação do *C. difficile* ao ambiente hospedeiro e essa resistência ao oxigênio pode contribuir para patogênese do *C. difficile* (GIORDANO; HASTIE; CARLSON, 2018). Sabendo que a cepa ICC-45 produz mais cisteína dessulfurase do que a NAP/027 (LIBA5756), estudos adicionais são necessários para confirmar hipótese de que a cepa ICC-45 também possa resistir a possíveis condições microaerófilas no trato gastrointestinal.

As proteínas de choque térmico (Hsp) são uma família de chaperonas moleculares constitutivas ou induzidas pelo estresse, evolutivamente conservadas, sendo responsáveis

tanto pela manutenção da homeostase proteica em células procarióticas e quanto em eucarióticas (JOHNSTON *et al.*, 2018). Sabe-se que as chaperonas GroEL e a GroES fazem parte dos membros da família das Hsp60. Assim, um estudo realizado por Hennequin e colaboradores (2001), mostrou que em cultura de células epiteliais a GroEL (Hsp60) do *C. difficile* está envolvida na adesão celular (HENNEQUIN *et al.*, 2001). Em 2017, Jain *et al.* mostrou que a inativação do gene *dnaK* em *C. difficile* 630 produz um fenótipo sensível à temperatura. Nesse mesmo estudo, foi mostrando ainda que essa inativação aumenta a capacidade de formação de biofilme do *C. difficile* (JAIN *et al.*, 2017). Sabendo-se que a Cwp84 modula a formação de biofilme juntamente com a chaperona DnaK e que a ICC-45 produz essas proteínas em quantidades maiores do que a NAP/027 (LIBA5756), então estudos do nosso grupo em andamento pode confirmar essa hipótese da cepa ICC-45 produzir mais biofilme do que a NAP/027 (LIBA5756).

A cepa ICC-45 produz quase 3x mais a flavorubredoxina redutase do óxido nítrico (NO) quando comparada a NAP1/027 (LIBA5756) e da mesma forma, a ICC-45 também produz quantidades de flavoproteína putativa do óxido nítrico redutase a mais do que a NAP1/027 (LIBA5657). Um estudo de proteômica com cepas de *C. difficile* realizado por Chong e colaboradores (2014), mostraram uma maior expressão de proteínas redox-ativas tanto nas cepas CD26A54_R quanto na CD26A54_S e, entre essas proteínas estão a flavoproteína putativa do óxido nítrico redutase e a flavorubredoxina redutase do óxido nítrico no qual estão envolvidas com os efeitos do óxido nítrico (CHONG *et al.*, 2014). Nossos resultados mostram que o sobrenadante da cepa ICC-45 quando injetados em alça ileal de camundongo causa formação de NO tanto quanto o sobrenadante da cepa NAP1/027 (LIBA5657) (COSTA *et al.*, 2016). Pelo fato da ICC-45 produzir essas proteínas que o protegem contra os efeitos do NO em quantidade maior, essa cepa apesar de produzir mais estresse oxidativo deve ser mais resistente aos efeitos do NO produzidos pelo hospedeiro.

No grupo das proteínas que estão envolvidas com metabolismo e atividade catalítica, a ICC-45 produz mais trealose-6-fosfato, fosfo-transpiruvato-proteína-fosfotransferase e alanina racemase quando comparada a NAP1/027 (LIBA5756). Uma pesquisa realizada por Collins *et al.* (2018), mostrou que a trealose dietética tem participação importante na disseminação de dois ribotipos epidêmicos de *C. difficile* (RT027 e RT078). Esse estudo, mostrou que as cepas de ribotipos 027 continham uma mutação pontual no repressor de trealose do gene *treA* e essa mutação, aumentou a sensibilidade desse ribotipo à trealose em mais de 500 vezes. Os autores afirmaram ainda que a trealose dietética aumenta a virulência da cepa RT027 em um modelo de infecção em camundongos. As cepas RT078 adquiriram um cluster de quatro

genes envolvidos no metabolismo da trealose, que foi necessária e suficiente para o crescimento em baixas concentrações de trealose. Então, os autores propuseram que a implementação da trealose na dieta humana tenha desempenhado um papel relevante no surgimento dessas duas cepas epidêmicas e hipervirulentas (COLLINS *et al.*, 2018).

Um estudo com a fosfotransferase proteica de fosfoenolpiruvato (Pts) realizado por Nebenzahl *et al.* (2016), mostrou que em *Streptococcus pneumoniae* essa proteína por estar presente na sua parede celular atua como uma adesina. Os autores propuseram que peptídeos de ligação a Pts derivados de suas moléculas-alvo podem ser consideradas alvos para o desenvolvimento de vacinas para a pneumonia causada por essa bactéria (NEBENZAHL *et al.*, 2016). Então, esses dados podem ser importantes para o desenvolvimento de vacinas ativas que ofereçam oportunidades de prevenir a CDI já que tanto as cepas ICC-45 quanto a NAP/027 (LIBA5756) apresentam essa proteína.

Em 2017, Shrestha e colaboradores, mostraram que uma alanina racemase (Alr2) do *C. difficile* afeta a germinação de esporos e acomoda a serina como substrato. Nesse estudo, os esporos mutante de *alr2* germinam mais prontamente em resposta a L-alanina como co-germinante e que também a D-alanina funcionou como co-germinante. Os autores mostraram ainda que Alr2 poderia interconverter L- e D-serina e que por sua vez, a Alr2 ligue-se a L- e D-serina com uma afinidade 2 vezes menor que a L- e D-alanina. Com esses resultados, foi sugerido que esporos do *C. difficile* podem responder a um conjunto diversificado de aminoácidos co-germinativos de e que a Alr2 pode acomodar a serina como substrato (SHRESTHA; LOCKLESS; SORG, 2017).

No grupo das proteínas que apresentam outras atividades, a cepa ICC-45 secreta mais D-prolina redutase, fator de alongamento de transcrição (GreA) e pilin do que a NAP1/027 (LIBA5756), porém, essa secreta bem mais proteína de choque frio (CspB) quando comparada a ICC-45. Um estudo realizado por Bouillaut e colaboradores (2013) sugeriram que a D-prolina redutase em *C. difficile* atue como um regulador do metabolismo central controlando a utilização preferencial de prolina e glicina para produzir energia através das reações de Stickland (BOUILLAUT; SELF; SONENSHEIN, 2013). Em 2012, Li *et al.* mostrou que a proteína fator de alongamento de transcrição (GreA) em *E. coli* inibe a agregação de várias proteínas do substrato sob condições de choque térmico, possuindo assim, atividade de chaperona funcional. Nesse mesmo estudo, a GreA suprimiu a agregação de proteínas e promoveu a reativação de proteínas desnaturadas (LI *et al.*, 2012). Em outra pesquisa com análise semiquantitativa de proteínas envolvidas com estresse por calor na cepa de *C. difficile* 630, foi mostrado que a proteína GreA confere resistência ao calor e às tensões

oxidativas, fornecendo ligação entre a RNA polimerase/aparelho de transcrição e controle de qualidade de proteína, interagindo com subunidades de ribossomo e chaperonas incluindo a DnaK, DnaJ, GroES e ClpX (TERNAN *et al.*, 2014).

Em 2016, Maldarelli *et al.* mostraram que o Pili tipo IV promove a formação precoce de biofilme por *C. difficile* (MALDARELLI *et al.*, 2016). Em outro estudo recente, foi mostrado que o pili do tipo IV promove adesão celular e persistência do *C. difficile* em um modelo de infecção em camundongos. Isso contribui para a persistência do *C. difficile* no intestino do hospedeiro (MCKEE *et al.*, 2018). Isso sugere hipóteses para o nosso grupo de pesquisa que vem desenvolvendo trabalhos com as cepas ICC-45 e NAP1/027 (LIBA5756) para avaliar quem dessas cepas é mais virulenta e produz mais biofilme e quais proteínas contribuem diretamente para a patogênese e a formação de biofilme. Estudos mostram que a proteína de choque frio (CspB) juntamente com outras da família Csp (CspA, CspC) controlam a germinação de esporos em *C. difficile* (FIMLAID *et al.*, 2015; KEVORKIAN *et al.*, 2017; KEVORKIAN; SHEN, 2017). A partir desses dados, pode ser sugerido que a ICC-45 possa produzir mais biofilme do que a NAP1/027 (LIBA5756), já que a ICC-45 apresenta quantidades maiores de proteínas Cwp84, DnaK e pili (proteínas que estão envolvidas com a formação de biofilmes) do que a NAP1/027 (LIBA5756). Porém, essa hipótese ainda precisa ser confirmada com estudos adicionais que o nosso grupo vem desenvolvendo.

No Blast2Go observou-se que as proteínas detectadas exclusivamente na cepa ICC-45 estão igualmente ligadas com processos metabólicos e celulares. Já as proteínas compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756), 36% delas estão envolvidas com processos metabólicos. Nossos dados estão de acordo com a literatura, no qual Parreira *et al.* (2016) ao realizar uma análise comparativa do transcriptoma de RNAseq na enterite necrótica por *C. perfringens* durante a colonização *in vivo* e condições *in vitro*, mostrou que a maioria dos transcritos observados estão envolvidos em processos metabólicos (57.2%). Da mesma forma, Dresler *et al.* (2017) ao realizar uma análise de proteomas liberados em cultivo celular de oito ribotipos de *C. difficile*, os autores identificaram um total de 662 proteínas, dessas mais de 120 também estavam envolvidas em vias metabólicas (DRESLER *et al.*, 2017).

No Blast2Go nas categorias funcionais das proteínas identificadas somente na cepa ICC-45, observou-se seis tipos de funções (ligase, porfobilinogênio sintase, ligante de DNA, ligante de íons manganês, mutase e epimerase) variando entre 16% a 17%. Na categorização funcional das 26 proteínas selecionadas e compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756), observou-se nove funções (oxidoreductase, ligante de íons ferro, peptidase do

tipo cisteína, ligante de ATP, ligante de proteína, transferase, toxina, ligante de DNA e ligante de vitamina), dessas a maior parte estão envolvidas com atividade oxidoreductase (26%). De forma diferente, Jain *et al.* (2011) ao realizar um estudo de análise proteômica quantitativa da resposta ao estresse térmico em cepa de *C. difficile* 630, observou que das 447 proteínas identificadas, a maioria estavam envolvidas com funções de síntese proteica (19,5%) e metabolismo de aminoácidos e moléculas relacionadas (15%) (JAIN *et al.*, 2011). Esses resultados divergentes podem ser explicados pelo fato da cepa 630 ter sido submetida a estresse térmico e, conseqüentemente, produzida mais proteínas relacionadas com síntese proteica e metabolismo de aminoácidos.

Nas análises dos componentes celulares das proteínas selecionadas e compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756), observou-se que mais da metade proteínas estão vinculadas a célula (57%). Em 2016, estudo de Parreira *et al.* a partir de uma análise comparativa do transcriptoma de RNAseq na enterite necrótica por *C. perfringens*, observou vários transcritos fazendo parte dos componentes celulares, como, complexo proteico, atividade transportadora específica do substrato, ligante de carboidratos, atividade do transportador de drogas, parte celular e outros componentes (PARREIRA *et al.*, 2016). Esses dados de processos metabólicos e celulares podem sugerir adaptações tanto do *C. perfringens* quanto do *C. difficile* em ambiente intestinal do hospedeiro durante a CDI no caso do *C. difficile* e na enterite necrótica no caso da infecção por *C. perfringens*.

Na análise da localização celular das proteínas realizada pelo programa Psortb, observou-se que as proteínas identificadas apenas na cepa ICC-45, (60%) são de origem citoplasmática. Na localização celular das 26 proteínas selecionadas e compartilhadas entre as cepas ICC-45 e NAP1/027 (LIBA5756), observou também que mais da metade são de origem citoplasmática (58%). Esses nossos resultados estão de acordo com os achados de um estudo de análise proteômica de um filtrado de cultura comercial de *C. difficile* realizado por Moura *et al.* (2013), no qual observaram que das 101 proteínas identificadas, a maioria foram também de origem citoplasmática (72%) (MOURA *et al.*, 2013).

Todas as proteínas identificadas nas cepas ICC-45 e NAP1/027 (LIBA5756) foram analisadas no banco de dados STRING e apresentaram um padrão diferente nas interações entre as outras proteínas, ou seja, uma mesma área selecionada nas duas cepas mostrou semelhança nas interações da proteína A de camada superficial (SlpA) com as demais. No interatoma das proteínas exclusivas da ICC-45 com as 26 selecionadas e compartilhadas entre essas cepas, 3 proteínas da ICC-45 interagiram com as compartilhadas das cepas ICC-45 e NAP1/027 (LIBA5756) e duas não apresentaram nenhuma interação com as demais. Dentre

as 3 interações, a proteína IscR (regulador transcricional) interagiu com a chaperona DnaK, essa interagiu com a chaperona GroL, que conseqüentemente, interagiu com proteínas de patogenicidade (TcdA, TcdB, FliC, Cwp84 e SlpA), por sua vez, a IscR também interagiu com cisteína desulfurase (iscS). O ácido desamino-aminolevulínico desidratase (HemB) interagiu com a ribulose-fosfato 3-epimerase (rpe1). A rpe1 apresentou-se interagindo com outras duas proteínas, sendo elas a nitroreductase (CD2572) e a oxidoreductase de piruvato-ferredoxina (Pfo). No interatoma das 26 proteínas compartilhadas entre ICC-45 e NAP1/027 (LIBA5756), a maneira como as proteínas interagiram foram semelhante.

Nesse sentido, o que explica as interações das proteínas da cepa ICC-45 juntamente com as compartilhadas entre as cepas, como, por exemplo, a interação da IscR com a IscS é que ambas as proteínas estão envolvidas no metabolismo da cisteína. Esse sistema que abriga as vias biossintéticas do cluster de ferro-enxofre são acionadas quando há condições de estresse ou até mesmo falta de ferro. As interações da IscR com a IscS possibilitaram uma ligação direta com as chaperonas DnaK e GroL, já que essas também são moléculas induzidas em momentos de estresse térmico. As interações da DnaK e GroL, por sua vez, influenciaram nas proteínas de patogenicidade (TcdA, TcdB, FliC, Cwp84 e SlpA), isso sugere que todos esses fatores de virulência estão interligados diretamente com a patogenicidade da infecção causada por *C. difficile*, conseqüentemente, as interações das 3 proteínas exclusivas da cepa ICC-45 com os fatores de virulência possam estar influenciando na virulência dessa cepa e participando de múltiplos processos celulares (SANTOS *et al.*, 2014; GIORDANO *et al.*, 2018).

Os achados desse estudo mostram a importância da proteômica para esclarecer as possíveis proteínas envolvidas na patogênese da doença e na comparação entre cepas e complementa dados obtidos na genômica. É importante enfatizar que no mundo globalizado o surgimento e disseminação de novas cepas capazes de gerar surtos, requerem estudos detalhados que possam levar a compreensão de patogênese, identificação de biomarcadores e proteínas que possam levar ao desenvolvimento de vacinas.

7 CONCLUSÃO

Os resultados do presente estudo, obtidos na análise exoproteômica dos sobrenadantes das cepas ICC-45 e NAP1/027 (LIBA5756), permitem inferir:

- Existe uma homologia de 97,5% entre as proteínas das cepas ICC-45 e NAP1/027 (LIBA5756), considerada cepa hipervirulenta;
- As proteínas relacionadas com os fatores de virulência, toxina A e B, proteína da camada S (S-layer), proteína de parede celular e de superfície celular, cisteína protease (Cwp84), flagelos (FliC - flagelina) e enzimas hidrolíticas, foram identificadas em maior quantidade no sobrenadante da cepa ICC-45 quando comparada a NAP1/027 (LIBA5756);
- As proteínas exclusivas da cepa ICC-45 estão igualmente envolvidas com processos metabólicos e celulares. Já as proteínas compartilhadas entre as cepas, 36% delas estão ligadas a processos metabólicos;
- As proteínas exclusivas (IscR, HemB e Rpe1) da cepa ICC-45 interagem com as proteínas de patogenicidade (TcdA, TcdB, FliC, Cwp84 e SlpA) compartilhadas entre ICC-45 e NAP1/027 (LIBA5756).

8 REFERÊNCIAS

AEBERSOLD, R.; MANN, M. Mass spectrometry-based proteomics. **Nature.**, v. 422, n. 6928, p. 198–207, 2003.

AKTORIES, K.; SCHWAN, C.; JANK, T. *Clostridium difficile* Toxin Biology. **Annu Rev Microbiol.**, v. 71, p. 281–307, 2017.

ALDAPE, M. J. *et al.* Effects of ciprofloxacin on the expression and production of exotoxins by *Clostridium difficile*. **J Med Microbiol.**, p. 741–747, 2013.

ARMENGAUD, J.; DUPORT, C. Exoproteomics of Pathogens: Analysis of Toxins and Other Virulence Factors by Proteomics. **Methods Enzymol.**, v. 586, p. 211–227, 2017.

AWAD, M. J. *et al.* *Clostridium difficile* virulence factors: Insights into an anaerobic spore-forming pathogen. **Gut Microbes.**, v. 5, p. 579–593, 2014.

ARAÚJO JUNQUEIRA, A.F.T *et al.* Adenosine Deaminase Inhibition Prevents *Clostridium difficile* Toxin A-Induced Enteritis in Mice. **Infect Immun.**, v. 79, n. 2, p. 653–662, 2011.

ANTUNES, A. *et al.* Global transcriptional control by glucose and carbon regulator CcpA in *Clostridium difficile*. **Nucleic Acids Res.**, v. 40, n. 21, p. 10701–10718, 2012.

ARONSSON, B.; MÖLLBY, R.; Nord, C.E. Occurrence of toxin-producing *Clostridium difficile* in antibiotic-associated diarrhea in Sweden. **Med Microbiol Immunol.**, v. 170, p. 27–35, 1981.

AKANA, J. *et al.* D-ribulose 5-phosphate 3-epimerase: Functional and structural relationships to members of the ribulose-phosphate binding (beta/alpha)8-barrel superfamily. **Biochemistry**, v. 45, n. 8, p. 2493–2503, 2006.

BABAN, S. T. *et al.* The Role of Flagella in *Clostridium difficile* Pathogenesis: Comparison between a Non-Epidemic and an Epidemic Strain. **PLoS ONE.**, v. 8, n. 9, 2013.

BALASSIANO, I.; T. *et al.* Anaerobe Characterization of *Clostridium difficile* strains isolated from immunosuppressed inpatients in a hospital in Rio de Janeiro, Brazil. **Anaerobe.**, v. 15, p. 61–64, 2009.

BALASSIANO, I.; T. *et al.* *Clostridium difficile*: a problem of concern in developed countries and still a mystery in Latin America. **J Med Microbiol.**, v. 61, p. 169–179, 2012.

BANERJEE, S.; MAZUMDAR, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. **Intern J Anal Chem.**, v. 2012, p. 1–40, 2012.

BARRETO A.R. *et al.* Fucoidin Prevents *Clostridium difficile* Toxin-A-Induced Ileal Enteritis in Mice. **Dig Dis Sci.**, v. 53, p. 990–996, 2008.

BARTLETT, A. H.; PARK, P. W. Proteoglycans in host-pathogen interactions: Molecular mechanisms and therapeutic implications. **Expert Rev Mol Med.**, v. 12, p. 1–25, 2010.

BARTLETT, J.G. *et al.* Symptomatic Relapse After Oral Vancomycin Therapy of Antibiotic-Associated Pseudomembranous Colitis. **Gastroenterology.**, v. 78, p. 431–434, 1980.

BARTLETT, J. G. Annals of Internal Medicine Review Narrative Review: The New Epidemic of *Clostridium difficile* – Associated Enteric Disease. **Ann Intern Med.**, v. 145, p. 758–764, 2006.

BRADSHAW, W. J. *et al.* The structure of the S-layer of *Clostridium difficile*. **J Cell Commun Signal.**, v. 12, p. 319 - 331, 2018.

BARTLETT, J. G.; GERDING, D. N. Clinical Recognition and Diagnosis of *Clostridium difficile* Infection. **Clin Infect Dis.**, v. 46, p. S12–S18, 2008.

BOETZKES, A.; FELKEL, K. W.; ZEISER, J.; JOCHIM, N.; JUST, I.; PICH, A. Secretome analysis of *clostridium difficile* strains. **Arch Microbiol.**, v. 194, n. 8, p. 675–687, 2012.

BOUILLAUT, L. *et al.* Integration of metabolism and virulence in *Clostridium difficile*. **Res Microbiol.**, v. 166, n. 4, p. 375–383, 2015.

BRAUN, V. *et al.* Definition of the single integration site of the pathogenicity locus in *Clostridium difficile*. **Gene.**, v. 181, n. 1–2, p. 29–38, 1996.

BRITO, G. A. C. *et al.* Mechanism of *Clostridium difficile* Toxin A – Induced Apoptosis in T84 Cells. **J Infect Dis.**, v. 186, p. 1438–1447, 2002a.

BRITO, G. A. C. *et al.* *Clostridium difficile* Toxin A Alters In Vitro – Adherent Neutrophil Morphology and Function. **J Infect Dis.**, v. 185, p. 1297–1306, 2002b.

BROUWER, M. S. M. *et al.* Horizontal gene transfer converts non-toxigenic *Clostridium difficile* strains into toxin producers. **Nature Communications.**, v. 4, p. 1–6, 2013.

BURDON, D. W. *et al.* Faecal toxin and severity of antibiotic-associated pseudomembranous colitis 2. **J clin pathol.**, v. 34, n. 0021–9746; 5, p. 548–551, 1981.

BARDAĞ-GORCE, F. *et al.* delta-aminolevulinic dehydratase is a proteasome interacting protein. **Exp Mol Pathol.**, v. 91, p. 485–489, 2012.

BLONDEAU, J. M. What have we learned about antimicrobial use and the risks for *Clostridium difficile*-associated diarrhoea? **J Antimicrob Chemoth.**, 63:238-42, 2009.

BASHIRI, G. *et al.* Elongation of the Poly- γ -glutamate Tail of F420 Requires Both Domains of the F420: γ -Glutamyl Ligase (FbiB) of *Mycobacterium tuberculosis*. **J Biol Chem.**, v. 291, n. 13, p. 6882–6894, 2016.

BOUILLAUT, L.; SELF, W. T.; SONENSHEIN, A. L. Proline-Dependent Regulation of *Clostridium difficile* Stickland Metabolism. **J Bacteriol.**, v. 195, n. 4, p. 844–854, 2013.

CAPRIOTTI, A. L. *et al.* Intact protein separation by chromatographic and/or electrophoretic techniques for top-down proteomics. **J Chromatog.**, v. 1218, n. 49, p. 8760–8776, 2011.

CARTER, G. P. *et al.* The Anti-Sigma Factor TcdC Modulates Hypervirulence in an Epidemic BI / NAP1 / 027 Clinical Isolate of *Clostridium difficile*. **PLoS Pathog.**, v. 7, n. 10, 2011.

CARTER, G. P. *et al.* Binary Toxin Production in *Clostridium difficile* Is Regulated by CdtR, a LytTR Family Response Regulator. **J Bacteriol.**, v. 189, n. 20, p. 7290–7301, 2007.

CATHERMAN, A. D.; SKINNER, O. S.; KELLEHER, N. L. Top Down proteomics: Facts and perspectives. **Biochem Biophys Res Commun.**, v. 445, n. 4, p. 683–693, 2014.

CAVALCANTE, I. C. *et al.* Effect of Novel A2A Adenosine Receptor Agonist ATL 313 on *Clostridium difficile* Toxin A-Induced Murine Ileal Enteritis. **Infect Immun.**, v. 74, n. 5, p. 2606–2612, 2006.

CHALLAMALLA, P. *et al.* Negative ion mode mass spectrometry - an overview. **J Chem Biol Phys Science**, v. 2, n. 3, p. 1462–1471, 2012.

CHANDRASEKARAN, R.; LACY, D. B. The role of toxins in *clostridium difficile* infection. **FEMS Microbiol Ver.**, v. 41, n. 6, p. 723–750, 2017.

CHEN, B. *et al.* Proteomics progresses in microbial physiology and clinical antimicrobial therapy. **Eur J Clin Microbiol Infect Dis.**, v. 36, n. 3, p. 403–413, 2017.

CHEN, J. W. *et al.* Proteomic comparison of historic and recently emerged hypervirulent *clostridium difficile* strains. **J Proteome Res**, v. 12, n. 3, p. 1151–1161, 2013.

CHEN, P. *et al.* Structural basis for recognition of frizzled proteins by *Clostridium difficile* toxin B. **Science.**, v. 669, p. 664–669, 2018.

CHEN, S. *et al.* The Role of Rho GTPases in Toxicity of *Clostridium difficile* Toxins. **Toxins (Basel)**, v. 7, p. 5254–5267, 2015.

CHENG, K. *et al.* Recent development of mass spectrometry and proteomics applications in identification and typing of bacteria. **Proteomics Clin Appl.**, v. 10, n. 4, p. 346–357, 2016.

CHILTON, C. H. *et al.* Co-amoxiclav induces proliferation and cytotoxin production of *Clostridium difficile* ribotype 027 in a human gut model. **J Antimicrob Chemother.**, v. 67, p. 951–954, 2012.

CHILTON, C. H. *et al.* Comparative proteomic analysis of *clostridium difficile* isolates of varying virulence. **J Med Microbiol.**, v. 63, p. 489–503, 2014.

CHONG, P. M. *et al.* Proteomic analysis of a NAP1 *Clostridium difficile* clinical isolate resistant to metronidazole. **PLoS ONE**, v. 9, n. 1, 2014.

COSTA, C. L. *et al.* A MLST Clade 2 *Clostridium difficile* strain with a variant TcdB induces severe inflammatory and oxidative response associated with mucosal disruption. **Anaerobe.**, v. 40, p. 76–84, 2016.

COSTA, C. L. *et al.* Molecular epidemiology of *Clostridium difficile* infection in a Brazilian cancer hospital. **Anaerobe.**, v. 48, p. 232–236, 2017.

COSTA, C.L. *et al.* International Journal of Infectious Diseases Community-acquired diarrhea associated with *Clostridium difficile* in an HIV-positive cancer patient: first case report in Latin America. **Int J Infect Dis.**, v. 26, p. 138–139, 2014.

COX, J.; MANN, M. Quantitative, High-Resolution Proteomics for Data-Driven Systems Biology. **Ann Rev Bioch.**, v. 80, n. 1, p. 273–299, 2011.

CHAVES-OLARTE, E. *et al.* R-Ras Glucosylation and Transient RhoA Activation Determine the Cytopathic Effect Produced by Toxin B Variants from Toxin A-negative Strains of *Clostridium difficile*. **J. Biol Chem.**, v. 278, n. 10, p. 7956–7963, 2003.

COMMICHAU, F. M. *et al.* Novel Activities of Glycolytic Enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. **Mol Cell Proteomics.**, p. 1350–1360, 2009.

COLLINS, J. *et al.* Dietary trehalose enhances virulence of epidemic *Clostridium difficile*. **Nature.**, v. 553, n. 7688, p. 291–294, 2018.

DARKOH, C. *et al.*, Toxin Synthesis by *Clostridium difficile* Is Regulated through Quorum. **MBio.**, v. 6, n. 2, p. 1–10, 2015.

DESAI, K. *et al.* Epidemiological and economic burden of *Clostridium difficile* in the United States : estimates from a modeling approach. **BMC Infect Dis.**, v. 16, p. 303, 2016.

DI BELLA, S. *et al.* *Clostridium difficile* toxins A and B: Insights into pathogenic properties and extraintestinal effects. **Toxins.**, v. 8, n. 5, p. 1–25, 2016.

DINEEN, S. S. *et al.* Repression of *Clostridium difficile* toxin gene expression by CodY. **Mol Microbiol.**, v. 66, p. 206–219, 2007.

DINGLE, K. E. *et al.* Clinical *clostridium difficile*: Clonality and pathogenicity locus diversity. **PLoS ONE.**, v. 6, n. 5, 2011.

DON GJ, D. A. The association between antibiotic- associated diarrhoea and C. difficile toxin in children. **Aust N Z J Med.**, v. 11, p. 433–4, 1981.

DRESLER, J. *et al.* Analysis of proteomes released from in vitro cultured eight *Clostridium difficile* PCR ribotypes revealed specific expression in PCR ribotypes 027 and 176 confirming their genetic relatedness and clinical importance at the proteomic level. **Gut Pathogens.**, v. 9, p. 1–11, 2017.

DUBBERKE, E. R.; OLSEN, M. A. Burden of *Clostridium difficile* on the Healthcare System. **Clin Infect Dis.**, v. 55, p. 88–92, 2012.

DUBBERKE, E. R.; WERTHEIMER, A. I. Review of Current Literature on the Economic Burden of *Clostridium difficile* Infection. **Infect Control Hosp Epidemiol.**, v. 30, n. 01, p. 57–66, 2009.

DUNBAR, S. A.; ZHANG, H.; TANG, Y. W. Advanced techniques for detection and identification of microbial agents of gastroenteritis. **Clin Lab Med.**, v. 33, n. 3, p. 527–552, 2013.

DUPUY, B.; SONENSHEIN, A. L. Regulated transcription of *Clostridium difficile* toxin genes. **Mol Microbiol.**, v. 27, p. 107–120, 1998.

DARWIN, K. H. *et al.* The Proteasome of Mycobacterium tuberculosis Is Required for Resistance to Nitric Oxide The Proteasome of Mycobacterium tuberculosis Is Required for Resistance to Nitric Oxide. **Science.**, v. 302, p. 1963–1966, 2003.

EGERER, M. *et al.* Auto-catalytic Cleavage of *Clostridium difficile* Toxins A and B Depends on Cysteine Protease Activity. **J Biol Chem.**, v. 282, n. 35, p. 25314–25321, 2007.

ELLIOTT, B. *et al.* Infection, Genetics and Evolution *Clostridium difficile* infection: Evolution, phylogeny and molecular epidemiology. **Infect Genet Evol.**, v. 49, p. 1–11, 2017.

ERICKSON, A. R. *et al.* Integrated Metagenomics/Metaproteomics Reveals Human Host-Microbiota Signatures of Crohn's Disease. **PLoS ONE**, v. 7, n. 11, 2012.

EL-ANEED, A.; COHEN, A.; BANOUB, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. **Applied Spectroscopy Reviews.**, v. 44, p. 210–230, 2009.

FENN, J. B. *et al.* Electrospray ionization for mass spectrometry of large biomolecules. **Science.**, v. 246, n. 4926, p. 64–71, 1989.

FIMLAID, K.; A. *et al.* Identification of a Novel Lipoprotein Regulator of *Clostridium difficile* Spore Germination. **PLoS Pathog.**, v. 11, p. 23 – 9, 2015.

KEVORKIAN, Y.; SHEN, A. Revisiting the Role of Csp Family Proteins in Regulating *Clostridium difficile* Spore Germination. **J Bacteriol.**, v. 199, n. 22, p. 1–20, 2017.

KEVORKIAN, Y. *et al.* Regulation of *Clostridium difficile* spore germination by the CspA pseudoprotease domain. **Biochimie.**, v. 122, n. 802, p. 243–254, 2017.

GENISYUERREK, S. *et al.* Structural determinants for membrane insertion , pore formation and translocation of *Clostridium difficile* toxin B. **Mol Microbiol.**, v. 79, p. 1643–1654, 2011.

GEORGE, R. H. *et al.* Identification of *Clostridium difficile* as a cause of pseudomembranous colitis. **Br Med J.**, v. 1, n. 6114, p. 695, 1978.

GEORGE, W. L. *et al.* *Clostridium difficile* and Cytotoxin in Feces of Patients with Antimicrobial Agent-associated Pseudomembranous Colitis. **Infection.**, v. 10, n. 4, p. 205–207, 1982.

GEORGE, W. L.; SUTTER, V. L.; CITRON, D. Selective and differential medium for isolation of Selective and Differential Medium for Isolation of *Clostridium difficile*. **J Clin Microbiol.**, v. 9, n. 2, p. 214–219, 1979.

GESSEL, M. M.; NORRIS, J. L.; CAPRIOLI, R. M. MALDI imaging mass spectrometry: Spatial molecular analysis to enable a new age of discovery. **J Proteomics**, v. 107, p. 71–82, 2014.

GUARNER, F.; MALAGELADA, J. R. Gut flora in health and disease. **Lancet.**, v. 361, n. 9356, p. 512–519, 2003.

GENTH, H.; AKTORIES, K.; JUST, I. Monoglucosylation of RhoA at Threonine 37 Blocks Cytosol-Membrane Cycling. **J. Biol Chem.**, v. 274, p. 29050–29056, 1999.

GIORDANO, N. *et al.* Cysteine desulfurase IscS2 plays a role in oxygen resistance in *Clostridium difficile*. **Infect Immun.**, v. 86, p. 326–18, 2018.

GIORDANO, A. N.; HASTIE, J. L.; CARLSON, P.; E.. Transcriptomic profiling of *Clostridium difficile* grown under microaerophilic conditions. **Pathog Dis.**, v. 76, 2018.

HAMMOND, G. A.; JOHNSON, J. L. The toxigenic element of *Clostridium difficile* strain VPI 10463. **Microb Pathog.**, v. 19, n. 4, p. 203–213, 1995.

HU, Q. *et al.* The Orbitrap: A new mass spectrometer. **J Mass Spectr.**, v. 40, n. 4, p. 430–443, 2005.

HENNEQUIN, C. *et al.* GroEL of *Clostridium difficile* is involved in cell adherence. **Microbiology.**, v. 147, n. 2001, p. 87–96, 2001.

JANK, T.; AKTORIES, K. Structure and mode of action of clostridial glucosylating toxins: the ABCD model. **Trends Microbiol.**, v. 16, n. 5, p. 222–229, 2008.

JAIN, S. *et al.* Proteomic analysis of the insoluble subproteome of *Clostridium difficile* strain 630. **FEMS Microbiol Lett.**, v. 312, p. 151–159, 2011.

JAIN, S. *et al.* Quantitative proteomic analysis of the heat stress response in *Clostridium difficile* strain 630. **J Proteome Res.**, v. 10, n. 9, p. 3880–3890, 2011.

JAIN, S. *et al.* Inactivation of the dnaK gene in *Clostridium difficile* 630 Δ erm yields a temperature-sensitive phenotype and increases biofilm-forming ability. **Sci Rep**, v. 7, p. 1–13, 2017.

JI, Q. *et al.* *Staphylococcus aureus* CymR Is a New Thiol-based Oxidation-sensing Regulator of Stress Resistance and Oxidative Response . **J Biol Chem.**, v. 287, n. 25, p. 21102–21109, 2012.

JEDRZEJAS, M. J. *et al.* Structure and mechanism of action of a novel phosphoglycerate mutase from *Bacillus stearothermophilus*. **EMBO J.**, v. 19, n. 7, p. 1419–1431, 2000.

JOHNSTON, C. L. *et al.* Using Single-Molecule Approaches to Understand the Molecular Mechanisms of Heat-Shock Protein Chaperone Function. **J Mol Biol.**, v. 18, p. S0022-2836, 2018.

KARAS, M.; HILLENKAMP, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. **Anal chem.**, v. 60, n. 20, p. 2299–2301, 1988.

KARASAWA, T. *et al.* Effect of Arginine on Toxin Production *Clostridium difficile* in Defined Medium. **Microbiol Immunol.**, v. 41, n. 8, p. 581–585, 1997.

KARLSSON, S. *et al.* Expression of *Clostridium difficile* Toxins A and B and Their Sigma Factor TcdD Is Controlled by Temperature. **Infect Immun.**, v. 71, n. 4, p. 1784–1793, 2003.

KELLIE, J. F. *et al.* The emerging process of Top Down mass spectrometry for protein analysis: biomarkers, protein-therapeutics, and achieving high throughput. **Mol BioSyst.**, v. 6, n. 9, p. 1532, 2010.

KELLY, C. P.; KYNE, L. The host immune response to *Clostridium difficile*. **J Med Microbiol.**, v. 60, p. 1070–1079, 2011.

KUIJPER, E. J. *et al.* Emergence of *Clostridium difficile*-associated disease in North America and Europe. **Clin Microbiol Infect.**, p. 2–18, 2006.

KOMMINENI, S. *et al.* Nitric oxide-sensitive and -insensitive interaction of *Bacillus subtilis* NsrR with a ResDE-controlled promoter. **Mol Microbiol.**, v. 78, p. 1280–1293, 2010.

KONDADI, P. K. *et al.* Contingency nature of *Helicobacter bizzozeronii* oxygen-insensitive NAD(P)H-nitroreductase (HBZC1-00960) and its role in metronidazole resistance. **Vet Res.**, v. 44, n. 1, p. 10–13, 2013.

LIMA, B.B *et al.* *Clostridium difficile* toxin A attenuates Wnt/ β -catenin signaling in intestinal epithelial cells. **Infect Immun.**, v. 82, n. 7, p. 2680–2687, 2014.

GEORGE, W. L. *et al.* Aetiology of Antimicrobial-Agent-Associated Colitis. **Lancet.**, v. 311, n. 8068, p. 802–803, 1978.

LAWSON, P. A. *et al.* Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prévot 1938. **Anaerobe.**, v. 40, p. 95–99, 2016.

LESSA FC, WINSTON LG, *et al.* Burden of *Clostridium difficile* infection in the United States. **N Engl J Med.**, v. 372, p. 2369–70, 2015.

LIU, T. S. *et al.* Protective role of HSP72 against *Clostridium difficile* toxin A-induced intestinal epithelial cell dysfunction. **AJP: Cell Physiol.**, v. 284, n. 4, p. C1073–C1082, 2003.

LOO, V. G *et al.* A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. **N Engl J Med.**, v. 353, p. 2442–2449, 2005.

LÖSSL, P.; VAN DE WATERBEEMD, M.; HECK, A. J. The diverse and expanding role of mass spectrometry in structural and molecular biology. **EMBO J.**, v. 35, n. 24, p. 2634–2657, 2016.

LUIZ, A. *et al.* Identification of *Clostridium difficile* Asymptomatic Carriers in a Tertiary Care Hospital. **Bio Med Res Internat.**, v. 2017, 2017.

LYERLY, D. M. *et al.* Characterization of cross-reactive proteins detected by Culturette Brand Rapid Latex Test for *Clostridium difficile*. **J Clin Microbiol.**, v. 26, n. 3, p. 397–400, 1988.

LYERLY, D. M.; KRIVAN, H. C.; WILKINS, T. D. *Clostridium difficile*: Its Disease and Toxins. **Clin Microbiol Ver.**, v. 1, n. 1, p. 1–18, 1988.

LYON, S. A.; HUTTON, M. L.; ROOD, J. I.; CHEUNG, J. K.; LYRAS, D. CdtR Regulates TcdA and TcdB Production in *Clostridium difficile*. **PLoS Pathog.**, p. 1–19, 2016.

LEITSCH, D. *et al.* Pyruvate: Ferredoxin oxidoreductase and thioredoxin reductase are involved in 5-nitroimidazole activation while flavin metabolism is linked to 5-nitroimidazole resistance in *Giardia lamblia*. **J Antimicrob Chemother.**, v. 66, n. 8, p. 1756–1765, 2011.

LI, K. *et al.* Transcription Elongation Factor GreA Has Functional Chaperone Activity. **PLoS ONE**, v. 7, n. 12, p. 1–8, 2012.

MA, W. *et al.* Antibacterial mechanism of daptomycin antibiotic against *Staphylococcus aureus* based on a quantitative bacterial proteome analysis. **J Proteomics**, v. 150, p. 242–251, 2017.

MESQUITA, A.M.R.C *et al.* Diarreia nosocomial e outras infecções adquiridas em hospital universitário. **R. Ci. md. Biol.**, n. 85, p. 54–61, 2011.

MAKAROV, A. Electrostatic axially harmonic orbital trapping a high performance technique of mass analysis. **Anal Chem.**, v. 72, n. 6, p. 1156–1162, 2000.

MAKAROV, A. *et al.* Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. **Anal Chem.**, v. 78, n. 7, p. 2113–2120, 2006.

MANI, N.; DUPUY, B. Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. **PNAS.**, v. 98, n. 10, 2001.

MANI, N. *et al.* Environmental Response and Autoregulation of *Clostridium difficile* TxeR , a Sigma Factor for Toxin Gene Expression. **J Bacteriol.**, v. 184, n. 21, p. 5971–5978, 2002.

MARTIN, J. S. H.; MONAGHAN, T. M.; WILCOX, M. H. *Clostridium difficile* infection: Epidemiology, diagnosis and understanding transmission. **Nat Rev Gastroenterol Hepatol.**, v. 13, n. 4, p. 206–216, 2016.

MATAMOUROS, S.; ENGLAND, P.; DUPUY, B. *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. **Mol Microbiol.**, v. 64, p. 1274–1288, 2007.

MATSUSHITA, O.; OKABE, A. Clostridial hydrolytic enzymes degrading extracellular components. **Toxicon.**, v. 39, n. 11, p. 1769–1780, 2001.

MCCOLLUM, D. L.; RODRIGUEZ, J. M. Detection, Treatment, and Prevention of *Clostridium difficile* Infection. **Clin Gastroenterol Hepatol.**, v. 10, n. 6, p. 581–592, 2012.

MERRIGAN, M. *et al.* Human hypervirulent *Clostridium difficile* strains exhibit increased sporulation as well as robust toxin production. **J Bacteriol.**, v. 192, n. 19, p. 4904–4911, 2010.

MONCRIEF, J. S.; BARROSO, L. A.; WILKINS, T. D. Positive Regulation of *Clostridium difficile* Toxins. **Infect Immun.**, v. 65, n. 3, p. 1105–1108, 1997.

MONOT, M. *et al.* *Clostridium difficile*: New Insights into the Evolution of the Pathogenicity Locus. **Scientific Reports**, v. 5, p. 1–13, 2015.

MOURA, H. *et al.* Proteomic Analysis and Label-Free Quantification of the Large *Clostridium difficile* Toxins. **Int J Proteomics.**, v. 2013, p. 1–10, 2013.

MANINA, G. *et al.* Biological and structural characterization of the *Mycobacterium smegmatis* nitroreductase NfnB, and its role in benzothiazinone resistance. **Mol Microbiol.**, v. 77, n. 5, p. 1172–1185, 2010.

MALDARELLI, G.; A. *et al.* Type IV pili promote early biofilm formation by *Clostridium difficile*. **Pathog Dis.**, v. 76, p. 1–10, 2016.

MCKEE, R.; W. *et al.* Type IV pili promote *Clostridium difficile* adherence and persistence in a mouse model of infection. **Infect. Immun.**, v. 7290, p. 17–43, 2018.

NAPOLITANO, L.; M.; EDMISTON, C. E. *Clostridium difficile* disease: Diagnosis, pathogenesis, and treatment update. **Surgery.**, v. 162, n. 2, p. 325–348, 2017.

NORÉN, T. *et al.* Rapid and sensitive loop-mediated isothermal amplification test for *Clostridium difficile* detection challenges cytotoxin B cell test and culture as gold standard. **J Clin Microbiol.**, v. 49, n. 2, p. 710–711, 2011.

NUKUI, M. *et al.* Structure and Molecular Mechanism of *Bacillus anthracis* Cofactor-Independent Phosphoglycerate Mutase: A Crucial Enzyme for Spores and Growing Cells of *Bacillus* Species. **Biophys J.**, v. 92, p. 977–988, 2007.

O'FARRELL, P. H. High resolution two-dimensional electrophoresis of proteins. **The J Biol chem.**, v. 250, n. 10, p. 4007–21, 1975.

HALL I. C; O'TOOLE, E. Intestinal flora in new-born infants with a description of a new pathogenic anaerobe, *Bacillus difficilis*. **Am J Dis Child.**, v. 49, p. 390–402, 1935.

OTTO, A. *et al.* Global relative and absolute quantitation in microbial proteomics. **Curr Opin Microbiol.**, v. 15, n. 3, p. 364–372, 2012.

OWENS, R. C. *et al.* Antimicrobial-associated risk factors for *Clostridium difficile* infection. **Clin Infect Dis**, v. 46, p. 19–31, 2008.

PAREDES-SABJA, D.; SHEN, A.; SORG, J. A. *Clostridium difficile* spore biology: Sporulation, germination, and spore structural proteins. **Trends Microbiol.**, v. 22, n. 7, p. 406–416, 2014.

PECHINE, S.; JANOIR, C.; COLLIGNON, A. Variability of *Clostridium difficile* Surface Proteins and Specific Serum Antibody Response in Patients with *Clostridium difficile*-Associated Disease. **J Clin Microbiol.**, v. 43, n. 10, p. 5018–5025, 2005.

PENDERS, J. *et al.* Quantification of *Bifidobacterium* spp., *Escherichia coli* and *Clostridium difficile* in faecal samples of breast-fed and formula-fed infants by real-time PCR. **FEMS Microbiol Letters.**, v. 243, n. 1, p. 141–147, 2005.

PENG, W.-P.; CHOU, S.-W.; PATIL, A. A. Measuring masses of large biomolecules and bioparticles using mass spectrometric techniques. **The Analyst.**, v. 139, n. 14, p. 3507–3523, 2014. Disponível

PETERSON, L. R. *et al.* Detection of Toxigenic *Clostridium difficile* in Stool Samples by Real-Time Polymerase Chain Reaction for the Diagnosis of *C. difficile*-Associated Diarrhea. **Clin Infect Dis.**, v. 45, n. 9, p. 1152–1160, 2007.

PIRES, R. N. *et al.* American Journal of Infection Control *Clostridium difficile* infection in Brazil : A neglected problem? **Am J Infect Control.**, v. 42, n. 4, p. 459–460, 2014.

PETTIT, L. J. *et al.* Functional genomics reveals that *Clostridium difficile* Spo0A coordinates sporulation , virulence and metabolism. **BMC Genomics.**, v. 25, p. 1–15, 2014.

PURWANTINI, E.; MUKHOPADHYAY, B. Conversion of NO₂ to NO by reduced coenzyme F420 protects mycobacteria from nitrosative damage. **PNAS.**, v. 106, n. 15, p. 6333–6338, 2009.

PANTALÉON, V. *et al.* The *Clostridium difficile* Protease Cwp84 Modulates both Biofilm Formation and Cell- Surface Properties. **PLoS One.**, v. 10, p. 1–20, 2015.

PARREIRA, V.; R. *et al.* Comparative transcriptome analysis by RNAseq of necrotic enteritis *Clostridium perfringens* during in vivo colonization and in vitro conditions. **BMC Microbiol.**, v. 16, p. 1–16, 2016.

NGOK-NGAM, P. *et al.*, Roles of *Agrobacterium tumefaciens* RirA in Iron Regulation, Oxidative Stress Response, and Virulence. **J Bacteriol.**, v. 191, n. 7, p. 2083–2090, 2009.

NEBENZAHL, Y. M. *et al.* *Streptococcus pneumoniae* cell-wall-localized phosphoenolpyruvate protein phosphotransferase can function as an adhesin: Identification of its host target molecules and evaluation of its potential as a vaccine. **PLoS ONE**, v. 11, n. 3, p. 1–33, 2016.

QIU, H. *et al.* Novel *clostridium difficile* anti-Toxin (TcdA and TcdB) humanized monoclonal antibodies demonstrate in vitro neutralization across a broad spectrum of clinical strains and in vivo potency in a hamster spore challenge model. **PLoS ONE.**, v. 11, n. 6, p. 1–21, 2016.

QUESADA-GÓMEZ, *et al.* Emergence of *Clostridium difficile* NAP1 in Latin America. **J Clin Microbiol.**, v. 48, n. 2, p. 669–670, 2010.

RABILLOUD, T.; LELONG, C. Two-dimensional gel electrophoresis in proteomics: A tutorial. **J Proteomics.**, v. 74, n. 10, p. 1829–1841, 2011.

REICHARDT, C. *et al.* Dramatic increase of diarrhea associated with *Clostridium difficile* in Germany: has the new strain PCR-ribotype 027 reached us? **Dtsch Med Wochenschr.**, v. 132, p. 223–228, 2007.

RIEGLER, M. *et al.* *Clostridium difficile* Toxin B Is More Potent than Toxin A in damaging human colonic epithelium in vitro. **J Clin Invest.**, v. 95, p. 2004–2011, 1995.

RITER, L. S. *et al.* Comparison of the Paul ion trap to the linear ion trap for use in global proteomics. **Proteomics**, v. 6, p. 1735–1740, 2006.

ROCHA, M. F. G. *et al.* *Clostridium difficile* Toxin A Induces the Release of Neutrophil Chemotactic Factors From Rat Peritoneal Macrophages: Role of Interleukin-1beta, Tumor Necrosis Factor Alpha, and Leukotrienes. **Infect Immun.**, v. 65, n. 7, p. 2740–2746, 1997.

ROCHA, M. F. G. *et al.* Absence of intestinal secretion on supernatants from macrophages stimulated with *Clostridium difficile* toxin B on rabbit ileum. **Toxicon**, v. 39, p. 335–340, 2001.

RUPNIK, M. *et al.* A novel toxinotyping scheme and correlation of toxinotypes with serogroups of *Clostridium difficile* isolates. **J Clin Microbiol.**, v. 36, p. 2240–2247, 1998.

RUPNIK, M.; JANEZIC, S. An Update on *Clostridium difficile* Toxinotyping. **J Clin Microbiol.**, v. 54, n. 1, p. 13–18, 2016.

RUPNIK, M.; WILCOX, M. H.; GERDING, D. N. *Clostridium difficile* infection: New developments in epidemiology and pathogenesis. **Nat Rev Microbiol.**, v. 7, n. 7, p. 526–536, 2009.

RYAN, A. *et al.* A Role for TLR4 in *Clostridium difficile* Infection and the Recognition of Surface Layer Proteins. **PLoS Pathog.**, v. 7, n. 6, p. E20–76, 2011.

SANDERS, N. L. *et al.* Appendectomy and *Clostridium difficile* colitis: Relationships revealed by clinical observations and immunology. **World J Gastroenterol.**, v. 19, n. 34, p. 5607–5614, 2013.

SCHÄFFLER, H.; BREITRÜCK, A. *Clostridium difficile* – From Colonization to Infection. **Front Microbiol.**, v. 9, p. 1–12, 2018.

SCHEEPERS, L. *et al.* The intestinal microbiota composition and weight development in children : the KOALA Birth Cohort Study. **Int J Obes (Lond).**, v. 39, n. 1, p. 16–25, 2014.

SCHENCK, L. P.; BECK, P. L.; MACDONALD, J. A. Gastrointestinal dysbiosis and the use of fecal microbial transplantation in *Clostridium difficile* infection. **World J Gastrointest Pathophysiol.**, v. 6, n. 4, p. 169–180, 2015.

SCIGELOVA, M. *et al.* A. Fourier Transform Mass Spectrometry. **Mol Cell Proteomics.**, v. 10, n. 7, p. M111.009431, 2011.

SEBAIHIA, M. *et al.* The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. **Nat Genetics.**, v. 38, n. 7, p. 779–786, 2006.

SILVA, R. O. S. *et al.* *Clostridium difficile* ribotypes in humans and animals in Brazil. **Mem Inst Oswaldo Cruz**, v. 110, n. 8, p. 1062–1065, 2015.

SMITH LM, K. N. Proteoform: a single term describing protein complexity. **Nat Methods.**, v. 10, n. 3, p. 186–187, 2013.

SMITS, W. K. *et al.* *Clostridium difficile* infection. **Nat Rev Dis Primers.**, v. 7, p. 1–47, 2016.

STAMPER, P. D. *et al.* Comparison of a commercial real-time PCR assay for tcdB detection to a cell culture cytotoxicity assay and toxigenic culture for direct detection of toxin-producing *Clostridium difficile* in clinical samples. **J Clin Microbiol.**, v. 47, n. 2, p. 373–378, 2009.

STARE, B. G.; DELME, M. Variant forms of the binary toxin CDT locus and tcdC gene in *Clostridium difficile* strains. **J Med Microbiol.**, v. 56, p. 329–335, 2007.

STECHER, B.; MAIER, L.; HARDT, W. D. “Blooming” in the gut: How dysbiosis might contribute to pathogen evolution. **Nat Rev Microbiol.**, v. 11, n. 4, p. 277–284, 2013.

SVENSSON, A. M.; LASALA, P. R. Pathology consultation on detection of *clostridium difficile*. **Am J Clin Pathol.**, v. 137, n. 1, p. 10–15, 2012.

SWINDELLS, J. *et al.* Evaluation of diagnostic tests for *Clostridium difficile* infection. **J Clin Microbiol.**, v. 48, n. 2, p. 606–608, 2010.

SEHR, P. *et al.* Glucosylation and ADP ribosylation of rho proteins: effects on nucleotide binding, GTPase activity, and effector coupling. **Biochemistry.**, v. 37, p. 5296–5304

SANTOS, J. A.; PEREIRA, P. J. B.; MACEDO-RIBEIRO, S. What a difference a cluster makes: The multifaceted roles of IscR in gene regulation and DNA recognition. **Bioch Biophys Acta.**, v. 1854, p. 1101–1112, 2015.

SHEPARD, W. *et al.* Insights into the Rrf2 repressor family – the structure of CymR , the global cysteine regulator of *Bacillus subtilis*. **FEBS J.**, v. 278, p. 2689–2701, 2011.

SANTOS, J. A. *et al.* The unique regulation of iron-sulfur cluster biogenesis in a Gram-positive bacterium. **PNAS.**, v. 111, p E2251–E2260, 2014.

SOBOTA, J. M.; IMLAY, J. A. Iron enzyme ribulose-5-phosphate 3-epimerase in *Escherichia coli* is rapidly damaged by hydrogen peroxide but can be protected by manganese. **PNAS.**, v. 108, n. 13, p. 5402–5407, 2011.

SELENGUT, J. D.; HAFT, D. H. Unexpected abundance of coenzyme F420-dependent enzymes in *Mycobacterium tuberculosis* and other actinobacteria. **J Bacteriol.**, v. 192, n. 21, p. 5788–5798, 2010.

SUN, X.; HIROTA, S.; A. The roles of host and pathogen factors and the innate immune response in the pathogenesis of *Clostridium difficile* infection. **Mol Immunol.**, v. 63, n. 2, p. 193–202, 2016.

SLIMINGS, C.; RILEY, T.; V. Antibiotics and hospital-acquired *Clostridium difficile* infection: Update of systematic review and meta-analysis. **J Antimicrob Chemother.**, v. 69, n. 4, p. 881–891, 2014.

SPIGAGLIA, P. Recent advances in the understanding of antibiotic resistance in *Clostridium difficile* infection. **Ther Adv Infect Dis.**, v. 3, n. 1, p. 23–42, 2016.

SHRESTHA, R.; LOCKLESS, S. W.; SORG, X. J. A. A *Clostridium difficile* alanine racemase affects spore germination and accommodates serine as a substrate. **J Biol Chem.**, v. 292, p. 10735–10742, 2017.

SECCO, D.; A. *et al.* *Clostridium difficile* infection among immunocompromised patients in Rio de Janeiro, Brazil and detection of moxifloxacin resistance in a ribotype 014 strain. **Anaerobe.**, v. 28, p. 85-9, 2014.

TAN, K. S.; WEE, B. Y.; SONG, K. P. Evidence for holin function of tcdE gene in the pathogenicity of *Clostridium difficile*. **J Med Microbiol.**, v. 50, n. May, p. 613–619, 2018.

TAO, L. *et al.* Frizzled proteins are colonic epithelial receptors for *C. difficile* toxin B. **Nature.**, v. 538, p. 350–355, 2016.

THOLEY, A.; BECKER, A. Top-down proteomics for the analysis of proteolytic events - Methods, applications and perspectives. **Bioch Bioph Acta.**, v. 1864, n. 11, p. 2191–2199, 2017.

TINOCO-VERAS, C. M. *et al.* Transforming Growth Factor β 1/SMAD Signaling Pathway Activation Protects the Intestinal Epithelium from *Clostridium difficile* Toxin A-Induced Damage. **Infect Immun.**, v. 85, n. 10, p. 1–13, 2017.

TUBAON, R. M.; HADDAD, P. R.; QUIRINO, J. P. Sample Clean-up Strategies for ESI Mass Spectrometry Applications in Bottom-up Proteomics: Trends from 2012 to 2016. **Proteomics.**, v. 17, n. 20, p. 1–27, 2017.

TERNAN, N. G. *et al.* Semiquantitative Analysis of Clinical Heat Stress in *Clostridium difficile* Strain 630 Using a GeLC/MS Workflow with in PAI Quantitation. **PloS One.**, v. 9, n. 2, p. 88–96, 2014.

VOTH, D. E.; BALLARD, J. D. *Clostridium difficile* Toxins: Mechanism of Action and Role in Disease. **Clin. Microbiol. Rev.**, v. 18, n. 2, p. 247–263, 2005.

VINUSHA, K. S. *et al.* Proteomic studies on lactic acid bacteria: A review. **Biochem Bioph Rep.**, v. 14, p. 140–148, 2018.

WARNY, M. *et al.* Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. **Lancet.**, v. 366, p. 1079–1084, 2005.

WILM, M. Principles of Electrospray Ionization. **Mol Cell Proteomics**, v. 10, n. 7, p. M111.009407, 2011.

WRIGHT, A. *et al.* Proteomic analysis of cell surface proteins from *Clostridium difficile*. **Proteomics**, v. 5, n. 9, p. 2443–2452, 2005.

WESSEL, D.; FLÜGGE, U. I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. **Anal Biochem.**, v. 138, p. 141–143, 1984.

WRIGHT, A. *et al.* Proteomic analysis of cell surface proteins from *Clostridium difficile*. **Proteomics.**, v. 5, n. 9, p. 2443–2452, 2005.

YU, Y. *et al.* Diagnosing inflammation and infection in the urinary system via proteomics. **J Translat Med.**, v. 13, n. 1, p. 1–14, 2015.

YOSHINO, Y. *et al.* *Clostridium difficile* flagellin stimulates toll-like receptor5 , and toxin B promotes fl agellin-induced chemokine production via TLR5. **Life Sci.**, v. 92, n. 3, p. 211–217, 2013.

ZHANG, H.; GE, Y. Comprehensive Analysis of Protein Modifications by Top-Down Mass Spectrometry. **Cir Cardiovasc Genet.**, v. 4, n. 6, p. 711–711, 2011.

ZHANG, Y. *et al.* Protein Analysis by Shotgun/Bottom-up Proteomics. **Chem Rev.**, v. 31, n. 9, p. 1713–1723, 2013.

ZUBAREV, R. A.; MAKAROV, A. Orbitrap Mass Spectrometry. **Anal Chem**, v. 85, n. 11, p. 5288–5296, 2013.

ZEISER, J. *et al.* Substrate specificity of clostridial glucosylating toxins and their function on colonocytesanalyzed by proteomics techniques. **J Proteome Res.**, v. 12, p. 1604 – 18, 2013.

ANEXO I - Proteínas identificadas no sobrenadante das culturas das cepas ICC-45 e NAP1/027 (LIBA5756) de *C. difficile* pelo programa Scaffold.

Nº de proteínas	Identificação Proteínas (<i>C. difficile</i>)	Nº de acesso	Probabilidade de identificação
1	cell surface protein (S-layer precursor protein)	gi 255307831	100%
2	toxin A	gi 255305655	100%
3	toxin B	gi 126698238	100%
4	S-layer protein	gi 126700409	99%
5	acetyl-CoA acetyltransferase	gi 126698643	100%
6	NAD-specific glutamate dehydrogenase	gi 126697752	100%
7	pyruvate-ferredoxin oxidoreductase	gi 126700296	100%
8	Aminotransferase	gi 126701292	100%
9	cold shock protein CspB	gi 126698954	100%
10	3-hydroxybutyryl-CoA dehydrogenase	gi 126698642	100%
11	Rubryerythrin	gi 126699128	100%
12	chaperone DnaK	gi 126700078	100%
13	Enolase	gi 126700790	100%
14	isocaprenoyl-CoA:2-hydroxyisocaproate CoA-transferase	gi 126697964	100%
15	electron transfer flavoprotein subunit beta	gi 126697969	100%
16	oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase subunit C	gi 126697967	100%
17	putative surface protein	gi 255308293	100%
18	electron transfer flavoprotein subunit alpha	gi 126697970	100%
19	cell surface protein	gi 255307829	100%
20	formate-tetrahydrofolate ligase	gi 123453210	100%
21	50S ribosomal protein L7/L12	gi 126697631	100%

22	gamma-aminobutyrate metabolism dehydratase/isomerase	gi 126699959	100%
23	aldehyde dehydrogenase family protein	gi 531109986	100%
24	60 kDa chaperonin	gi 255305190	100%
25	trehalose-6-phosphate hydrolase	gi 126700708	100%
26	translation inhibitor endoribonuclease	gi 126700129	100%
27	Peptidase	gi 123452052	100%
28	peptidase D	gi 126698287	100%
29	subunit of oxygen-sensitive 2-hydroxyisocaproyl-CoA dehydratase	gi 255305394	100%
30	aspartate aminotransferase	gi 126698938	100%
31	glyceraldehyde-3-phosphate dehydrogenase	gi 126700794	100%
32	D-lactate dehydrogenase	gi 126697963	100%
33	pyrroline-5-carboxylate reductase	gi 126700900	100%
34	propanediol utilization phosphotransacylase	gi 126700297	100%
35	oligopeptide ABC transporter, substrate-binding protein	gi 255307707	100%
36	phosphoglycerate kinase	gi 126700793	100%
37	fructose-1,6-bisphosphate aldolase	gi 126697972	100%
38	alpha/beta hydrolase	gi 126699797	100%
39	anaerobic nitric oxide reductase flavorubredoxin	gi 126698752	100%
40	succinate-semialdehyde dehydrogenase	gi 126699960	100%
41	N-acetylmuramoyl-L-alanine amidase	gi 126700400	100%
42	toxin A	gi 260685955	100%
43	electron transfer flavoprotein subunit alpha	gi 126698640	100%
44	electron transfer flavoprotein subunit beta	gi 260686277	100%
45	Ferritin	gi 126699811	100%
46	phosphoenolpyruvate-protein phosphotransferase	gi 260687988	100%
47	elongation factor G	gi 126697637	100%
48			

	NAD-dependent 4-hydroxybutyrate dehydrogenase
49	putative nitric oxide reductase flavoprotein
50	Rubrerythrin
51	glutamyl-aminopeptidase
52	cysteine desulfurase
53	glucose-6-phosphate isomerase
54	proline racemase
55	putative cell wall hydrolase
56	ABC transporter, substrate-binding lipoprotein
57	50S ribosomal protein L1
58	butyryl-CoA dehydrogenase
59	formate acetyltransferase
60	cell surface protein
61	hypothetical protein CD630_16600
62	3-hydroxybutyryl-CoA dehydratase
63	ferredoxin/flavodoxin oxidoreductase subunit gamma
64	M24 family peptidase
65	elongation factor Ts
66	hypothetical protein CdifA_19096
67	acyl-CoA dehydrogenase
68	inosine 5-monophosphate dehydrogenase
69	beta-lactamase family protein
70	NifU-like protein
71	manganese-dependent inorganic pyrophosphatase
72	cell surface protein (putative cell surface-associated cysteine protease)
73	O-acetyl-serine thiol-lyase A

gi 255307361	100%
gi 255306635	100%
gi 126699078	100%
gi 126699768	100%
gi 126698876	100%
gi 126700904	100%
gi 126700857	100%
gi 123452506	100%
gi 255305875	100%
gi 126697629	100%
gi 260686276	100%
gi 255305763	100%
gi 1001999562	100%
gi 126699264	100%
gi 126698641	100%
gi 126697690	100%
gi 126700227	100%
gi 126699756	100%
gi 255308707	100%
gi 126697968	100%
gi 126699953	100%
gi 531115368	100%
gi 255306237	100%
gi 126697905	100%
gi 260210525	100%
gi 126699198	100%

74	trigger fator
75	glycine cleavage system protein H
76	Transketolase
77	Chain D, Alanine Racemase
78	elongation factor Tu
79	oligopeptide family ABC transporter substrate-binding protein
80	Rubrerythrin
81	2-hydroxyacyl-CoA dehydratase
82	6-phosphofructokinase
83	formate acetyltransferase
84	50S ribosomal protein L10
85	tellurium resistance protein
86	pyruvate kinase
87	D-proline reductase PrdA
88	putative aminotransferase
89	polynucleotide phosphorylase/polyadenylase
90	hypothetical protein CdifA_05660
91	ribosome recycling fator
92	putative phosphate butyryltransferase
93	50S ribosomal protein L5
94	putative amino acid racemase
95	Decarboxylase
96	butyrate kinase
97	30S ribosomal protein S16
98	DNA-directed RNA polymerase subunit beta
99	ferredoxin-NADP(+) reductase subunit alpha
100	hypothetical protein CdifA_19448
101	triosephosphate isomerase
102	tellurium resistance protein

gi 260688508	100%
gi 126698309	100%
gi 126699939	100%
gi 645985739	100%
gi 126697638	100%
gi 126698435	100%
gi 126698405	100%
gi 126699356	100%
gi 126701018	100%
gi 260688484	100%
gi 126697630	100%
gi 126699238	100%
gi 126701017	100%
gi 126700863	100%
gi 123453053	100%
gi 126698917	100%
gi 255306059	100%
gi 126699754	100%
gi 123453214	100%
gi 126697652	100%
gi 255307859	100%
gi 260687271	100%
gi 126697685	100%
gi 123453547	100%
gi 126697633	100%
gi 260686749	100%
gi 255308768	100%
gi 255308207	100%
gi 126699239	100%

103	glycine reductase complex component C subunit beta	gi 126699967	100%
104	adenylate kinase	gi 260685465	100%
105	50S ribosomal protein L3	gi 126697640	100%
106	toxin B	gi 260685953	100%
107	Peptidase	gi 260688699	100%
108	NUDIX family hydrolase	gi 260686089	100%
109	hypothetical protein CD630_24460	gi 400927444	100%
110	putative bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase	gi 255305718	100%
111	indolepyruvate oxidoreductase subunit	gi 255307406	100%
112	bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase complex subunit alpha/beta	gi 126698308	100%
113	30S ribosomal protein S8	gi 126697654	100%
114	cysteinyl-tRNA synthetase	gi 260685418	100%
115	cell wall protein	gi 400927472	100%
116	acetyl-CoA decarboxylase/synthase complex subunit gamma	gi 255305729	100%
117	Pilin	gi 255308544	100%
118	Oxidoreductase	gi 126699141	100%
119	nitroreductase-family protein	gi 126698718	100%
120	serine hydroxymethyltransferase	gi 126700341	100%
121	chaperonin GroES	gi 126697766	100%
122	translation initiation factor IF-3	gi 126698264	100%
123	hypothetical protein CDR20291_3527	gi 260688868	100%
124	phosphate butyryltransferase	gi 126697684	100%
125	thiamine-phosphate synthase	gi 126699205	100%
126	pyochelin synthetase F	gi 255305434	100%
127	DNA-directed RNA polymerase subunit alpha	gi 126697669	100%

128	flagellin subunit	gi 260685615	100%
129	tellurium resistance protein	gi 126699408	100%
130	bacteriocin-associated integral membrane family protein	gi 531109313	100%
131	cell surface protein (S-layer precursor protein)	gi 123451639	97%
132	hypothetical protein CD630_08070	gi 126698387	100%
133	hypothetical protein CD630_35710	gi 126701198	100%
134	diguanylate kinase signaling protein	gi 126699025	100%
135	30S ribosomal protein S3	gi 126697646	100%
136	50S ribosomal protein L4	gi 760237126	100%
137	threonine--tRNA ligase	gi 531119503	100%
138	50S ribosomal protein L11	gi 126697628	100%
139	cell-wall hydrolase	gi 126700384	100%
140	4Fe-4S binding domain-containing protein	gi 126697687	100%
141	DNA/RNA helicase	gi 400927363	100%
142	transcription elongation factor GreA	gi 126701179	100%
143	pyruvate phosphate dikinase	gi 260687649	100%
144	peptidase T	gi 126698630	100%
145	adenylosuccinate synthetase	gi 126701282	100%
146	4-aminobutyrate aminotransferase	gi 255307170	100%
147	elongation factor P	gi 126698842	100%
148	heat shock protein 90	gi 126697845	100%
149	rubredoxin oxidoreductase	gi 260686125	100%
150	ferredoxin/flavodoxin oxidoreductase subunit alpha	gi 126697688	100%
151	carbon monoxide dehydrogenase/acetyl-CoA synthase complex, dihydrolipoyl dehydrogenase subunit	gi 260686019	100%
152	GMP synthase	gi 126697769	100%
153	1-phosphofructokinase	gi 531111417	100%

154	n-acetylmuramoyl-l-alanine amidase
155	cell division protein FtsZ
156	heat shock protein
157	tellurium resistance protein
158	proline reductase
159	cell wall binding repeat 2 family protein
160	Nitroreductase
161	S-adenosylmethionine synthetase
162	phosphoglyceromutase
163	sugar family ABC transporter substrate-binding protein
164	putative polysaccharide deacetylase
165	DNA polymerase III subunit beta
166	Translaldolase
167	iron-only hydrogenase, catalytic subunit
168	hypothetical protein CdifA_11162
169	Xaa-Pro dipeptidase
170	30S ribosomal protein S13
171	50S ribosomal protein L24
172	Oxidoreductase
173	aspartate aminotransferase
174	putative signaling protein
175	hypothetical protein CD630_19530
176	sigma-54 factor interaction domain-containing protein
177	cell surface protein
178	hypothetical protein CD196_1925
179	Thioredoxin
180	dihydrodipicolinate synthase 3
181	50S ribosomal protein L21

gi 260686827	100%
gi 126700260	100%
gi 123451155	100%
gi 123450620	100%
gi 260688442	100%
gi 531118578	100%
gi 126700191	100%
gi 126697702	100%
gi 255308206	100%
gi 126700259	100%
gi 123452999	100%
gi 123452711	100%
gi 126699947	100%
gi 260688570	100%
gi 255307139	100%
gi 126699965	100%
gi 126697666	100%
gi 400927242	100%
gi 126700638	100%
gi 760239936	100%
gi 255305513	100%
gi 126699562	100%
gi 255305861	100%
gi 760239296	100%
gi 260683661	100%
gi 126699296	100%
gi 126700845	100%
gi 126698756	100%

182	30S ribosomal protein S4	gi 126697668	100%
183	delta-aminolevulinic acid dehydratase	gi 255308448	100%
184	phosphohexomutase	gi 126699936	100%
185	hypothetical protein CdifA_02508	gi 255305433	100%
186	hemagglutinin/adhesin	gi 126698095	83%
187	hydrolase beta-lactamase-like	gi 126700357	100%
188	30S ribosomal protein S7	gi 126697636	99%
189	Rrf2 family transcriptional regulator	gi 126698875	100%
190	threonine dehydratase II	gi 126700130	100%
191	hypothetical protein CD630_19700	gi 126699579	100%
192	ribulose-phosphate 3-epimerase	gi 126700194	100%
193	deoxyguanosinetriphosphate triphosphohydrolase-like protein	gi 255306687	99%
194	DNA primase	gi 260686668	99%
195	putative hydrolase	gi 123450945	100%
196	hypothetical protein CD630_02790	gi 126697851	100%
197	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide isomerase	gi 260686765	100%