

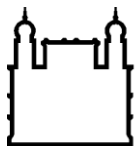
MINISTÉRIO DA SAÚDE
FUNDAÇÃO OSWALDO CRUZ
INSTITUTO OSWALDO CRUZ

Doutorado em Programa de Pós-Graduação em Biologia Computacional e Sistemas

CARACTERIZAÇÃO DE MECANISMOS DE REGULAÇÃO
TRANSCRICIONAL DOS GENES DO SISTEMA DE SECREÇÃO DO
TIPO VI EM *Klebsiella pneumoniae*

VICTOR AUGUSTO ARAÚJO BARBOSA

Rio de Janeiro
Agosto de 2019



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

Victor Augusto Araújo Barbosa

Caracterização de mecanismos de regulação transcricional dos genes do sistema de secreção do tipo VI em *Klebsiella pneumoniae*

Tese apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Doutor em Ciências

Orientadora: Dra. Leticia Miranda Lery Santos

RIO DE JANEIRO

Agosto de 2019

Barbosa, Victor Augusto Araújo.

Caracterização de mecanismos de regulação transcricional dos genes do sistema de secreção do tipo VI em *Klebsiella pneumoniae* / Victor Augusto Araújo Barbosa. - Rio de Janeiro, 2019.

170 f.; il.

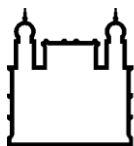
Tese (Doutorado) - Instituto Oswaldo Cruz, Pós-Graduação em Biologia Computacional e Sistemas, 2019.

Orientadora: Leticia Miranda Lery Santos.

Bibliografia: Inclui Bibliografias.

1. Sistema de secreção do tipo VI. 2. T6SS. 3. Regulação transcricional .
4. *Klebsiella pneumoniae*. 5. Genoma bacteriano. I. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da Biblioteca de Manguinhos/ICICT com os dados fornecidos pelo(a) autor(a).



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

AUTOR: Victor Augusto Araújo Barbosa

CARACTERIZAÇÃO DE MECANISMOS DE REGULAÇÃO TRANSCRICIONAL DOS GENES DO SISTEMA DE SECREÇÃO DO TIPO VI EM *Klebsiella pneumoniae*

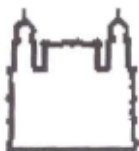
ORIENTADORA: Prof. Dra. Leticia Miranda Lery Santos

Aprovada em: ____/____/____

EXAMINADORES:

Prof. Dr. Thiago Estevam Parente (FIOCRUZ-IOC)
Prof. Dr. Bianca Cruz Neves (UFRJ)
Prof. Dr. Wanda Maria Almeida von Kruger (UFRJ)
Prof. Dr. Fabio Faria da Mota (FIOCRUZ-IOC) - Suplente
Prof. Dr. Gisele Cardoso de Amorim (UFRJ) - Suplente

Rio de Janeiro, 27 de agosto de 2019



Ministério da Saúde

Fundação Oswaldo Cruz
Instituto Oswaldo Cruz

Ata da defesa de tese de doutorado em Biologia Computacional e Sistemas de **Victor Augusto Araújo Barbosa**, sob orientação da Dr^a. Leticia Miranda Lery Santos. Ao vigésimo sétimo dia do mês de agosto de dois mil e dezenove, realizou-se às nove horas, no Auditório Carlos Chagas/FIOCRUZ, o exame da tese de doutorado intitulada: **“Caracterização de mecanismos de regulação transcricional dos genes do sistema de secreção do tipo VI em *Klebsiella pneumoniae*”**, no programa de Pós-graduação em Biologia Computacional e Sistemas do Instituto Oswaldo Cruz, como parte dos requisitos para obtenção do título de Doutor em Ciências - área de concentração: Genômica Funcional, Evolução e Filogenômica, na linha de pesquisa: Genoma, transcriptoma, proteoma e metaboloma. A banca examinadora foi constituída pelos Professores: Dr. Thiago Estevam Parente Martins - IOC/FIOCRUZ (Presidente), Dr^a. Bianca Cruz Neves - UFRJ/RJ, Dr^a. Wanda Maria Almeida Von Kruger - UFRJ/RJ e como suplentes: Dr. Fabio Faria da Mota – IOC/FIOCRUZ e Dr^a. Gisele Cardoso de Amorim – UFRJ/RJ. Após arguir o candidato e considerando que o mesmo demonstrou capacidade no trato do tema escolhido e sistematização da apresentação dos dados, a banca examinadora pronunciou-se pela APROVAÇÃO da defesa da tese de doutorado. De acordo com o regulamento do Curso de Pós-Graduação em Biologia Computacional e Sistemas do Instituto Oswaldo Cruz, a outorga do título de Doutor em Ciências está condicionada à emissão de documento comprobatório de conclusão do curso. Uma vez encerrado o exame, a coordenadora do Programa, Dr^a. Ana Carolina Ramos Guimarães, assinou a presente ata tomando ciência da decisão dos membros da banca examinadora. Rio de Janeiro, 27 de agosto de 2019.

Dr. Thiago Estevam Parente Martins (Presidente da Banca):

Dr^a. Bianca Cruz Neves (Membro da Banca):

Dr^a. Wanda Maria Almeida Von Kruger (Membro da Banca):

Dr^a. Ana Carolina Ramos Guimarães (Coordenadora do Programa):

DEDICATÓRIA

Aos meus pais por não terem medido esforços, em momento algum, para que eu pudesse começar, seguir e terminar mais esta jornada.

À minha prima Aleida Moura que me acompanha desde a tenra infância e sempre acreditou no meu potencial de ir cada vez mais longe.

Aos meus amigos com os quais moldei meus pensamentos libertários, que de alguma forma me fizeram ter vontade e força para trilhar a carreira docente.

Às minhas orientadoras da Graduação, Dra. Waleska Ferreira de Albuquerque, e do Mestrado, Dra. Maria Christina Sanchez Muratori, por acreditarem sempre no meu trabalho e me incentivarem a ir cada vez mais longe.

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES pelo auxílio financeiro.

À Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro – FAPERJ.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq.

À Fundação Oswaldo Cruz – Fiocruz.

Ao Laboratório de Microbiologia Celular – LAMICEL IOC Fiocruz.

À minha orientadora, Dra. Leticia Miranda Lery Santos, por todos os ensinamentos e pelo suporte fornecidos nestes quatro anos de Doutorado.

Aos coordenadores da Pós-graduação em Biologia Computacional e Sistemas, IOC Fiocruz.

A todos os meus amigos que de alguma forma contribuíram com suporte emocional ou financeiro.

Aos familiares que me incentivaram na conclusão de mais essa jornada.

Há uma diferença muito grande entre saber e acreditar que se sabe.

Saber é ciência. Acreditar que se sabe é ignorância.

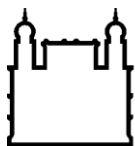
Mas, cuidado! Saber mal, não é ciência.

Saber mal pode ser muito pior que ignorar.

Na verdade, sabe-se somente quando se sabe pouco, pois com o saber, cresce a
dúvida.

viii Que é preciso idolatrar, sempre!

Antônio Abujamra



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

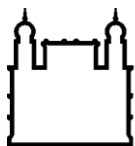
CARACTERIZAÇÃO DE MECANISMOS DE REGULAÇÃO TRANSCRICIONAL DOS GENES DO SISTEMA DE SECREÇÃO DO TIPO VI EM *Klebsiella pneumoniae*

RESUMO

TESE DE DOUTORADO EM BIOLOGIA COMPUTACIONAL E SISTEMAS

Victor Augusto Araújo Barbosa

Klebsiella pneumoniae é um patógeno oportunista que causa principalmente infecções respiratórias e do trato urinário. A ocorrência frequente de isolados virulentos resistentes a múltiplas drogas levou a inclusão dessa espécie na lista da OMS das principais prioridades para pesquisa e desenvolvimento de alternativas terapêuticas. O conhecimento abrangente dos mecanismos moleculares subjacentes à virulência da *K. pneumoniae* pode levar à proposta de medicamentos mais eficientes e específicos. O Sistema de Secreção Tipo VI (T6SS) contribui para a competição bacteriana, invasão celular e colonização *in vivo*. Apesar dos estudos que mostram o envolvimento do T6SS na patogênese da *K. pneumoniae*, pouco se conhece sobre a regulação de sua expressão. O entendimento dos mecanismos regulatórios pode fornecer pistas sobre a função desse sistema e contribuir para o desenvolvimento de novas abordagens terapêuticas para o tratamento de infecções por *K. pneumoniae*. Esse trabalho teve como objetivo identificar mecanismos de regulação transcricional dos genes do T6SS de *K. pneumoniae*. Para isso, analisamos os genomas de três cepas (Kp52.145, HS11286 e NTUH-K2044) de forma a prever e padronizar a anotação dos genes de T6SS através de buscas de similaridade. Foram encontrados 38 genes do T6SS em Kp52.145, 29 em HS11286 e 30 em NTUH-K2044. Nas adjacências dos genes do T6SS foram encontrados genes envolvidos em sistemas de captação de ferro, sugerindo que o T6SS de *K. pneumoniae* também pode desempenhar um papel na importação de íons. Foram identificadas 17 regiões promotoras dependentes de σ^{70} em Kp52.145, 12 em HS11286 e 12 em NTUH-K2044. Identificamos ainda 17, 12 e 15 sequências promotoras a partir dos sítios putativos de σ^{54} . Também identificamos 165 sítios de ligação para reguladores transcricionais em Kp52.145, 125 em HS11286 e 134 em NTUH-K2044. Nossos resultados *in silico* sugerem que o T6SS de *K. pneumoniae* seja regulado em resposta a sinais ambientais e devem direcionar experimentos *in vitro* que testem a resposta de *K. pneumoniae* a variações de temperatura (H-NS), de nutrientes (GcvA e Fis), estresse oxidativo (OxyR) e osmolaridade (RscAB e OmpR).



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

CHARACTERIZATION OF THE TRANSCRIPTIONAL REGULATORY MECHANISMS OF THE TYPE VI SECRETION SYSTEM GENES IN *Klebsiella pneumoniae*

ABSTRACT

PHD THESIS IN COMPUTATIONAL AND SYSTEMS BIOLOGY

Victor Augusto Araújo Barbosa

Klebsiella pneumoniae is an opportunistic pathogen that mainly causes respiratory and urinary tract infections. The frequent occurrence of virulent and multiple drug-resistant isolates led the WHO to include this species in the list of top priorities for research and development of therapeutic alternatives. A comprehensive knowledge of the molecular mechanisms underlying *K. pneumoniae* virulence may lead to the proposal of more efficient and specific drugs. One of its virulence factors is the Type VI Secretion System (T6SS), which contributes to bacterial competition, cell invasion and *in vivo* colonization. Despite the few studies showing the involvement of T6SS in *K. pneumoniae* pathogenesis, little is known concerning the regulation of its expression. The understanding of regulatory mechanisms may give more clues about the function of the system and the possibilities of future interference in this process. This work aimed to standardize the annotation of T6SS genes in *K. pneumoniae* strains and identify mechanisms of their transcriptional regulation through computational predictions. We analyzed the genomes of Kp52.145, HS11286 and NTUH-K2044 strains to perform a broad prediction and re-annotation of T6SS genes through similarity searches. 38 genes were found in Kp52.145, while 29 in HS11286 and 30 in NTUH-K2044. Genes coding for iron uptake systems are encoded in adjacencies of T6SS, suggesting that *K. pneumoniae* T6SS might also play a role in ion import. Some of the T6SS genes are comprised of syntenic regions. 17 σ^{70} -dependent promoter regions were identified in Kp52.145, 12 in HS11286, and 12 in NTUH-K2044. In addition, 17, 12 and 15 σ^{54} -dependent promoter sequences were found in Kp52.145, HS11286 and NTUH-K2044. Using Virtual Footprint we retrieved a total of 165 binding sites for transcriptional regulators in Kp52.145 and 125 in HS11286 and 134 in NTUH-K2044. The data presented herein are derived from computational analysis. Although future experimental studies are required to confirm those predictions, they suggest that *K. pneumoniae* T6SS might be regulated in response to environmental signals that are indeed sensed by the bacteria inside the human host: temperature (H-NS), nutrition limitation (GcvA and Fis), oxidative stress (OxyR) and osmolarity (RscAB and OmpR).

Sumário

RESUMO	ix
1 INTRODUÇÃO.....	17
1.1 Sistema de secreção do Tipo VI (T6SS)	17
1.2 Componentes do T6SS.....	20
1.2.1 Complexo de Membrana - CM (TssL, TssM e TssJ)	20
1.2.2 Ponta do sistema - Tssl e PAAR.....	21
1.2.3 Complexo da base (<i>baseplate</i> : TssE, TssF, TssG e TssK)	23
1.2.4 Tubo (TssD) e cauda (TssB/TssC).....	24
1.2.5 TssA	25
1.2.6 Reciclagem – TssH.....	26
1.3 Componentes acessórios.....	28
1.4 Efetores e proteínas de imunidade do T6SS.....	29
1.5 O papel do T6SS nas interações bactéria-bactéria	30
1.6 O papel do T6SS nas interações com células do hospedeiro	31
1.7 O papel do T6SS na captação de íons metálicos.....	32
1.8 Genes órfãos do T6SS.....	33
1.9 Regulação do T6SS.....	33
1.10 Regulação transcricional.....	36
1.11 <i>Klebsiella pneumoniae</i>	38
1.12 T6SS em <i>Klebsiella pneumoniae</i>	39
1.13 Justificativa	41
2 OBJETIVOS.....	42
2.1 Geral.....	42
2.2 Específicos	42
3 MÉTODOS	43
3.1 Sequências genômicas de <i>K. pneumoniae</i>	43
3.2 Identificação <i>in silico</i> de genes codificadores do T6SS	43
3.2.1 Identificação dos componentes principais.....	43
3.2.2 Identificação dos componentes acessórios	46
3.3 Análise das sequências dos genes putativos do T6SS	47
3.3.1 Alinhamento de sequências	47
3.3.2 Determinação conteúdo GC.....	47
3.4 Construção das sintenias.....	47
3.5 Identificação das sequências promotoras	47
3.5.1 Sequências promotoras de sigma-70 (σ^{70}).....	47
3.5.2 Sequências promotoras de sigma-54 (σ^{54}).....	48
3.6 Reguladores transcricionais codificados nos genomas de <i>K. pneumoniae</i>	49

3.7 Predição de sítios de ligação de fatores transcricionais nas sequências promotoras dos genes do T6SS.....	50
4 Resultados e Discussão	52
4.1 Identificação de genes do T6SS nas cepas de <i>K. pneumoniae</i>	52
4.2 TssA, TssA1 e TagA em <i>K. pneumoniae</i>	59
4.3 ATPases Clp: componentes TssH identificados em <i>K. pneumoniae</i>	61
4.4 Contexto genético dos genes órfãos.....	66
4.5 Sintenia entre <i>loci</i> do T6SS em <i>K. pneumoniae</i>	69
4.6 Regiões promotoras.....	73
4.6.1 Sítios de ligação de fatores sigma σ	73
5 PERSPECTIVAS	88
6 CONCLUSÕES.....	89
7 REFERÊNCIAS BIBLIOGRÁFICAS.....	90
8 APÊNDICES.....	107
APÊNDICE A - Genes putativos do T6SS em Kp52.145.....	107
APÊNDICE B - Genes putativos do T6SS em HS11286.....	108
APÊNDICE C – Genes putativos do T6SS em NTUH.....	109
APÊNDICE D – Sequências promotoras σ^{70}	111
APÊNDICE E - Sequências promotoras σ^{54}	117
APÊNDICE F - Sequências promotoras para σ^{70} de genes do T6SS de Kp52.145.....	123
APÊNDICE G- Sequências promotoras para σ^{70} de genes do T6SS de HS11286.....	124
APÊNDICE H - Sequências promotoras para σ^{70} de genes do T6SS de NTUH-K2044	125
APÊNDICE I- Sequências promotoras para σ^{54} de genes do T6SS de Kp52.145.....	126
APÊNDICE J - Sequências promotoras para σ^{54} de genes do T6SS de HS11286.....	127
APÊNDICE K- Sequências promotoras para σ^{54} de genes do T6SS de NTUH-K2044.....	128
APÊNDICE L – Sítios putativos dos promotores dependentes de σ^{70} em Kp52.145	129
APÊNDICE M – Artigo publicado na revista <i>BMC Genomics</i>	159

ÍNDICE DE FIGURAS

Figura 1: Arquitetura do T6SS.....	19
Figura 2: Estrutura do complexo de membrana do T6SS de <i>E. coli</i>	21
Figura 3: Complexo da ponta do T6SS, composto pelas proteínas TssI e PAAR.....	23
Figura 4: Cauda do T6SS. A.....	24
Figura 5: Arquitetura dos domínios das proteínas TssA, TssA1 e TagA.	26
Figura 6: Arquitetura dos domínios da proteína ClpV.....	27
Figura 7: Co-localização de genes que codificam pares de proteína efetor-imunidade em <i>E. coli</i> EAEC 042, UPEC CFT073 e UPEC UTI89.	30
Figura 8: Representação esquemática dos diversos sistemas que modulam a expressão e ativação de T6SS em bactérias.....	35
Figura 9: Interação da RNA polimerase com subunidades σ^{70} e σ^{54}	37
Figura 10: Representação de fatores de virulência em <i>K. pneumoniae</i> ..	39
Figura 11: Estratégia usada pelo T6SS-HMMER/SecReT6 para prever e identificar genes que codificam o T6SS.....	45
Figura 12: Representação das regiões dos motivos para identificação de sítios de ligação do fator σ^{70}	48
Figura 13: Representação da construção de uma matriz de peso-posição no Virtual Footprint.	50
Figura 14: Características gerais dos genomas de <i>K. pneumoniae</i> analisados e a localização dos genes que codificam componentes do T6SS...55	
Figura 15: Representação esquemática das sequências de aminoácidos de TssA e TagA de <i>K. pneumoniae</i> , ressaltando os domínios conservados.....	60
Figura 16: Re-anotação da sequência de aminoácidos da ClpB da cepa HS11286, tendo como base às ClpBs de Kp52145 e NTUH-K2044.	62
Figura 17: Alinhamento das sequências de aminoácidos das proteínas ClpB de <i>F. tularensis</i> e <i>K. pneumoniae</i>	63
Figura 18: Alinhamento múltiplo das proteínas ClpV, ClpB e ClpA de <i>K. pneumoniae</i> Kp52.145.....	65
Figura 19: Análise do contexto genômico de genes órfãos do T6SS em Kp52.145, HS11286 e NTUH-K2044.	68
Figura 20: Regiões que codificam genes relacionados com o T6SS em <i>K. pneumoniae</i> Kp52.145, <i>E. coli</i> E042 e <i>P. ananatis</i> LMG106.....	69

Figura 21: Análise da sintenia entre as regiões que codificam genes relacionados com o T6SS em <i>K. pneumoniae</i> Kp52.145 (Kp52), NTUH-K2044 (NT) e HS11286 (HS).....	70
Figura 22: Organização de genes de T6SS, sítios de início da transcrição nos promotores dependentes de σ^{70} e sítios putativos de ligação de reguladores transcricionais em genomas de <i>K. pneumoniae</i> Kp52.145 (a), HS11286 (b) e NTUH-K2044 (c).....	77
Figura 23: Sítios de início da transcrição dependentes de σ^{70} e locais putativos de ligação de reguladores transcricionais em genomas de <i>K. pneumoniae</i> Kp52.145 (a), HS11286 (b) e NTUH-K2044 (c).....	82
Figura 24: Logotipos dos sítios de ligação preditos para os reguladores transcricionais H-NS, Fis, OxyR e RcsAB nos promotores σ^{70} de <i>K. pneumoniae</i> Kp.52.145, HS11286 e NTUH-K2044.....	83
Figura 25: Sítios de início da transcrição dependentes de σ^{54} e locais putativos de ligação de reguladores transcricionais em genomas de <i>K. pneumoniae</i> Kp52.145 (a), HS11286 (b) e NTUH-K2044 (c).....	84
Figura 26: Panorama geral de estímulos específicos e reguladores transcricionais que podem regular a transcrição do T6SS em <i>K. pneumoniae</i>	86

LISTA DE TABELAS

Tabela 1: Componentes do T6SS	18
Tabela 2: Números de acesso dos cromossomos e plasmídeos de três cepas de <i>K. pneumoniae</i> analisadas neste estudo.....	43
Tabela 3: Genes que codificam componentes putativos do T6SS de Kp52.145.	56
Tabela 4: Genes que codificam componentes putativos do T6SS de HS11286.	58
Tabela 5: Genes que codificam componentes putativos do T6SS de NTUH-2044.	59
Tabela 6: Identidade e similaridade entre as ATPases ClpB, ClpA e ClpX identificadas como componente TssH no genoma de três <i>K. pneumoniae</i>	62
Tabela 7: Conteúdo de GC das regiões genômicas do T6SS em Kp52.145, HS11286 e NTUH-K2044.	72
Tabela 8: Reguladores transcricionais avaliados neste estudo e informações sobre as matrizes de peso-posição correspondentes.	75
Tabela 9: Características de reguladores transcricionais preditos como possíveis reguladores do T6SS de <i>K. pneumoniae</i>	80

Lista de Siglas e Abreviaturas

AAA+ – *ATPases Associated with various cellular Activities*

ABM – *Actin Binding Motif*

ACD – *Actin Cross-linking Domain*

bEBPs – *bacterial Enhancer Binding Proteins*

Clp – *Chaperone like protease*

CM – *Complexo de membrana*

COG – *Cluster of Orthologs Groups*

FHA – *ForkHead-Associated*

FPI – *Francisella pathogenicity island*

Fur – *Ferric-Uptake Regulator*

Hcp – *Hemolysin-coregulated protein*

HMM – *hidden Markov models*

LPS – *lipopolissacarídeos*

ME – *Membrana externa*

MI – *Membrana interna*

ORF – *Open Reading Frame*

PAAR – *Proline-Alanine-Alanine-arginine*

PC – *Parede Celular*

PDB – *Protein Data Bank*

PGB – *Peptidoglycan-Binding*

Pld – *Phospholipase*

PWM – *Position Weight Matrix*

QS – *quorum-sensing*

RNAP – *RNA polimerase*

SCI – *Salmonella Centrisome Island*

SecReT6 – *Type VI Secretion system Resource*

T6SS – *Type Six Secretion System*

Tag – *type six secretion associated genes*

Tle – *Type VI lipase effectors*

TPP – *Threonine Phosphorylation*

Tse1-3 – *Type six exported 1-3*

Tse2 – *Type six effector 2*

TssA-M – *Type Six Secretion System A-M*

VgrG – *Valine-glycine repeat G*

1 INTRODUÇÃO

1.1 Sistema de secreção do Tipo VI (T6SS)

O sistema de secreção do tipo VI (T6SS) foi inicialmente descrito em 2005 quando demonstrado que *Vibrio cholerae* O37 exibia citotoxicidade mediante contato com *Dictyostelium* sp ou macrófagos de mamíferos (Pukatzki et al. 2006). Nesse estudo, foi verificado que as proteínas Hcp (*Hemolysin-coregulated protein*) e VgrG (*Valine-glycine repeat G*) eram secretadas de forma independente da atividade dos mecanismos de secreção conhecidos. Um conjunto de genes em *V. cholerae*, que apresenta similaridade de sequência com genes que codificam as proteínas gp27 e gp5, que compõem a placa de base do bacteriófago T4, foi identificado. Estas similaridades serviram de evidências de que o T6SS de *V. cholerae* provavelmente possui uma estrutura similar ao complexo de proteínas da cauda do bacteriófago T4 de *Escherichia coli*. Através de ensaios de secreção e detecção de proteínas por *western blot*, bem como ensaios de ligação cruzada, foi demonstrado que os produtos destes genes eram responsáveis por translocar as proteínas VgrG e Hcp do interior da bactéria para o meio extracelular (Pukatzki et al. 2006, 2007).

Diversos outros estudos dessa época mostraram o envolvimento deste sistema de secreção na virulência e patogênese bacteriana (Mougous et al. 2006; de Bruin et al. 2007; Nano e Schmerk 2007; Shalom et al. 2007; Bingle et al. 2008). Em *Pseudomonas aeruginosa*, foi sugerido que o T6SS é ativo em infecções crônicas e contribui para a patogenicidade em pacientes com fibrose cística (Mougous et al. 2006). A ilha de patogenicidade de *Francisella* sp (FPI, *Francisella pathogenicity island*) codifica proteínas que compõem o T6SS (de Bruin et al. 2007; Nano e Schmerk 2007). Em *Burkholderia pseudomallei*, foi observado que a expressão de alguns genes putativos do T6SS é induzida pelo contato com macrófagos (Shalom et al. 2007). Genes do T6SS foram encontrados em cerca de 25% das bactérias Gram-negativas cujos genomas tinham sido sequenciados (Bingle et al. 2008). Desde então, cada vez mais estudos relacionam a função do T6SS às interações ambiente-bactéria-hospedeiro e elucidam a arquitetura do complexo e estrutura tridimensional de seus componentes.

Um sistema T6SS funcional é composto por produtos de pelo menos 13 genes conservados (Boyer et al. 2009). Conforme foram sendo descobertos e caracterizados, esses genes receberam nomes, por vezes relacionados a alguma característica de sua sequência primária, em outros casos relacionados a um fenótipo associado. Com o tempo, percebeu-se que genes homólogos de outras espécies

(ortólogos) tinham recebido nomes diferentes – dificultando a comparação entre eles e a caracterização geral do sistema. Assim, em 2007, foi proposta uma padronização da nomenclatura dos genes do T6SS, utilizando a nomenclatura *tss* (*type six secretion*), sendo as 13 famílias de genes denominadas de *tssA* a *tssM* (Tabela 1) (Shalom et al. 2007). Esses genes codificam um complexo macromolecular que atravessa as membranas interna (MI) e externa (ME) da bactéria, formando uma estrutura que lembra uma “seringa” - composta por uma “cauda” que se estende na região citoplasmática e uma “agulha” que atravessa as membranas, alcançando o meio extracelular (Ballister et al. 2008; Leiman et al. 2009). Além desses 13 componentes, outros já foram descritos e participam principalmente regulando a expressão e atividade do sistema (Tabela 1).

Tabela 1: Componentes do T6SS (Adaptado de (Shalom et al. 2007; Boyer et al. 2009; Leiman et al. 2009; Shneider et al. 2013; Weber et al. 2016; Zoued et al. 2017; Santin et al. 2018)).

Nomenclatura padronizada	Nome usual	Função	COG
TssA	SciA	Coordenação da montagem do sistema	COG3515
TssB	VipA	Componente da bainha da cauda	COG3516
TssC	VipB	Componente da bainha da cauda	COG3517
TssD	Hcp	Componente do tubo	COG3157
TssE	SciD	Complexo da base	COG3518
TssF	SciC	Complexo da base	COG3519
TssG	SciB	Complexo da base	COG3520
TssH	ClpV	ATPase / reciclagem da bainha	COG0542
TssI	VgrG	Ponta do sistema	COG3501
TssJ	SciN	Complexo de membrana	COG3521
TssK	SciO	Complexo da base	COG3522
TssL	MotB	Complexo de membrana	COG3455
TssM	IcmF	Complexo de membrana	COG3523
PAAR	-	Pode se associar a ponta do sistema e ter função efetora ou carrear algum efetor	COG4104
TagA	-	Interrupção biogênese do tubo/cauda	-
TagB	-	-	COG1357
TagC	-	-	-
TagD	-	PAAR-like	-
TagE	PpkA	Regulador pós-transcricional	COG0515
TagF	SciT	Regulador pós-transcricional	COG3913
TagG	-	Regulador pós-transcricional	COG0631
TagH	Fha I	Regulador pós-transcricional	COG3456
TagI	-	Lipoproteína ancorada à ME	-
TagJ	SciE	Incorporação de TssB ao sistema	COG4455
TagK	-	-	-
TagL	SciZ	Ancorar CM à parede celular e MI	COG2885
TagX	-	transglicosilase lítica	

Legenda: Esta tabela lista os componentes do T6SS, incluindo os 13 componentes conservados e proteínas acessórias. Ainda, esta tabela associa a nomenclatura padronizada Tss a outros nomes usuais de cada componente, bem como a localização ou função de cada componente no complexo e domínios funcionais de acordo com o COG: *Clusters of Orthologous Groups*. TssA-M – *Type Six Secretion System A-M*; PAAR - *Proline-Alanine-alanine-Arginine Repeats*; MI – Membrana interna; ME – Membrana externa; PC – Parede Celular.

Devido à similaridade entre as sequências desses componentes com proteínas de bacteriófagos, foi sugerido que o T6SS poderia formar um complexo similar a um bacteriófago (Ballister et al. 2008; Leiman et al. 2009). Bacteriófagos possuem uma estrutura conhecida como “cabeça” e uma “cauda” e interagem com a membrana externa bacteriana para injetar seu DNA no interior da bactéria. Por outro lado, o T6SS transloca proteínas para fora da bactéria. Assim, foi sugerido que o T6SS seria como um bacteriófago invertido: ancorado na membrana interna bacteriana por um complexo de membrana e teria uma “cauda” polimerizada no citoplasma bacteriano (Figura 1). Mediante um estímulo, essa “cauda” poderia injetar proteínas no meio extracelular ou numa outra célula alvo (translocação). Dessa forma, o T6SS agiria da mesma forma que um bacteriófago, mas a translocação seria no sentido oposto. Estudos de biologia estrutural têm demonstrado que o T6SS, de fato, é um complexo macromolecular, cujo arranjo tridimensional lembra uma “seringa” com agulha (Figura 1). Este sistema é compreendido pelo complexo de membrana (em inglês: *membrane complex*), ligado ao complexo da base (em inglês: *baseplate complex*;) e uma longa cauda citoplasmática tubular (em inglês: *tail components*) (Brunet et al. 2015b; Weber et al. 2016; Zoued et al. 2017).

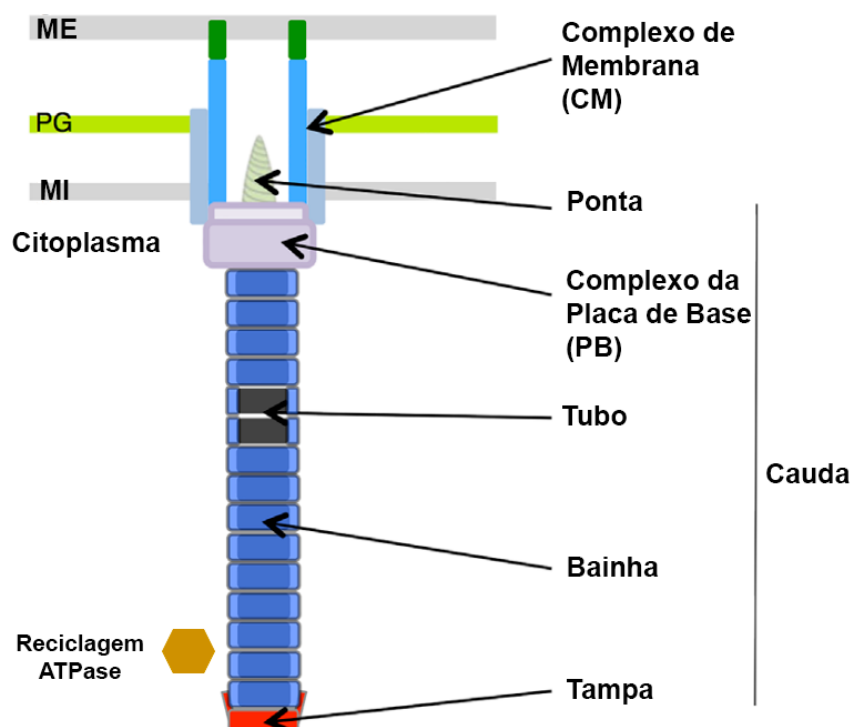


Figura 1: Arquitetura do T6SS. Representação esquemática da estrutura do T6SS. ME: Membrana Externa. MI: Membrana Interna. PG: Peptidoglicano Adaptado de (Zoued et al. 2017).

Este complexo de secreção teve sua biogênese revisada por Zoued et al. 2017, que descreveram que sua montagem acontece da membrana externa para o citoplasma. Inicialmente o complexo de membrana (componentes TssJ, TssM, TssL) é organizado através das membranas externa e interna. A ponta (*spike*: TssI e PAAR) e o complexo da base (TssE, TssF, TssG, TssK) são recrutados e interagem com o complexo de membrana. Em seguida, a cauda citoplasmática é polimerizada, sendo constituída pelo tubo interno (*tube*: TssD) e bainha (*sheath*: TssB e TssC). A proteína TssA (*cap* ou tampa) é essencial para o recrutamento do complexo da base para o complexo de membrana, bem como para a polimerização da cauda a partir do complexo da base e coordenação entre a bainha e o tubo. Proteínas efetoras podem ser carregadas por TssD no interior do tubo central ou por TssI e PAAR na ponta do sistema. Mediante um estímulo, ocorrem mudanças conformacionais na cauda, que levam a translocação das proteínas TssD, TssI, PAAR e dos efetores, para o meio extracelular ou para o interior de uma célula alvo. Essas proteínas efetoras geralmente possuem atividade relacionada com a modulação do metabolismo ou vias de sinalização da célula alvo, de forma a favorecer a sobrevivência da bactéria. Após todo este ciclo de montagem e translocação do T6SS, a bainha é desmontada através da ação de uma ATPase (TssH) e suas proteínas recicladas para que uma nova cauda seja polimerizada (Zoued et al. 2017).

1.2 Componentes do T6SS

1.2.1 Complexo de Membrana - CM (TssL, TssM e TssJ)

O complexo de membrana do T6SS é formado pelas proteínas TssJ, TssL e TssM (Durand et al. 2015). A TssJ é uma lipoproteína de membrana externa, enquanto TssL e TssM são proteínas ancoradas na membrana interna. A proteína TssJ é caracterizada pela presença do domínio T6SS-SciN e pertence ao COG3521. A proteína TssL possui o domínio N-terminal DotU e COG3455; e a proteína TssM o domínio lcmF e COG3523.

Dados recentes mostraram que a proteína TssJ é a que primeiro se posiciona na membrana externa, recrutando a seguir TssM e TssL (Rapisarda et al. 2019). Uma vez formado, este complexo de membrana corresponde a uma estrutura de aproximadamente 1,7 MDa, sendo composto por cinco dímeros heterotriméricos, além de cinco subunidades adicionais de TssJ (Durand et al. 2015; Rapisarda et al. 2019). A arquitetura deste complexo consiste em: 1) no *core*, uma estrutura composta pelos heterotrímeros, em forma de cinco duplos pilares que atravessam a região

periplasmática e atravessam a membrana interna; 2) uma base, composta pelos domínios citoplasmáticos de TssL e TssM; e 3) uma ponta, que contém as cinco subunidades adicionais de TssJ (Figura 2).

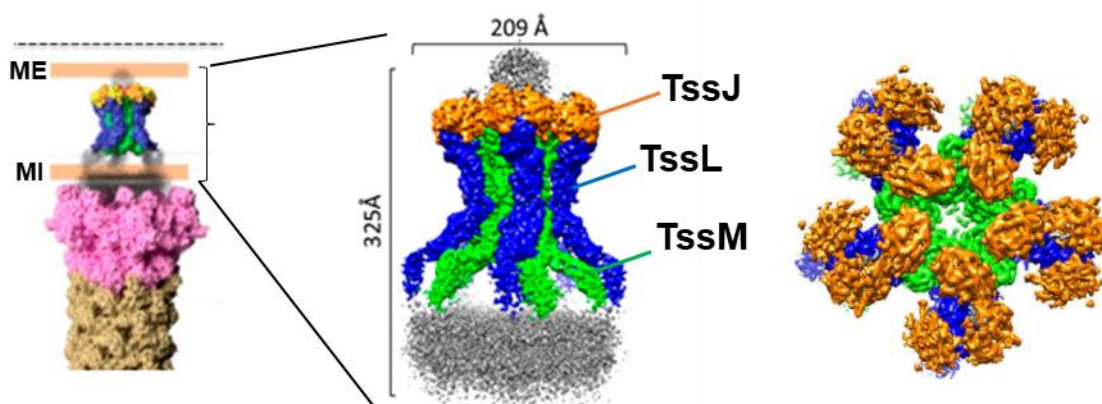


Figura 2: Estrutura do complexo de membrana do T6SS de *E. coli*. No centro, visão global da densidade eletrônica do T6SS. À esquerda está representada a posição do complexo de membrana em relação às membranas bacterianas, externa (ME) e interna (MI). À direita, são representadas as visões lateral e superior da região do *core*, em cores: pilares internos em verde, pilares externos em azul e parte superior em laranja. A ponta não estruturada e a base estão em cinza. Figura adaptada de (Rapisarda et al. 2019).

A região da base do complexo de membrana possui grande flexibilidade, o que propicia o posicionamento do *baseplate* (TssKFG-E-TssI) (Rapisarda et al. 2019). Além de servir como ancoramento do sistema de secreção, este complexo também funciona como um portão que permite a secreção pelo T6SS, mas impede extravasamentos do periplasma e a entrada de produtos tóxicos para dentro da bactéria (Durand et al. 2015; Rapisarda et al. 2019). O complexo de membrana é montado numa conformação onde os pilares ficam orientados para o centro do complexo, fechando-o na altura da membrana externa. Dessa forma, a célula fica protegida contra a liberação ou entrada de conteúdos indevidos.

Em algumas espécies, como *E. coli*, *K. pneumoniae* e *Y. pestis*, a proteína acessória TagL, ancorada à membrana interna, se associa ao complexo de membrana, interagindo com TssL (Aschtgen et al. 2010; Santin et al. 2019). Após a montagem do complexo de membrana, TagL interage com TssL. Em seguida, o complexo *baseplate* é recrutado, para dar continuidade ao processo de montagem do T6SS, polimerizando a bainha do complexo (Santin et al. 2019).

1.2.2 Ponta do sistema - TssI e PAAR

A ponta do T6SS é uma estrutura formada por um trímero da proteína TssI associada a proteína PAAR. A proteína TssI geralmente tem aproximadamente 850

aminoácidos, apresenta os domínios conservados VgrG, Phage_GPD, RHS e DUF2345 e pertence ao COG3501. Algumas TssI apresentam uma extensão na região C-terminal, de sequência variável e que pode conter um domínio extra, relacionado a sua função efetora nas células alvo ou à interação com efetores. Em *V. cholerae*, por exemplo, foram identificadas três proteínas TssI que apresentam extensões C-terminais com atividades enzimáticas distintas: TssI-1, TssI-2 e TssI-3 (Pukatzki et al. 2006). Por exemplo, a proteína TssI-1 possui o domínio de *crosslinking* de actina (ACD – Actin Cross-linking Domain) contendo um motivo de ligação a actina (ABM – Actin Binding Motif) na porção C-terminal (Ma et al. 2009; Dutta et al. 2019). Quando *V. cholerae* é fagocitada por células hospedeiras, TssI-1 é translocada pelo T6SS para o citosol da célula alvo, levando a um desarranjo do citoesqueleto de actina, e assim evitando a fagocitose de outras bactérias por macrófagos (Ma et al. 2009; Dutta et al. 2019). Em *Aeromonas hydrophila*, a TssI-1 contém um domínio de ribosilação de ADP (Suarez et al. 2010a). Em *E. coli* CFT073, a TssI-1 possui o domínio VgrG, Phage_GPD, T6SS_Vgr e DUF2345, que apresenta função de carreadora de efetores (Leiman et al. 2009). Por outro lado, a proteína PAAR (Proline-Alanine-Alanine-aRginine), possui o domínio PAAR_1, pertence ao COG4104 e pode apresentar papel efetor ou de carrear efetores para translocação. A análise da estrutura deste complexo revelou que elas apresentam um formato pontiagudo (Figura 3).

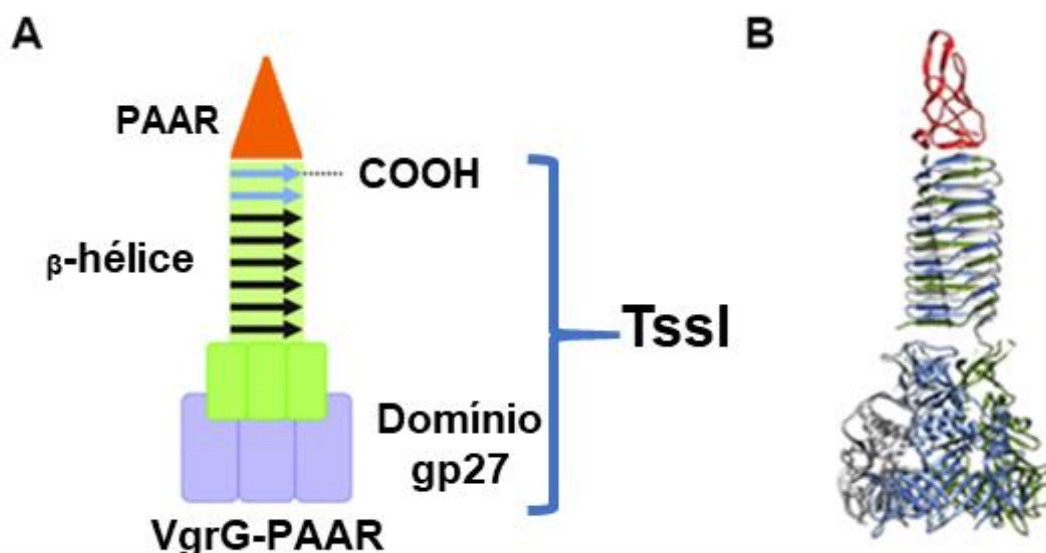


Figura 3: Complexo da ponta do T6SS, composto pelas proteínas TssI e PAAR. A – Representação esquemática identificando os domínios da TssI: em verde domínio gp5, em roxo o domínio gp27. A proteína PAAR está representada em laranja. **B –** Estrutura em fita do complexo TssI-PAAR de *E. coli* O6. TssI, ID PDB: 2P57; PAAR, ID PDB: 4JIW. Adaptado de (Shneider et al. 2013; Zoued et al. 2014).

Essa ponta em forma de triângulo corresponde a estrutura formada pela região C-terminal da TssI, que interage por ligações de hidrogênio com a proteína PAAR (Leiman et al. 2009; Shneider et al. 2013). Ambas as proteínas da ponta do T6SS são translocadas para a célula alvo. A importância de PAAR para a atividade antibacteriana do T6SS de *V. cholerae* e *Acinetobacter baylyi* foi demonstrada quando verificado que mutantes *paar* apresentam diminuição de 100 a 10.000 vezes na capacidade de matar *E. coli* (Shneider et al. 2013). O mesmo estudo também demonstrou que mutações em *paar* reduzem de 70% a 90% a secreção de TssD. A proteína PAAR pode ainda carrear efetores, como as proteínas Tde1 e Tde2 em *Agrobacterium tumefaciens* (Ma et al. 2014; Bondage et al. 2016).

1.2.3 Complexo da base (*baseplate*: TssE, TssF, TssG e TssK)

O *baseplate* é formado pelas proteínas TssE, TssF, TssG e TssK. Alguns autores consideram ainda que as proteínas da ponta também fazem parte do *baseplate* (Brunet et al. 2015b). As subunidades do *baseplate* formam um complexo, que é recrutado e se encaixa no complexo de membrana (Brunet et al. 2015b). Essa estrutura funciona como uma plataforma para a montagem da cauda (Cherrak et al. 2018).

A proteína TssE, classificada no COG3518, é localizada no complexo da base, possui domínio gpW_gp25 e função de ligar-se à TssI. A proteína TssF, classificada no COG3519, se localiza entre as proteínas TssK e TssG e, possui domínio VI_chp_6

e função de auxiliar na polimerização da bainha. A proteína TssG, classificada no COG3520, é ligada à TssF, possui domínio VI_chp_1 e função de auxiliar na polimerização da bainha. A proteína TssK, classificada no COG3522, integra a membrana interna, possui domínio VI_chp_4 e função de manter a estabilidade do complexo da base antes, durante e após a contração do T6SS.

1.2.4 Tubo (TssD) e cauda (TssB/TssC)

O tubo e a cauda são estruturas longas, montadas no citoplasma bacteriano. O tubo é formado pela proteína TssD, que apesar da padronização da nomenclatura das Tss, ainda se utiliza o nome usual Hcp (*Hemolysin-coregulated pilus*). A proteína TssD forma hexâmeros em formato de anéis, que por sua vez se empilham formando um tubo, com diâmetro interno de aproximadamente 40Å. O revestimento externo do tubo é feito por dodecâmeros das proteínas TssB e TssC (anteriormente denominadas VipA e VipB, respectivamente), formando uma bainha ao redor do tubo (Figura 4).

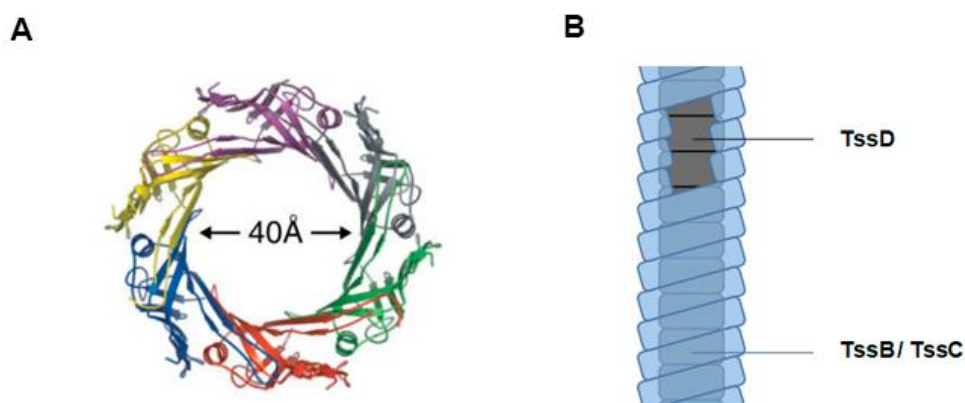


Figura 4: Cauda do T6SS. **A.** Vista superior de uma representação em fita do hexâmero TssD1 [Protein Data Bank (PDB) código ID 1Y12]. As subunidades individuais são coloridas em cores diferentes para destacar sua organização. A TssD1 forma um anel hexamérico com um diâmetro interno de aproximadamente 40Å. **B.** Representação do tubo (TssD em cinza) e da bainha (TssB/TssC em azul) do T6SS. Adaptado de (Mougous et al. 2006; Cherrak et al. 2019).

A proteína TssD apresenta domínio VI_effect_Hcp1 e pertence ao COG3157. Inicialmente foi imaginado que as proteínas efetoras do T6SS poderiam passar pela luz do tubo de TssD. Essa hipótese foi reforçada quando análises de microscopia eletrônica mostraram que a proteína efetora Tse2 (*Type six effector 2*) de *P. aeruginosa* interage com a superfície interna de TssD. Novos estudos encontraram hexâmeros de TssD contendo proteína efetora no seu interior, no interior da bactéria, antes mesmo

desses hexâmeros formarem o tubo (Blondel et al. 2009; Ma et al. 2017). Assim, hoje acredita-se que um dos mecanismos de translocação do T6SS, seja mediado por TssD, que carrega efetores (Mougous et al. 2006; Ballister et al. 2008; Silverman et al. 2013). Ainda, considerando que a proteína TssD também é translocada para a célula alvo ou meio extracelular, sua presença em sobrenadantes de cultura é usada como um marcador da atividade do T6SS (Zhou et al. 2012). TssD também possui função efetora, com múltiplas funções em diferentes bactérias (Ma et al. 2017). Foi demonstrado que a TssD de *Aeromonas hydrophila* ATCC 7966 pode impedir a maturação de macrófagos, inibindo o processo de fagocitose (Suarez et al. 2010b). Este processo ocorre com a inibição da produção de citocinas pró-inflamatórias e indução de interleucina-10, o que pode impedir a maturação dos macrófagos (Suarez et al. 2010a). Proteínas TssD de *E. coli* RS218 desempenham papéis na aderência e invasão celular (Zhou et al. 2012). Essas proteínas podem induzir rearranjo de citoesqueleto de actina, apoptose e liberação de interleucina-6 (IL-6) e interleucina-8 (IL-8) de células endoteliais microvasculares do cérebro humano (Zhou et al. 2012). Em *V. cholerae* V52, a deleção do gene *tssD* torna a bactéria avirulenta para *Dictyostelium* sp (Pukatzki et al. 2006).

As proteínas da bainha, TssB e TssC contêm o domínios VI_chp_5 e VI_chp_2 e pertencem aos COGs 3516 e 3517, respectivamente (Bönemann et al. 2009; Basler e Mekalanos 2012; Brunet et al. 2014). Apesar da semelhança funcional entre caudas de bacteriófagos contráteis e o T6SS, TssB e TssC não são homólogas com a proteína gp18 da bainha da cauda do bacteriófago T4 (Kube et al. 2014). Utilizando técnicas de crio-eletro tomografia de células bacterianas inteiras, foi mostrado que a estrutura formada pelas proteínas TssB e TssC apresenta ao menos duas conformações: estendida e contraída (Bönemann et al. 2009; Basler et al. 2012). A bainha é montada ao redor do tubo, na conformação estendida. Mediante estímulo, uma mudança conformacional na estrutura do complexo TssB/TssC para a conformação contraída, impulsiona o tubo formado por TssD em direção ao meio extracelular (Kube et al. 2014).

1.2.5 TssA

A coordenação da biogênese do T6SS, desde o acoplamento entre o complexo de membrana e o *baseplate*, a polimerização da cauda a partir do *baseplate* e a extensão do tubo e seu revestimento com a bainha, é mediado pela proteína TssA em algumas espécies (Brunet et al. 2015b).

TssA é uma proteína com o domínio ImpA-N no N-terminal, classificada no COG3515. Em *E. coli*, *V. cholerae* e *A. tumefaciens*, TssA possui ainda um domínio T6SS_VasJ na região C-terminal. Por outro lado, em *P. aeruginosa* uma proteína TssA foi identificada com a presença de um domínio com homologia a gp6 (subunidade do *baseplate* de bacteriófagos), sendo denominada TssA1 (Zoued et al. 2017).

Diferente da TssA, a TssA1 interage com componentes do complexo de base, interagindo com TssK. Portanto, TssA e TssA1 não compartilham propriedades funcionais. Ainda, alguns T6SS possuem proteína com domínio ImpA na porção N-terminal e uma região hidrofóbica (HR) que pode servir como âncora de membrana além de um domínio de função desconhecida VasL no C-terminal (Zoued et al. 2017). Nesse caso, a proteína é denominada TagA (Figura 5) (Zoued et al. 2017; Santin et al. 2018). Quando a extremidade distal da bainha do T6SS se aproxima da membrana oposta, a TagA é capturada e interrompe a polimerização da bainha, prendendo a bainha estendida à membrana como uma rolha (Santin et al. 2018).

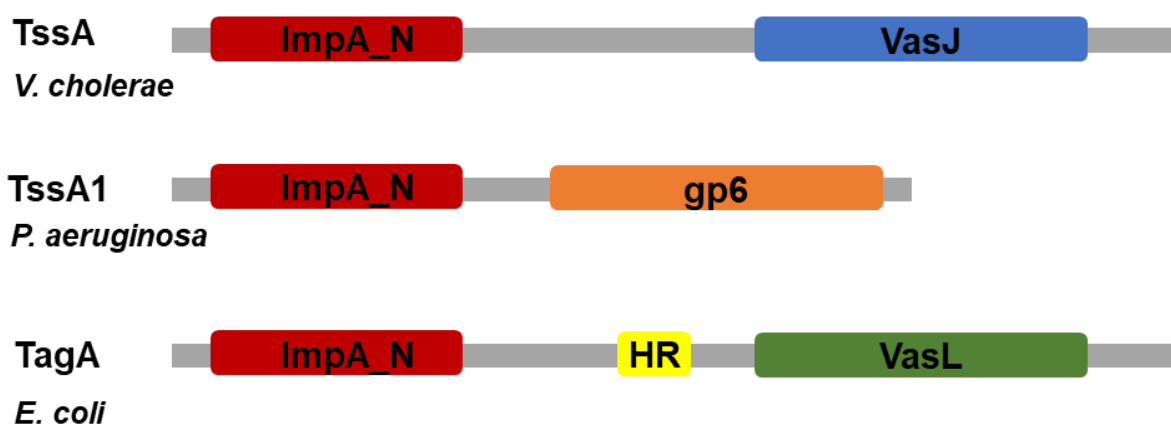


Figura 5: Arquitetura dos domínios das proteínas TssA, TssA1 e TagA. O domínio ImpA_N conservado (acesso: PF06812) é representado em vermelho, enquanto os domínios do C-terminal são representados em outras cores: T6SS_VasJ, domínio de função desconhecida (acesso: PF16989) em azul; gp6, região de homologia do bacteriófago gp6 em laranja; HR, região hidrofóbica em amarelo; VasL, domínio de função desconhecida (acesso PF12486) em verde. Adaptado de (Zoued et al. 2017).

1.2.6 Reciclagem – TssH

Após a montagem do T6SS e a contração da bainha que propulsiona o tubo para fora da bactéria, as proteínas TssD, TssI, PAAR e efetores são liberados no meio extracelular ou interior da célula alvo. Por outro lado, no citoplasma da bactéria, as proteínas da bainha TssB e TssC são desmontadas pela ação de uma TssH (Clp ATPase) – de forma a serem recicladas e usadas para a montagem de um novo T6SS (Bönemann et al. 2009). Proteínas Clp (*Chaperone like protease*) funcionam como subunidades reguladoras para selecionar e ligar substratos, desnovelá-los e

translocá-los para uma câmara proteolítica (Sauer et al. 2004). A proteína TssH possui um ou dois domínios AAA. O domínio AAA é definido por uma sequência conservada de 200-250 aminoácidos, que inclui as sequências características de motivos *Walker A / P-loop* e *Walker B*, responsáveis por transformar a energia da hidrólise de ATP em uma força mecânica para remodelar substratos (Figura 6) (Hanson e Whiteheart 2005; Schlieker et al. 2005). TssH tem como substratos as proteínas (TssB / TssC) da bainha do T6SS. Sabe-se que a porção N-terminal da TssH interage com a proteína TssC, dentro de um sulco hidrofóbico (Bönemann et al. 2009; Pietrosiuk et al. 2011).



Figura 6: Arquitetura dos domínios da proteína ClpV. Os domínios Clp_N (acesso: PF02861.20) e ClpB_D2 (acesso: PF10431.9) são representados em verde e marrom, enquanto os domínios AAA (acesso: PF00004.29) e AAA_2 (acesso: PF07724.14) são representados em cores vermelho e azul, respectivamente. As setas preta e azul representam os domínios *Walker A* e *Walker B*, respectivamente.

O T6SS é montado e, mediante um estímulo, é disparado, secretando as proteínas efetoras. Em seguida, as proteínas citoplasmáticas são reaproveitadas para a montagem de um novo complexo, que vai translocar mais efetores. Esse ciclo, se repete enquanto durar o estímulo. Considerando o papel de TssH na reciclagem de componentes do T6SS, podemos dizer que TssH é importante para a atividade do sistema, mas não para sua montagem (Bönemann et al. 2009; Basler e Mekalanos 2012; Kapitein et al. 2013). De fato, foi demonstrado que a deleção do gene *tssH-1* em *P. aeruginosa*, anula a secreção do tubo de TssD-1 (Mougous et al. 2006), o que reforça a importância da ATPase para o funcionamento do complexo de secreção. A atividade hidrolítica da TssH sobre a molécula de adenosina trifosfato (ATP) foi determinada inativando a atividade do domínio AAA-1, introduzindo uma mutação no motivo *Walker B*. Com essa mutação, a proteína TssH não funciona como fonte de energia para facilitar a secreção de TssD-1 (Mougous et al. 2006). De forma similar, em *V. cholerae*, a TssH foi descrita como necessária para a secreção de TssD e TssI-2 (Bönemann et al. 2009).

Curiosamente, bactérias como *Campylobacter jejuni* e *Fransicella tularensis*, apresentam um T6SS funcional, porém não tem homólogos de TssH, sugerindo que estas bactérias utilizam outro componente com atividade ATPase (Shrivastava e Mande 2008; Barker et al. 2009; Lertpiriyapong et al. 2012). No genoma destes organismos deficitários em TssH, foram identificados genes *clpB* (Brodmann et al. 2017). Assim como a TssH, ClpB é uma chaperona molecular hexamérica, com

atividade de ATPase (Shrivastava e Mande 2008). Foi demonstrado que ClpB pode cumprir o mesmo papel que TssH na reciclagem de componentes citoplasmáticos do T6SS (Shrivastava e Mande 2008; Barker et al. 2009; Lertpiriyapong et al. 2012). TssH e ClpB compartilham o COG0542 e os domínios AAA, AAA2 e ClpB_D2-small (Brodmann et al. 2017).

1.3 Componentes acessórios

Análises *in silico* demonstraram a presença de genes do T6SS em bactérias Gram-negativas, com o número de ortólogos variando em cada bactéria (Shrivastava e Mande 2008; Li et al. 2015). Além disso, os genomas de algumas bactérias podem codificar o conjunto de genes do T6SS em mais de um *locus*. Comumente, esses *loci* estão dentro de ilhas de patogenicidade, por exemplo a *Pseudomonas aeruginosa* HSI (Hcp1 Secretion Island) e *Salmonella typhimurium* SCI (Salmonella Centrisome Island), bem como nos genomas de *E. coli* enteroagregativa e *V. cholerae* (Cascales 2008; Chen et al. 2015; Lin et al. 2016). Além dos genes que codificam os 13 componentes mais conservados do T6SS, esses *loci* genômicos podem codificar toxinas, antitoxinas, adaptadores e proteínas auxiliares, bem como efetores adicionais (Journet e Cascales 2016).

Alguns desses componentes adicionais, tem sido denominados Tag (*type six secretion associated genes*) (Shalom et al. 2007). É possível que algumas destas proteínas possam ser secretadas pelo T6SS ou que sejam ancoradas ao envelope celular devido à presença de domínios transmembranares. Essas proteínas acessórias, podem funcionar também como locais de modificação lipídica, na ancoragem à membrana externa ou apresentar domínios de ligação a peptidoglicano semelhantes a MotB / OmpA (Shalom et al. 2007).

Em *P. aeruginosa* coexistem três T6SS, denominados T6SS - H1 a H3. A proteína TagJ é única para o T6SS-H1, portanto ausente nos T6SSs H2 e H3. Essa proteína TagJ interage com uma α -hélice na porção N-terminal de TssB, formando um subcomplexo do T6SS e modulando a incorporação da TssB ao T6SS (Lossi et al. 2012; Förster et al. 2014). Em *P. aeruginosa*, o T6SS é ativado após a fosforilação da proteína TagH, uma proteína com domínio ForkHead-Associated (FHA) e por TagE, uma treonina quinase que atravessa a membrana (Hsu et al. 2009). Com papel contrário ao da TagE, as proteínas TagF e TagG agem inibindo a ativação do T6SS em vias dependente e independente da fosforilação da treonina, ao formar um complexo com a proteína TagH (Lin et al. 2018a).

Outro componente acessório identificado em *Acinetobacter baumannii*, é um componente estrutural associado à membrana, chamado TagX. Esta proteína foi caracterizada como uma hidrolase de peptidoglicano com atividade de endopeptidase. A TagX cumpre o papel de permitir o trânsito dos componentes do T6SS através da camada de peptidoglicano da bactéria produtora de T6SS (Weber et al. 2016).

1.4 Efetores e proteínas de imunidade do T6SS

Conforme mencionado anteriormente, os sistemas de secreção podem ser usados em vários estágios da vida de uma bactéria, incluindo funções relacionadas a facilitar sua interação e sobrevivência em células hospedeiras eucarióticas, bem como conferir vantagens em processos de competição com outras bactérias (Kanamaru 2009). Considerando o amplo espectro de células alvo do T6SS, uma grande diversidade de efetores e toxinas estão associados com o T6SS, tornando-o uma arma versátil (Journet e Cascales 2016).

Para proteger a si e às células irmãs da intoxicação por efetores secretados, os organismos cujo T6SS apresenta atividade antibacteriana, expressam também proteínas de imunidade cognatas, que neutralizam a toxicidade de tais efetores (Alcoforado Diniz et al. 2015). Portanto, essas proteínas de imunidade impedem a ação de uma proteína efetora específica contra a própria bactéria que a produz. Nos *loci* do T6SS em que estes pares são descritos, frequentemente os genes que codificam as proteínas efetoras e as proteínas de imunidade são adjacentes. Por exemplo, em *E. coli* os genes *tle1* e *tli1*, que codificam as proteínas Tle1 (efetora) e Tli1 (imunidade) são vizinhos (Figura 7) (Alcoforado Diniz et al. 2015; Yang et al. 2018), assim como em *K. pneumoniae* HS11286, as proteínas Tle1^{KP} e Tli1^{KP} são adjacentes (Liu et al. 2017). Estes pares de Efetores-Imunizadores (imunizantes) podem ser identificados por meio de análises computacionais e de secretomas (Hood et al. 2010; Russell et al. 2012; Dong et al. 2013; English et al. 2014).

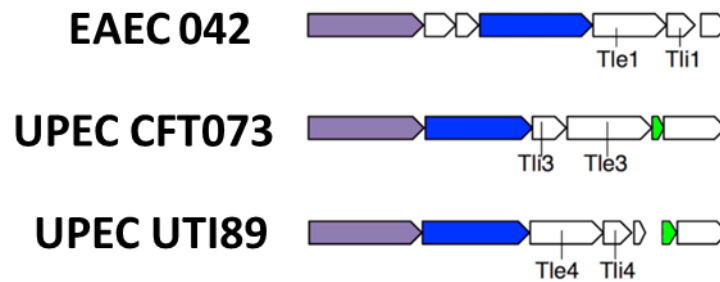


Figura 7: Co-localização de genes que codificam pares de proteína efetor-imunidade em *E. coli* EAEC 042, UPEC CFT073 e UPEC UTI89. Nas cores magenta, azul e verde estão representados os genes que codificam as proteínas TssH, TssI e PAAR. Os genes que codificam proteínas efetoras e de imunidade estão indicados com seus nomes. As setas brancas sem indicação representam genes com função desconhecida. Tle: *type VI lipase effector*, Tli: *type VI lipase immunity* (Journet e Cascales 2016).

1.5 O papel do T6SS nas interações bactéria-bactéria

Já foram descritas proteínas efetoras do T6SS, com atividade antibacteriana, com atividades de amidase, muramidase de peptidoglicano e fosfolipase (Russell et al. 2011; Shang et al. 2012; Ma et al. 2017). Em *P. aeruginosa* foram descritas a presença de três proteínas exportadas pelo T6SS, denominadas Tse1-3 (*Type six exported* 1-3). Tse1 e Tse3 são enzimas periplasmáticas com atividade lítica, que degradam peptidoglicanos via atividades do tipo amidase e muramidase (Russell et al. 2011). A Tse2 codificada na ilha genômica T6SS-H1, demonstra capacidade de impedir o crescimento de *Burkholderia thailandensis* e *E. coli* (Hood et al. 2010).

Em *E. coli*, a proteína efetora Tle1 possui atividade de fosfolipase A1 e A2 (Flaugnatti et al. 2016). Tle1 é translocada pelo T6SS sendo carregada por TssI e tem como alvo lipídios da membrana interna. Em *V. cholerae* V52 a proteína codificada pelo gene *vasX*, homólogo da colicina, auxilia na morte de *E. coli* mediada por T6SS (Miyata et al. 2013). O efetor VasX, juntamente com Tse4 de *P. aeruginosa* tem a capacidade para formar poros ou canais na membrana de células bacterianas (Miyata et al. 2013; LaCourse et al. 2018). Em *K. pneumoniae* HS11286, por ensaios de competição intra e inter-específicos com expressão do gene efetor *tle1^{KP}* projetado para atingir o periplasma, foi descrito que o T6SS provoca retardo no crescimento de *E. coli* (Liu et al. 2017). Muitos outros efetores com atividade antibacteriana vem sendo descritos na literatura (Russell et al. 2014; Alcoforado Diniz et al. 2015; Trunk et al. 2019).

1.6 O papel do T6SS nas interações com células do hospedeiro

Diversos efetores associados ao T6SS possuem ação em células eucarióticas (Hachani et al. 2016). Abaixo são relatados apenas alguns desses efetores já estudados.

As células fagocíticas são determinantes na capacidade do hospedeiro para controlar infecções bacterianas. Uma estratégia empregada por bactérias invasoras para evitar o sistema imune de células hospedeiras, é direcionar sua captação para células hospedeiras que são fagócitos não-profissionais, como células epiteliais e endoteliais que revestem superfícies mucosas e vasos sanguíneos (Cossart 2004). Foi demonstrado que o T6SS-H2 de *P. aeruginosa* PAO1 pode agir facilitando a captação bacteriana por células HeLa e células epiteliais pulmonares, modulando a interação com células hospedeiras eucarióticas (Sana et al. 2012). Nesta mesma bactéria, a proteína Tssl2b é translocada pelo T6SS e interage com microtúbulos no complexo de anel de γ -tubulina (γ TuRC) de células eucarióticas. Essa interação precede a internalização da *P. aeruginosa* que é dependente dos microtúbulos de α e β - tubulina (Sana et al. 2015). Outros exemplos são os patógenos humano *B. pseudomallei* e *B. thailandensis*, que são bactérias intracelulares facultativas, com habilidade de escapar para o citosol de células hospedeiras. Estas espécies podem estimular, via T6SS-5, a formação de células gigantes multinucleadas facilitando a disseminação bacteriana, o acesso a nutrientes e o escape de mecanismos de defesa e de antibióticos (Whiteley et al. 2017).

Dois tipos de fosfolipases D secretadas pelo T6SS de *P. aeruginosa* foram identificadas como efetores de virulência, tendo como alvos células hospedeiras tanto procarióticas como eucarióticas. Cada uma dessas fosfolipases pode provocar a internalização bacteriana em células não fagocíticas (Bleves et al. 2014). Com atividade nas membranas eucarióticas, tem-se identificado em *V. cholerae*, *P. aeruginosa* e *K. pneumoniae* os efetores TseL, Tle5 e Tle1^{KP} respectivamente (Russell et al. 2012; Dong et al. 2013; Liu et al. 2017). A secreção de efetores pelo T6SS ressalta a diversidade de substratos secretados via T6SS e mecanismos diferentes ainda não elucidados completamente, que podem realizar diferentes funções importantes no estabelecimento de infecções em hospedeiros eucarióticos (Sana et al. 2012; Dong et al. 2013).

1.7 O papel do T6SS na captação de íons metálicos

Uma característica do T6SS é a presença de múltiplos sistemas em algumas espécies bacterianas (Boyer et al. 2009). Está bem estabelecido que alguns desses sistemas funcionam para competir por seus nichos entre diversas espécies bacterianas (Anderson et al. 2017) e células eucarióticas (Hachani et al. 2016), enquanto a atividade de muitos desses transportadores permanece desconhecida. Os íons metálicos são componentes essenciais para crescimento bacteriano, assim a aquisição de íons metálicos altera o metabolismo celular, participando do transporte de micronutrientes, da regulação transcricional e resistência a estresses oxidativos (Aguirre e Culotta 2012; Lisher e Giedroc 2013).

Em *Yersinia pseudotuberculosis* e *B. thailandensis* foi descrito um T6SS (T6SS-4 em ambas as espécies), que funciona para importar íons Zn^{2+} e Mn^{2+} respectivamente do meio extracelular, agindo como antioxidantes para diminuir efeitos de radicais hidroxila prejudiciais para a bactéria (Wang et al. 2015; Si et al. 2017a). Foi identificado que a expressão do T6SS em *Y. pseudotuberculosis* é ativada pelas proteínas reguladoras OxyR e OmpR (Zhang et al. 2013; Wang et al. 2015), que detectam estímulos ambientais relacionados ao estresse oxidativo (Zheng 1998) e alterações osmóticas (Gueguen et al. 2013).

O transporte de Zn^{2+} através da membrana externa pode ocorrer mediado por porinas inespecíficas ou por receptores dependentes de TonB. Ao alcançar o periplasma o íon atravessa a membrana citoplasmática por afinidade a transportadores (Si et al. 2017a). Outro mecanismo que facilita a absorção de íons metálicos durante a infecção é a presença de compostos quelantes similares aos sideróforos (Bobrov et al. 2014). Tanto em *B. thailandensis* quanto em *Y. pseudotuberculosis* o transporte dos íons é facilitado pela presença de proteínas quelantes, YezP em *Y. pseudotuberculosis* e TseM em *B. thailandensis* (Wang et al. 2015; Si et al. 2017a).

Em *P. aeruginosa*, após a observação de que um mutante sem os sistemas de captação de ferro mantem a capacidade de crescer em meios contendo quelantes do íon Fe^{3+} , foi sugerido uma via alternativa para o processo de captação deste íon (Lin et al. 2017). Diferente do relatado em *Y. pseudotuberculosis* (Wang et al. 2015) e *B. thailandensis* (Si et al. 2017a), onde uma proteína age como um quelante de íons, em *P. aeruginosa*, a proteína TseF é secretada pelo T6SS-H3 e incorporada a vesículas de membrana externa contendo Fe^{3+} . A proteína TseF facilita a captura do íon pelo recobrimento com os receptores de superfície celular FptA e OprF (Lin et al. 2017).

Estas funções de captura de íons metálicos demonstram a diversidade de funções que o T6SS bacteriano pode desempenhar, promovendo ganhos diversos às bactérias.

1.8 Genes órfãos do T6SS

Se por um lado, grande parte dos genes do T6SS encontram-se codificados em ilhas genômicas, por outro lado alguns genes do T6SS são encontrados fora dessas regiões. Esses genes do T6SS codificados em regiões distintas são comumente referidos como genes "órfãos" do T6SS (Shyntum et al. 2014). A expressão destes genes órfãos foi evidenciada em condições de infecção (Bent et al. 2016) e em isolados clínicos (Bruchmann et al. 2015). Portanto, a identificação desses genes em genomas bacterianos pode ser crucial para entender o papel do T6SS.

Por exemplo, na espécie *Pantoea ananatis*, o gene que codifica o componente TssD é localizado numa região que codifica genes com funções de um sistema toxina / antitoxina induzido pelo sistema SOS (Shyntum et al. 2014). Em *K. pneumoniae* ATCC BAA-2146 um dos genes que codifica Hcp é encontrado juntamente com genes que codificam um sistema de transporte de íons Fe (Bent et al. 2016).

1.9 Regulação do T6SS

O T6SS pode ser utilizado para mediar a competição interbacteriana, este sistema estaria adaptado às necessidades específicas de cada bactéria individualmente (Russell et al. 2012). Devido à demanda energética do ciclo de montagem, contração, desmontagem e remontagem do T6SS, foi hipotetizado que a célula bacteriana poderia apresentar uma diversidade de mecanismos regulatórios para a expressão dos componentes do T6SS. Desta forma, os *clusters* de genes que codificam estas máquinas devem ser regulados com precisão pelas funções do hospedeiro e / ou condições ambientais (Bernard et al. 2011).

Em *P. aeruginosa* as funções do T6SS podem ser reguladas nos níveis pós-transcricional e pós-traducional por diversos mecanismos. Pós-transcricionalmente pode ocorrer pela ligação da proteína RsmA ao mRNA, inibindo a tradução do RNA, numa via dependente do sensor de quinase RetS que regula fenótipos de virulência associados à infecção aguda em *P. aeruginosa* (Silverman et al. 2012). Pós-traducionalmente pode ocorrer pela via dependente de TPP (*Threonine Phosphorylation*) com fosforilação de TagE (uma serina-treonina quinase) pelo componente acessório TagR, regulando a secreção de TssD1 (Chen et al. 2015).

A definição dos sinais e condições que controlam a expressão e ativação do T6SS são variados e os ambientes que as bactérias ocupam pode ser fundamental para revelar seu papel nesses contextos (Silverman et al. 2012). As bactérias podem codificar várias 'cópias' funcionais de cada gene do T6SS. A expressão das diferentes isoformas de um componente pode variar de acordo com o alvo celular e/ou pela condição ambiental em que a bactéria se encontra (Journet e Cascales 2016). A expressão destes complexos de secreção pode ser regulada também, com precisão, por mecanismos transcricionais sensíveis a diversas alterações ambientais (Ishikawa et al. 2009). Tem sido reportado que essa expressão envolve uma variedade de sistemas regulatórios como sistemas de dois componentes (Lazzaro et al. 2017), sistemas *quorum-sensing* (QS) (Kitaoka et al. 2011; Salomon et al. 2013), formação de biofilme (Aubert et al. 2008; Moscoso et al. 2011), depleção de ferro (Brunet et al. 2011; Chakraborty et al. 2011), termorregulação (Pieper et al. 2009; Salomon et al. 2013) e salinidade (Salomon et al. 2013). Associados a estas condições descritas, fatores sigma alternativos e proteínas semelhantes a histonas, estão relacionados à expressão do T6SS (Bernard et al. 2011; Miyata et al. 2013). Exemplos da regulação do T6SS em bactérias estão ilustrados na Figura 8.

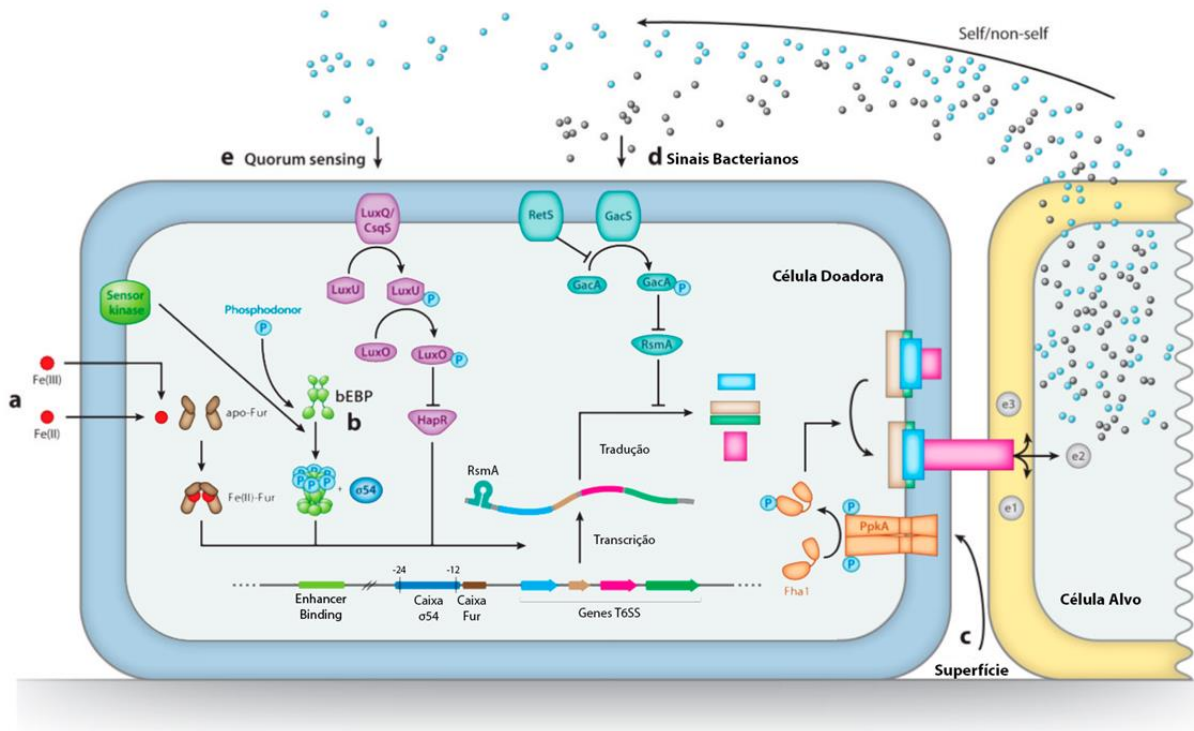


Figura 8: Representação esquemática dos diversos sistemas que modulam a expressão e ativação de T6SS em bactérias. A – O regulador transcricional Fur, reprime a transcrição de T6SS na presença de ferro; B – Função da *bEBP* (*bacterial Enhancer-Binding Proteins*) em conjunto com σ^{54} para ativar a transcrição de T6SS; C – A proteína TPP ativa pós-traducionalmente T6SS em resposta à associação de superfície; D – T6SS sendo modulado pós-transcricionalmente através da via Gac / Rsm; E – T6SS sendo transcrito via *Quorum Sensing*. À direita, uma bactéria alvo que está sendo intoxicada por efeitores Tse1-3 (e1-3). Adaptado de (Silverman et al. 2012).

A proteína reguladora de absorção férrica (Fur – *Ferric-Uptake Regulator*) é um modulador chave da expressão gênica em bactérias. Este regulador geralmente reprime a transcrição através de dimerização dependente de Fe (II) e subsequente ligação ao DNA a uma sequência consenso em regiões promotoras (Carpenter et al. 2009). Em *E. coli* EAEC a expressão do T6SS é reprimida transcricionalmente por Fur (Brunet et al. 2011) e em *Edwardsiella tarda* foi identificado que o regulador transcricional Fur pode reprimir a exportação de TssD (Chakraborty et al. 2011).

Outro mecanismo transcricional envolvido na ativação do T6SS em bactérias tem como base a densidade populacional celular, denominado de *quorum sensing* – QS (Ng e Bassler 2009). A regulação de genes controlados por QS é obtida por efeitos diretos ou indiretos que as moléculas sinalizadoras transmitem sobre as propriedades de ligação ao DNA de proteínas reguladoras dedicadas (Silverman et al. 2012).

Quando há baixa densidade populacional de células bacterianas em determinado ambiente, o regulador de resposta LuxO é fosforilado, o que ativa a expressão de pequenos RNAs reguladores específicos (sRNAs) em conjunto com o fator sigma σ^{54} , inibindo a expressão do sensor de quinase LuxR. Com ação contrária, em alta densidade populacional de células, o regulador LuxO é desfosforilado, desligando a transcrição dos sRNAs e permitindo a tradução de LuxR (Waters e Bassler 2005; Sheng et al. 2012).

Em *V. cholerae* o gene *vasH* codifica um ativador do fator σ^{54} . A proteína VasH e o fator σ^{54} atuam em conjunto para controlar a transcrição do gene *tssD* (Kitaoka et al. 2011). Em *V. cholerae* A1552 sorotipo O1, a expressão da proteína TssD é dependente da fase de crescimento em que a bactéria se encontra e pode envolver os reguladores de *quorum sensing* HapR, LuxR, RpoN e uma proteína regulatória global, CRP (Ishikawa et al. 2009). Em *P. aeruginosa* PAO1, o fator σ^{54} atua de maneira divergente nos T6SS-H2 e H3. Os dois operons putativos encontrados em T6SS-H3, chamados de "*left*" e "*right*", são co-regulados com T6SS-H2 pelos sistemas de *quorum sensing* (Sana et al. 2013). Nestes operons descritos, o fator σ^{54} participa da ativação da expressão de T6SS-H3 "*left*" e reprime T6SS-H3 "*right*" juntamente com T6SS-H2.

1.10 Regulação transcricional

Genes bacterianos podem ser conservados entre diferentes espécies. Contudo, a variedade de ambientes que as bactérias vivem, resulta em diferenças nos mecanismos de expressão gênica e nas quantidades dos produtos gerados. A regulação da expressão pode ocorrer pela ação de fatores de transcrição que regulam os promotores. As bactérias usam uma variedade de mecanismos para direcionar a RNA polimerase a promotores específicos, a fim de ativar a transcrição em resposta a sinais de crescimento ou sinais ambientais. A ativação pode ser devida a fatores que interagem em promotores específicos, aumentando assim a transcrição direcionada (Lee et al. 2012).

As bactérias contêm apenas uma forma da enzima RNA polimerase (RNAP), que é um complexo de cinco subunidades ($\alpha_2\beta\beta'\omega$). No entanto, existem múltiplas formas da subunidade do fator σ . Diferenças no promotor resultam em mudanças na disponibilidade de diferentes holoenzimas da RNAP, sob diferentes condições nutricionais ou ambientais (Haugen et al. 2008; Lee et al. 2012).

Para que o processo de transcrição tenha início é necessário o reconhecimento dos elementos promotores, localizados a montante do sítio de início do transcrito, pela holo RNA polimerase (Murakami e Darst 2003). Os elementos promotores podem variar, dependendo do tipo da subunidade σ , dando à bactéria capacidade de adaptar a expressão gênica em função das mudanças ambientais (Borukhov e Nudler 2003).

Nas bactérias um ou mais fatores σ controlam o início da transcrição de um subconjunto de genes com funções que variam entre as espécies. Por exemplo, no genoma de *Mycoplasma genitalium* é codificado 1 fator σ , em *E. coli* são codificados 7 fatores σ , em *B. subtilis* são 18 fatores σ e em *P. putida* são encontrados 24 fatores σ diferentes (Gruber e Gross 2003). Dentre os sete fatores sigmas encontrados em *E. coli*, destacamos duas classes que podem ser distinguidas pela sequência e modo de início de transcrição, os fatores σ^{70} e σ^{54} . O fator σ^{70} liga-se a sequências promotoras nas posições -35 e -10 a partir do sítio de início da transcrição, Figura 9A (Haugen et al. 2008).

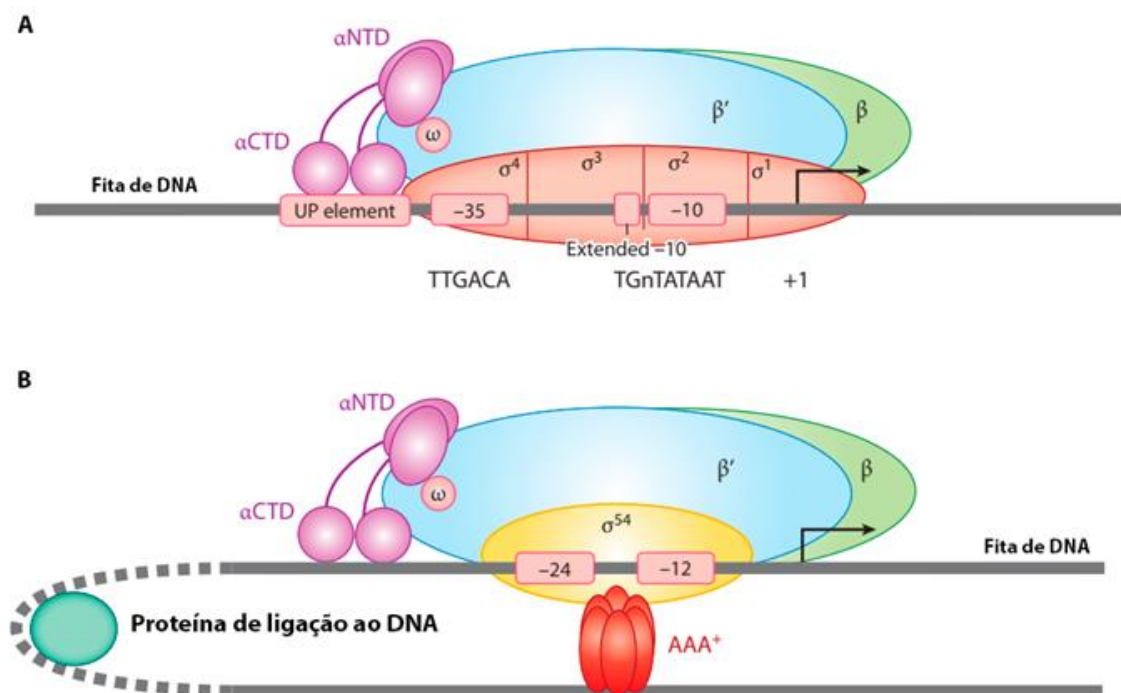


Figura 9: Interação da RNA polimerase com subunidades σ^{70} e σ^{54} . **A** – Representação esquemática da RNA polimerase ligada a um promotor. A localização do sítio de início de transcrição é indicada pela seta curva na posição +1. São evidenciados os sítios do elemento UP, -35, -10 e -10 estendido. **B** – RNA polimerase associada a σ^{54} liga-se aos elementos promotores -12 e -24. Ocorre interação com a proteína ativadora AAA+ (em vermelho), facilitada por uma volta na fita de DNA, provocado por uma proteína de flexão de DNA, representada em verde claro. NTD – Domínio N-terminal; CTD – Domínio C-terminal; UP – a montante. Adaptada de (Lee et al. 2012).

A regulação transcricional nos promotores σ^{70} , pode ocorrer por proteínas reguladoras que interagem com a RNAP e modulam a transcrição do promotor. Essas

proteínas reguladoras são chamadas de fatores transcricionais (Ishihama et al. 2016). Por outro lado, o fator σ^{54} se liga a sequências promotoras específicas nas posições -24 e -12 a partir do sítio de início da transcrição (Figura 9B), onde permanece em uma conformação fechada do DNA (Schumacher et al. 2006). Este fator necessita de proteínas ativadoras dependentes, denominadas de bEBPs – *bacterial Enhancer Binding Proteins*. Estas proteínas medeiam a tradução de sinais ambientais para mudanças na expressão gênica em bactérias (Rappas et al. 2007). Como representado na Figura 9, a reestruturação do complexo promotor σ^{54} -RNAP, de fechado para uma conformação aberta ou ativa depende ainda da hidrólise de ATP (Schumacher et al. 2006).

1.11 *Klebsiella pneumoniae*

Klebsiella pneumoniae é uma espécie presente na natureza, comensal do intestino e patógeno oportunista humano - causando principalmente infecções em indivíduos hospitalizados ou imunocomprometidos (Podschun e Ullmann 1998). Essa bactéria pode causar uma vasta gama de infecções, incluindo pneumonia, infecções do trato urinário, bacteremia e abscessos hepáticos (Yang et al. 2009). Devido à ocorrência frequente de múltiplos isolados resistentes a antibióticos, *K. pneumoniae* é considerada um problema de saúde pública global (Munoz-Price et al. 2013; Pereira et al. 2013).

Análises de centenas de genomas da espécie *K. pneumoniae* identificou aproximadamente 2.000 genes conservados, representando o que chamamos de *core* genoma (Holt et al. 2015; Martin e Bachman 2018). Por outro lado, o conjunto de genes que variam entre os isolados, é chamado de genoma acessório, (Martin e Bachman 2018). Na comparação entre genomas de cinco cepas de *K. pneumoniae*, foram identificados 743 genes acessórios na cepa Kp52.145 e 608 genes na cepa NTUH-K2044 (Lery et al. 2014). Genes responsáveis por processos específicos, como por exemplo a fixação de nitrogênio ou fatores de virulência fazem parte do genoma acessório (Fouts et al. 2008). A extensão do genoma acessório, e variações desse conjunto refletem a diversidade de ambientes e funções desempenhadas por bactérias dessa espécie, assim como a diversidade de fatores de virulência encontrados nas diferentes linhagens de *K. pneumoniae*. Os principais fatores de virulência bem estudados são a cápsula polissacarídica, fimbrias, adesinas, lipopolissacarídeos (LPS) e sideróforos (Figura 10).

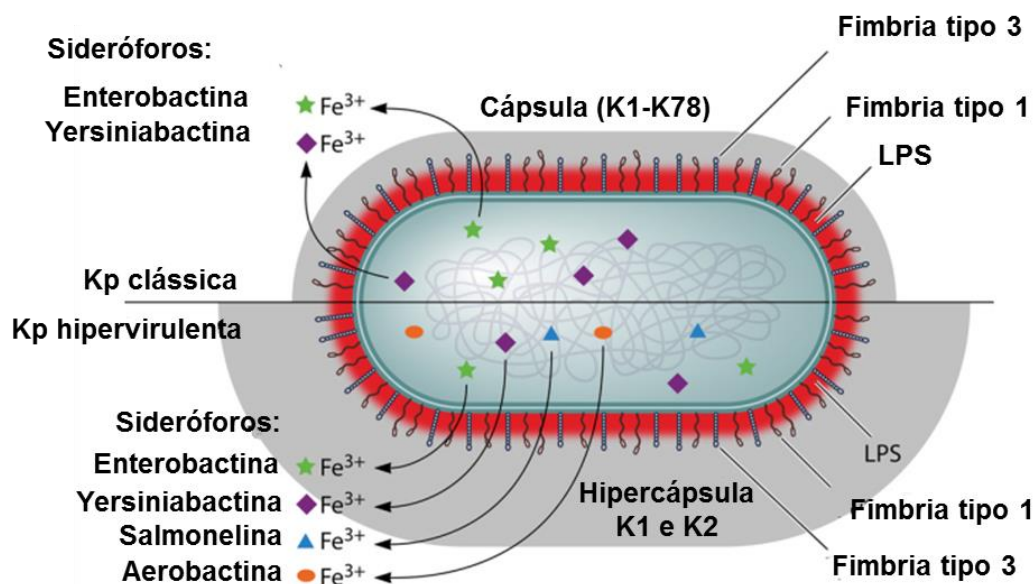


Figura 10: Representação de fatores de virulência em *K. pneumoniae*. A parte superior representa uma *K. pneumoniae* clássica com produção de cápsula modesta (em cinza), fímbrias do tipo 1 e 3, lipopolissacarídeos (em vermelho) e dois sideróforos (Enterobactin e Yersiniabactin). A parte inferior representa uma *K. pneumoniae* hipervirulenta, com hiperprodução de cápsula e sideróforos adicionais (*Salmochelina* e *Aerobactin*). Adaptado de (Paczosa e Mecsas 2016).

A heterogeneidade significativa entre as cepas de *K. pneumoniae* e os fatores de virulência que cada uma contém, é característica dessa espécie (Broberg et al. 2014; Paczosa e Mecsas 2016). Estes fatores podem determinar a gravidade de uma infecção e, portanto, a virulência da cepa infectante e a capacidade colonizadora de progredir para a infecção, definindo o potencial patogênico de uma determinada cepa (Holt et al. 2015). Recentemente, foram descritos fatores adicionais, como fosfolipases, sistemas de secreção do tipo II (T2SS) e tipo VI (T6SS) (Lery et al. 2014; Tomás et al. 2015; Liu et al. 2017).

Outros fatores foram recentemente identificados como importantes para a virulência de *K. pneumoniae*, incluindo proteínas de membrana externa, porinas, bombas de efluxo, sistemas de transporte de ferro e genes envolvidos no metabolismo da alantoína (Paczosa e Mecsas 2016). No entanto, esses fatores ainda não estão completamente caracterizados.

1.12 T6SS em *Klebsiella pneumoniae*

Pouco se sabe sobre o T6SS em *K. pneumoniae*. De uma perspectiva genômica, os genes que codificam os componentes putativos do T6SS estão presentes nos genomas de *K. pneumoniae* (Sarris et al. 2011; Lery et al. 2014; Liu et al. 2017). Em algumas cepas, os genes do T6SS são agrupados em dois *loci* (por

exemplo: NTUH-K2044, Kpn2146 e HS11286, enquanto em outras são encontrados em três *loci* (Kp52.145, MGH 78578, 342) (Sarris et al. 2011; Lery et al. 2014; Bruchmann et al. 2015; Liu et al. 2017). Como os genomas de *K. pneumoniae* foram anotados utilizando diferentes abordagens - alguns usando *pipelines* automáticas - não há anotação uniforme de genes do T6SS.

Em uma perspectiva funcional, a mutação de dois genes hipotéticos codificadores de proteínas de *K. pneumoniae* 43816 diminuiu a capacidade de infectar baço de rato (Lawlor et al. 2005). Essas proteínas foram anotadas como componentes putativos do T6SS. Recentemente foi descrito que a cepa HS11286, secreta um efetor de fosfolipase com atividade antibacteriana, via T6SS (Liu et al. 2017). Além disso, foi demonstrado que mutantes do T6SS em *K. pneumoniae* NTUH-K2044 reduziram significativamente a morte de *E. coli* e *S. thiphimurium* em ensaios de competição, reduzindo também a expressão de fímbrias tipo-1, a adesão e invasão de células epiteliais (Hsieh et al. 2018). No que se refere à regulação do T6SS, foi observado que genes de *K. pneumoniae* Kp52.145 foram expressos quando a bactéria colonizou pulmões de ratos, mas não quando a bactéria cresce no meio de cultura TCS (Lery et al. 2014). Em *K. pneumoniae* Kpn2146, análises de RNA-Seq revelaram que genes do T6SS tiveram sua expressão aumentada 24 horas após a infecção em macrófagos (Bent et al. 2016). Em *K. pneumoniae* HS11286 foi sugerido que concentrações sub-inibitórias de antibióticos podem regular a secreção de T6SS (Liu et al. 2017), similar ao que ocorre em *P. aeruginosa* (Jones et al. 2013). Na cepa NTUH-K2044, foi demonstrada que a proteína estruturante de nucleóide similar a histona (H-NS) liga-se à região reguladora e inibe a expressão do gene *tssD* (Hsieh et al. 2018).

Com base na evidência crescente para a relevância do T6SS para a patogênese de *K. pneumoniae* e divergentes características genômicas dos *loci* do T6SS entre as linhagens de *K. pneumoniae*, este trabalho objetivou identificar genes do T6SS em três cepas de *K. pneumoniae* e padronizar a anotação dos genes que constituem o T6SS. Além disso, a hipótese deste trabalho é que mecanismos transcricionais estejam envolvidos na modulação e ativação dos componentes do T6SS nestas três cepas. Para verificar esta hipótese, identificamos sítios de ligação de reguladores transcricionais a jusante dos sítios de início de transcrição dos genes do T6SS com o intuito de entender mais sobre a função do T6SS e sua regulação.

1.13 Justificativa

Diante do exposto acima, esse trabalho é importante pois hoje, há cepas de *K. pneumoniae* resistentes a todas as classes de antibióticos disponíveis, o que significa que infecções causadas por estas linhagens não são tratáveis. Assim, faz-se necessário compreender os mecanismos de virulência desta bactéria, contribuindo para a compreensão de um mecanismo envolvido na interação bactéria-hospedeiro para que no futuro, possam desenvolver novas formas de tratamento.

2 OBJETIVOS

2.1 Geral

Contribuir para a caracterização de mecanismos de regulação transcricional dos genes que codificam componentes do T6SS em *K. pneumoniae*.

2.2 Específicos

- 1) Identificar os genes que codificam o T6SS nas cepas de *K. pneumoniae* Kp52.145, HS11286 e NTUH-K2044;
- 2) Harmonizar a anotação dos genes que codificam componentes do T6SS conforme a nomenclatura padrão;
- 3) Avaliar a organização genômica destes genes;
- 4) Identificar sítios de início de transcrição e regiões promotoras dos genes que codificam membros do T6SS em *K. pneumoniae*;
- 5) Predizer sítios de ligação de proteínas reguladoras às regiões regulatórias dos genes do T6SS.
- 6) Correlacionar os reguladores transcricionais identificados com estímulos ambientais que possam ativar ou bloquear sua atividade.

3 MÉTODOS

3.1 Sequências genômicas de *K. pneumoniae*

A sequência de nucleotídeos dos genomas completos e anotação dos cromossomos e plasmídeos das cepas Kp52.145, HS11286 e NTUH-K2044 de *K. pneumoniae* foram baixadas do projeto NCBI RefSeq. Os números de acesso destas sequências estão apresentados na Tabela 2.

Tabela 2: Números de acesso dos cromossomos e plasmídeos de três cepas de *K. pneumoniae* analisadas neste estudo.

Cepa	Cromossomo	Plasmídeo 1	Plasmídeo 2
Kp52.145	NZ_FO834906	NZ_FO834904	NZ_FO834905
HS11286	NC_016845.1	NC_016838.1	NC_016846.1
NTUH-K2044	NC_012731.1	NC_006625.1	-

A cepa de *K. pneumoniae* Kp52.145 que é uma cepa do sorotipo K2, altamente virulenta da qual foram identificados importantes fatores de virulência, incluindo o grande plasmídeo de virulência que abriga o regulador do fenótipo mucóide (*rmpA*) (Nassif e Sansonetti 1986). A cepa HS11286 é um isolado de escarro humano (*human sputum*) proveniente do Hospital Huashan em Shanghai, China. Contém um cromossomo de 5.3 Mb e três plasmídeos de resistência a múltiplas drogas (110 kb), incluindo um produtor de carbapanemase e 3 plasmídeos pequenos (3 kb) (Liu et al. 2012). A cepa NTUH-K2044 é um isolado do sangue de um homem de 40 anos, previamente saudável, sofrendo de abscesso hepático primário adquirido na comunidade com meningite metastática. Foi isolada pela primeira vez na Ásia e depois foi encontrada em todo o mundo. Possui os genes *magA* e *rmpA*, pertence ao sorotipo capsular K1, possui alta virulência e hiper mucoviscosidade. Estes fatores tornam esse isolado muito adequado como uma cepa modelo para estudos genômicos (Wu et al. 2009).

3.2 Identificação *in silico* de genes codificadores do T6SS

3.2.1 Identificação dos componentes principais

A identificação de genes que codificam componentes putativos do T6SS pode ser realizada pela busca de sequências de aminoácidos e nucleotídeos similares a de componentes do T6SS já identificados (Sarris et al. 2011), bem como pela identificação das chamadas “ilhas de TssI”, onde é possível identificar componentes e efetores putativos, codificados nessas ilhas (Ma et al. 2017). Outro método de

identificação destes componentes nos genomas é através do uso de COGs (*Cluster of Orthologs Groups*), construídos a partir das sequências de aminoácidos dos componentes do T6SS identificados em outros genomas (Boyer et al. 2009).

Neste estudo, utilizamos o software T6SS-HMMER3 e o banco de dados SecReT6 (Type VI Secretion system Fonte: <http://db-mml.sjtu.edu.cn/SecReT6/index.php>) para identificar homólogos do T6SS nos três genomas de *K. pneumoniae*. O SecReT6 foi construído a partir da análise de mais de 11 mil componentes do T6SS e 906 regiões que agrupam estes genes, em 240 espécies (Li et al. 2015). Nesse banco de dados, os componentes do T6SS estão agrupados de acordo com COGs, em 13 grupos, correspondendo aos componentes TssA a TssM. Para cada um desses grupos, perfis baseados em modelos ocultos de Markov (HMM: *hidden Markov models*) foram gerados. Usamos o T6SS-HMMER3, que utiliza a abordagem representada na Figura 11, para detectar genes codificadores de proteínas de *K. pneumoniae* que correspondam a estes perfis.

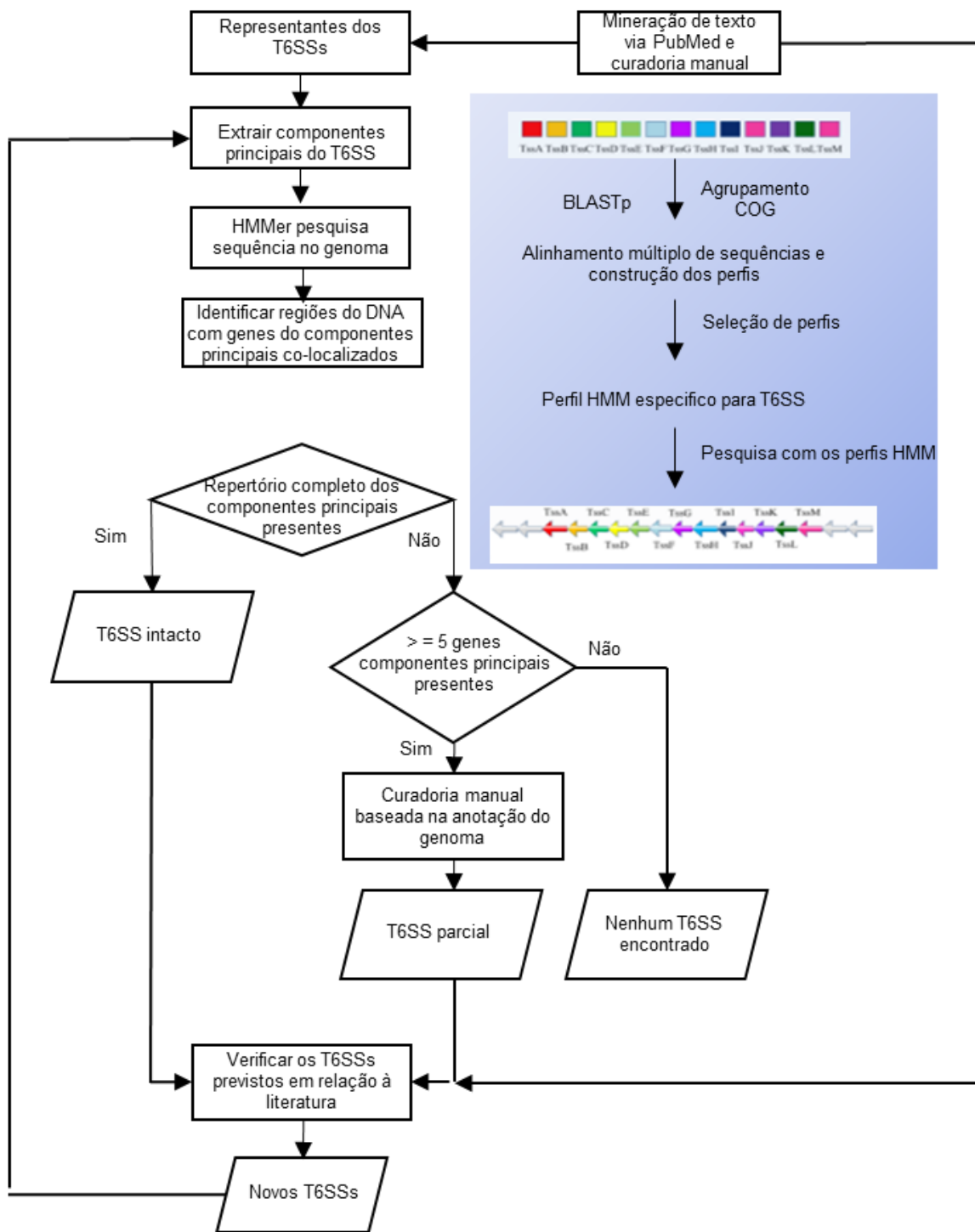


Figura 11: Estratégia usada pelo T6SS-HMMER/SecReT6 para prever e identificar genes que codificam o T6SS. O painel, no canto superior direito, ilustra a construção dos perfis *Hidden Markov Model* (HMM) para os 13 principais componentes do T6SS. Adaptado de (Li et al. 2015).

Como critério para identificar genes do T6SS o T6SS-HMMER avalia a organização destes em tandem. Neste caso, é necessária a detecção de ao menos cinco genes em tandem que codifiquem proteínas do T6SS. Entre estes cinco (ou mais) genes e o próximo gene que codifique um possível componente do sistema,

pode haver uma região variável que possivelmente codifica proteínas efetoras e de imunidade dependente do T6SS. A extensão dessa região variável pode ser personalizada no T6SS-HMMER. Neste estudo, testamos a distância padrão fornecida no T6SS-HMMER (15.000 pb), mas também avaliamos a distância de 16.000 pb. Outro critério utilizado para filtrar os resultados gerados foi o *e-value* $<10^{-18}$. O *e-value* reflete o número de ocorrências que se pode encontrar ao acaso, pesquisando num banco de dados de um tamanho específico. Quanto menor o *e-value*, ou quanto mais próximo de zero, mais significativa é a correspondência encontrada. Além de encontrar genes que codificam os componentes do T6SS em *locus*, o T6SS-HMMER também identifica estes genes órfãos, desagrupados dos *loci* do T6SS, juntamente com sua região vizinha de 5kb.

3.2.2 Identificação dos componentes acessórios

O T6SS-HMMER contém informações apenas sobre os 13 principais componentes do T6SS. Assim, para realizar a busca dos genes que codificam componentes acessórios, designados *tag* (*Type Six-Associated Genes*), nos três genomas de *K. pneumoniae*, utilizamos a ferramenta COGviewer (http://202.120.12.133/STEP/STEP_COG.html), disponível no VRprofile (http://202.120.12.134//STEP/Instruction_STEP.htm). O VRprofile é uma plataforma projetada para identificar e localizar genes do banco de dados mobilomeDB. Este banco é construído sobre conjuntos de genes conhecidos de sistemas de secreção bacterianos do tipo III / IV / VI / VII e elementos genéticos móveis, incluindo elementos integrativos e conjugativos, prófagos, integrons de classe I, elementos de inserção de sequências e ilhas de resistência a antibióticos em genomas bacterianos (Li et al. 2017).

O “COGviewer” pesquisa um conjunto de genes codificadores de proteínas, a partir de entradas de COGs fornecidas pelo usuário, utilizando o perfil HMM. Utilizamos os COGs: COG4104, COG1357, COG0515, COG3913, COG0631, COG3456, COG4455, COG2885, para identificar genes que codifiquem PAAR, TagB, TagE, TagF, TagG, TagH, TagJ e TagL, conforme apresentado na Tabela 1. Os componentes acessórios que não dispunham de identificadores COG, como TagX e MltE, foram pesquisados pelo BLASTp a partir do alinhamento das sequências de aminoácidos (ID: QCR57576.1 e NP_415711.2) anotadas do genoma de *Acinetobacter baumannii* ATCC17978 e *E. coli* K-12 MG1655, respectivamente.

3.3 Análise das sequências dos genes putativos do T6SS

3.3.1 Alinhamento de sequências

As sequências de aminoácidos codificadas pelos genes identificados como codificadores de componentes TssH (número de acesso: WP_002889795.1; WP_002891807.1; WP_016532438.1; WP_046043127.1; YP_005228284.1; YP_005226093.1; YP_005225441.1; YP_005226603.1; WP_004890358.1; WP_002891807.1; WP_002896522.1; WP_015874721.1), nas três cepas de *K. pneumoniae*, foram anotadas e alinhadas pelo Clustal (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Os domínios conservados foram identificados no CDD (Banco de Dados de Domínios Conservados) (Marchler-Bauer et al. 2012) e PFAM (Banco de Dados de Famílias de Proteínas) (El-Gebali et al. 2019). A sequência de aminoácidos da proteína ClpB de *Francisella tularensis* foi baixada do NCBI (número de acesso YP_170660.1) e utilizada como referência para comparação.

3.3.2 Determinação conteúdo GC

As sequências de nucleotídeos das regiões que agrupam genes putativos do T6SS, identificadas pelo T6SS-HMMER, foram importadas no programa Artemis (Carver et al. 2012). Neste programa, foi utilizada a ferramenta de análise da variação do conteúdo de GC das regiões anotadas.

3.4 Construção das sintenias

As sequências de nucleotídeos das regiões contendo os genes putativos do T6SS, identificadas pelo T6SS-HMMER e COGviewer, foram anotadas e alinhadas entre si pelo BLASTn com o recurso on-line “DoubleACTv2.0” (http://www.hpa-bioinfotools.org.uk/pise/double_actv2.html), para gerar um arquivo de comparação. Os alinhamentos foram visualizados com o *Artemis Comparison Tool* (Carver et al. 2005). As regiões de sintenia foram representadas em blocos vermelhos sólidos, representando identidade de sequência > 87%.

3.5 Identificação das sequências promotoras

3.5.1 Sequências promotoras de sigma-70 (σ^{70})

Nas regiões genômicas do T6SS, anotamos 250 pares de base à montante dos genes que codificam os componentes putativos do T6SS. Estas sequências foram analisadas pelo algoritmo de predição de sequências promotoras “BPPROM” (*bacterial promoter prediction*:

<http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfin> db). O BPROM é um programa de predição de promotores bacterianos a partir da detecção de sítios de ligação putativos do fator σ^{70} , com sensibilidade de 83% e especificidade de 84% no reconhecimento das sequências promotoras. Seu melhor uso é feito em análises de regiões imediatamente a montante do início da janela aberta de leitura (ORF – *Open Reading Frame*) (V. Solovyev 2011). Para identificar características significativas das regiões promotoras bacterianas, este programa utiliza cinco motivos de sequências conservadas. Dois motivos, mais conservados, correspondem aos elementos da sequência -10 e -35 dos promotores regulados por σ^{70} . O terceiro motivo está à montante da caixa -35 com comprimento de 7 pb, pesquisado na área -60 a -40. O quarto motivo é pesquisado a jusante do bloco -10, com comprimento de 7 pb, pesquisado na área -11 a +10; e o quinto motivo está entre as caixas -35 e -10 com comprimento de 5 pb pesquisado na área -31 a -22 (V. Solovyev 2011). Estes motivos estão representados na Figura 12. Para reduzir a taxa de previsão de falsos positivos, analisamos apenas os 250 pares de base a montante das ORFs previstas nas anotações originais dos genomas.

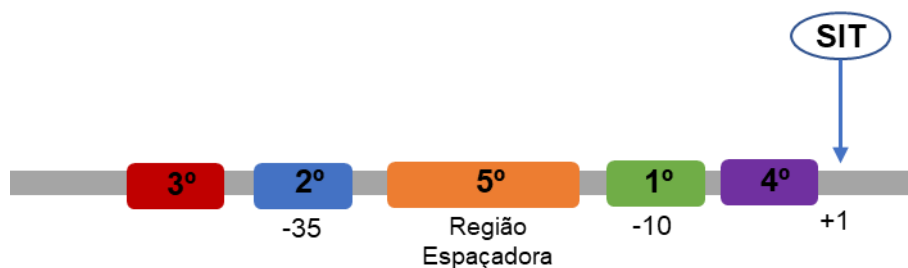


Figura 12: Representação das regiões dos motivos para identificação de sítios de ligação do fator σ^{70} . Na imagem podemos observar a representação dos cinco motivos pesquisados para identificar uma região promotora de σ^{70} pelo BPROM. As letras em maiúsculo representam as caixas -35, -10 e o SIT (Sítio de Início de Transcrição) que inicia na posição +1. Adaptado de (Gordon et al. 2003).

3.5.2 Sequências promotoras de sigma-54 (σ^{54})

Para reconhecer sítios putativos de σ^{54} , analisamos as sequências de nucleotídeos, com 250 pares de base, a montante das ORFs que codificam os componentes putativos do T6SS, no Virtual Footprint 3.0 (http://www.prodoric.de/vfp/vfp_promoter.php). Esta é uma ferramenta de pesquisa sensível para reconhecer padrões de DNA. Este programa foi projetado para analisar os sítios de ligação de fatores de transcrição em genomas bacterianos. O Virtual Footprint permite uma varredura de sequências baseada em matrizes de peso-posição, associada ao banco de dados PRODORIC (Munch 2003). Este banco de

dados possui perfis dos sítios de ligação de 163 fatores de transcrição, representados como matrizes de peso-posição (PWMs) e logotipos de sequência, onde o tamanho da letra da base indica o grau de conservação daquela posição no alinhamento. Todo o conjunto de dados do Virtual Footprint é curado manualmente e conta com resultados publicados extraídos da literatura científica.

No Virtual Footprint 3.0, inserimos as sequências de 250 pares de bases anotadas e selecionamos a matriz RpoN, referente ao fator σ^{54} . É preciso destacar que este tipo de análise de busca por sítios de ligação, pode gerar falsos positivos. Para diminuir a presença destes falsos positivos no Virtual Footprint, foram testados os valores 0,7 a 1,0 para o parâmetro de precisão (*Sensitivity*), 0,7 a 1,0 no uso das posições mais conservadas e a quantidade dessas posições mais conservadas variando de 5 a 8 (*Core Sensitivity/Size*). Após estes testes, utilizamos os valores de 1,0 para aumentar a precisão da busca. Para recuperar as posições mais conservadas utilizamos o valor 1,0 para as oito posições mais conservadas de cada matriz. Com estes ajustes, pudemos recuperar os sítios com até oito pares de base com o máximo de confiabilidade.

3.6 Reguladores transcricionais codificados nos genomas de *K. pneumoniae*

Como o objetivo final deste estudo é a identificação de sítios de ligação de reguladores transcricionais às regiões promotoras dos genes do T6SS, precisamos analisar os genomas de *K. pneumoniae* para identificar os reguladores transcricionais codificados neles. Além disso, cruzamos estes dados com a lista de reguladores transcricionais cujos sítios de ligação estão descritos por matrizes de peso-posição no banco de dados do Virtual Footprint.

No Virtual Footprint encontramos matrizes para sítios de ligação de 52 reguladores transcricionais, sendo que alguns reguladores possuem mais de uma matriz disponível. A sequência de aminoácidos desses 52 reguladores em *E. coli* K-12 (número de acesso: NC_000913.3) foram alinhadas com os genomas das cepas de *K. pneumoniae* usando a ferramenta tBLASTn. *Hits* que apresentaram identidade maior que 70% e cobertura maior que 90% foram considerados presentes em *K. pneumoniae*.

3.7 Predição de sítios de ligação de fatores transcricionais nas sequências promotoras dos genes do T6SS

Para identificar sítios de ligação de fatores transcricionais nas sequências promotoras dependentes dos fatores σ^{70} e σ^{54} , anotamos 250 pares de base a montante dos sítios de início de transcrição identificados pelo BROM e pelo Virtual Footprint (Apêndice D). Estas predições foram realizadas no Virtual Footprint 3.0. Na Figura 13, temos um exemplo de uma matriz disponível no Virtual Footprint para o regulador transcricional H-NS. Neste exemplo, podemos observar no alinhamento que há posições muito conservadas, indicando que H-NS reconhece estas sequências muito parecidas entre si. Esses padrões de ligação podem ser representados por alinhamentos ou por Logos, onde o tamanho da letra da base indica o grau de conservação da posição no alinhamento.

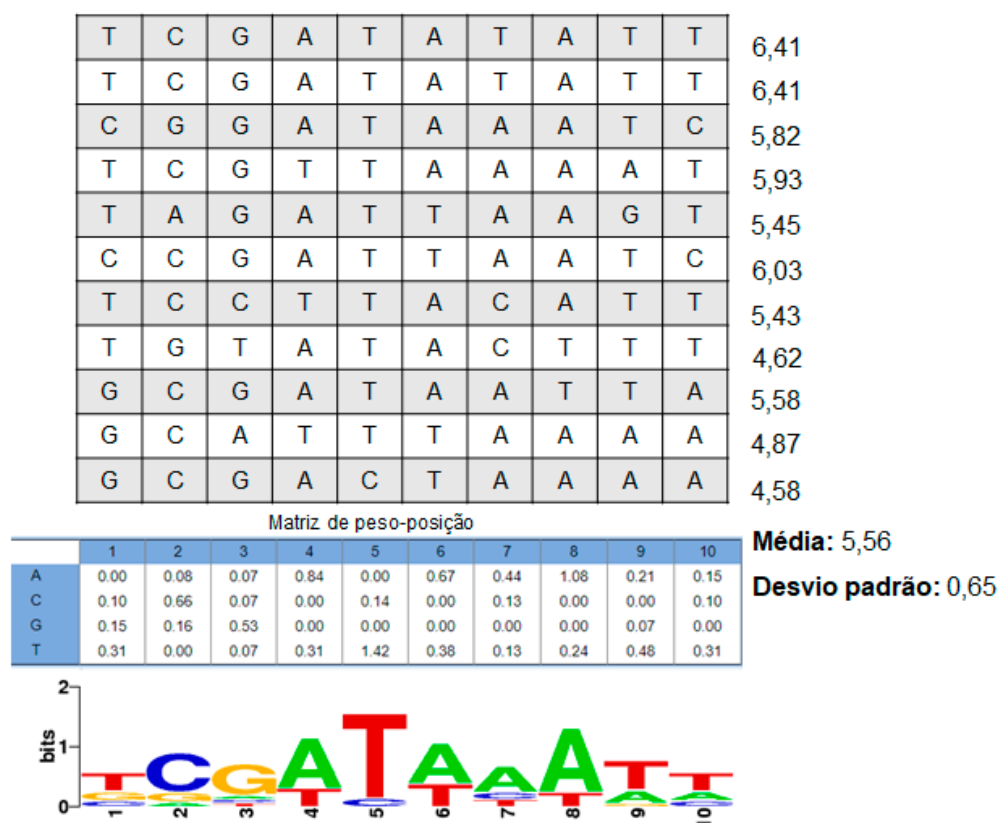


Figura 13: Representação da construção de uma matriz de peso-posição no Virtual Footprint. A construção é realizada a partir do alinhamento de sítios de ligação anotados para o fator transcricional H-NS. Cada sequência de nucleotídeos com 10 bases representa um sítio de ligação de H-NS ao DNA de *E. coli* k-12 determinado experimentalmente e representados em forma de Logos, onde o tamanho da letra da base indica o grau de conservação da posição no alinhamento. Ao lado direito estão representadas as pontuações das sequências de acordo com sua matriz de peso-posição.

A partir da sequência de sítios de ligação validados experimentalmente são construídas matrizes de peso-posição (Grote et al. 2009). Essas matrizes refletem a

frequência que cada base ocorre em cada posição de um sítio de ligação. Ainda, baseado nessa informação, uma pontuação pode ser atribuída ao comparar uma sequência de interesse com estas matrizes. Utilizamos 46 matrizes de peso posição, construídas a partir de sítios de ligação de 44 fatores transcricionais (encontrados nos genomas de *K. pneumoniae*) de *E. coli* K12 disponíveis (Tabela 3). Estas matrizes foram escolhidas por terem seus ligantes codificados nos três genomas de *K. pneumoniae*, objetos de estudo deste trabalho.

Para um sítio ser considerado relevante, na varredura das sequências promotoras inseridas, utilizamos o *score* somado ao desvio padrão de cada matriz de peso posição, para filtrar os resultados encontrados. Alternativamente, a informação contida nessas matrizes foi representada em formato de Logo, utilizando o programa WebLogo (<https://weblogo.berkeley.edu/logo.cgi>).

4 Resultados e Discussão

4.1 Identificação de genes do T6SS nas cepas de *K. pneumoniae*

Estudos anteriores já tinham identificado genes do T6SS em genomas de *K. pneumoniae* (Sarris et al. 2011; Lery et al. 2014; Liu et al. 2017; Hsieh et al. 2018; Chen et al. 2019). Sarris e colegas (2011) anotaram genes do T6SS nas três cepas de *K. pneumoniae* totalmente sequenciadas na época (342, NTUH-K2044 e MGH78578) e uma cepa parcialmente sequenciada (*K. pneumoniae subsp. rhinoscleromatis* ATCC13884) (Sarris et al. 2011). Quando os primeiros genes do T6SS foram identificados em *K. pneumoniae*, havia poucos genomas completos de *K. pneumoniae* disponíveis para comparação. A pouca disponibilidade de dados para comparação pode levar a uma anotação com nomenclatura inconsistente dos componentes do T6SS. Isto pode ocorrer pois a anotação as vezes é parcial, se referindo a algum domínio específico de uma família; em outros casos, a anotação é feita com base em genes que não tem nomes informativos ou são apenas “putativos” ou “hipotéticos”. Assim, apesar desses estudos prévios, faz-se necessário reanotar os genes e padronizar sua nomenclatura. Para o desenvolvimento desta tese nos concentramos nas cepas patogênicas humanas: Kp52.145 (uma cepa virulenta K2) (Nassif e Sansonetti 1986), HS11286 (resistente a múltiplas drogas) (Liu et al. 2012) e NTUH-K2044 (uma cepa virulenta K1) (Fang et al. 2004).

A análise aqui apresentada permitiu identificar genes e *loci* não identificados anteriormente, bem como propor uma padronização da anotação dos genes do T6SS usando nomenclatura TssA-M para os principais componentes do T6SS em *K. pneumoniae*. Essa nomenclatura foi introduzida por Shalom e cols em *B. pseudomallei* para superar o problema da não padronização histórica na nomenclatura dos genes T6SSs (Shalom et al. 2007).

Assim, os genomas das três cepas de *K. pneumoniae* mencionadas acima foram reanalisados para identificar e reanotar os genes putativos codificantes de proteínas do T6SS. O ponto de partida para esse processo de anotação foi o conjunto de resultados gerados pela análise que fizemos no banco de dados SeCReT6 e MobilomeDB. Os resultados das análises obtidas nesta etapa encontram-se no Apêndice A - C.

Os genomas destas *K. pneumoniae* possuem mais de uma região que contém genes do T6SS. A cepa Kp52.145 possui três regiões, enquanto as cepas HS11286 e NTUH-K2044 possuem duas regiões que codificam esses componentes. Por um lado, essa organização em múltiplos *loci* em *K. pneumoniae* já havia sido descrita em

estudos anteriores (Sarris et al. 2011; Lery et al. 2014; Liu et al. 2017). Por outro lado, vale a pena ressaltar que nessa reanálise dos genomas de *K. pneumoniae*, utilizamos parâmetro de colocação entre dois componentes do T6SS de 16.000 pb, o que nos permitiu identificar uma nova região com genes do T6SS (T6SS-2) na cepa HS11286 - devido a região variável dentro do *cluster* T6SS-2 ser maior que 15.000 pb.

Este tipo de organização dos genes do T6SS, em diferentes *loci*, também é observada em espécies bacterianas como *P. aeruginosa* e *B. pseudomallei* (Schell et al. 2007; Sana et al. 2013). Curiosamente, cada uma das três regiões contendo genes do T6SS (H1 a H3) em *P. aeruginosa* PAO1 é requerida para situações diferentes e regulada por diferentes mecanismos de regulação transcricional (Sana et al. 2013). T6SS-H1 tem como alvo células procarióticas, mas o T6SS-H2 e H3 medeiam a invasão de células epiteliais (Sana et al. 2013, 2015). A disposição dos genes do T6SS em mais de uma região nos genomas de *K. pneumoniae* sugere que estes T6SSs possam apresentar funções distintas e serem transcritos em diferentes condições ambientais em que a bactéria se encontre.

Em cada genoma, identificamos ainda genes que codificam componentes do T6SS, mas que não pertencem às regiões denominadas de *cluster* ou ilhas do T6SS. Alguns desses genes, chamados de genes órfãos, haviam sido reportados em trabalhos anteriores de Kp52.145, HS11286 e NTUH-2044. Na Figura 14, temos as posições dos genes que codificam os principais componentes do T6SS nas três cepas de *K. pneumoniae*, numa representação circular dos genomas.

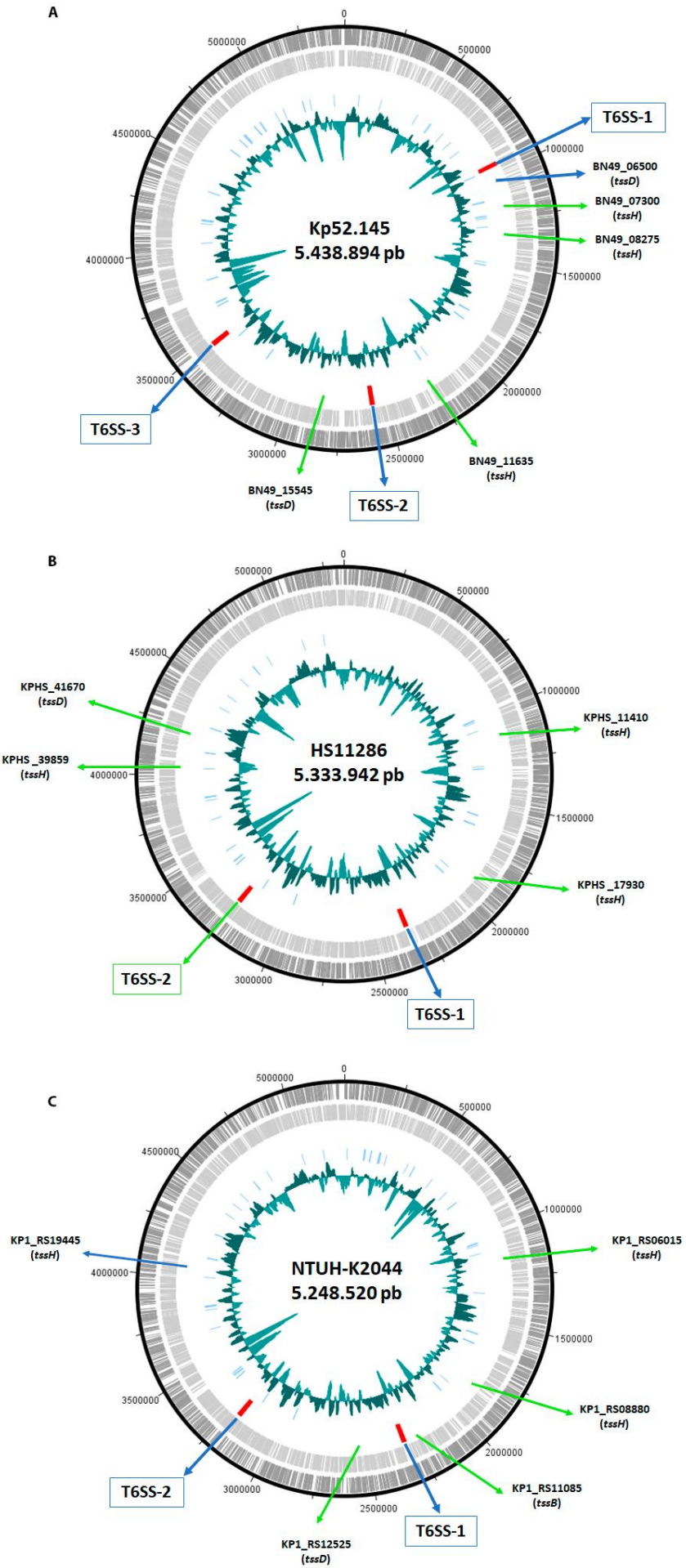


Figura 14: Características gerais dos genomas de *K. pneumoniae* analisados e a localização dos genes que codificam componentes do T6SS. Os cromossomos circulares estão representados como círculos e uma escala em kpb está representada. O nome de cada cepa e o tamanho total de cada cromossomo está indicado no centro do círculo. **A** - Kp52.145, **B** - HS11286 e **C** - NTUH-2044. Destacado em vermelho, estão as ilhas do T6SS. Nas setas em azul estão os genes que codificam os componentes do T6SS já identificados em outros trabalhos e as setas em verde representam os genes do T6SS identificados neste trabalho, como assinalado em cada imagem. As indicações com locus_tag indicam os genes órfãos do T6SS, encontrados distantes dos agrupamentos de genes do T6SS. Em cinza claro e escuro estão representados as CDS e em verde a representação do conteúdo GC.

Na Tabela 3, temos as anotações dos 38 genes da Kp52.145 encontrados utilizando o T6SS-HMMER e COGviewer.

Tabela 3: Genes que codificam componentes putativos do T6SS de Kp52.145.

Locus tag BN49_RS	Anotação	CP	Loci
05940	Type VI secretion system baseplate subunit	TssE	I
05945	Type VI secretion system lipoprotein	TssJ	I
05965	Type VI secretion system baseplate subunit	TssG	I
05970	Type VI secretion system baseplate subunit	TssF	I
05990	PAAR domain-containing protein	PAAR	I
05995	Hypothetical protein	TssA	I
06000	Type VI secretion protein VskK	TssM	I
06025	Type VI secretion system tip protein VgrG	TssI	I
06030	Hypothetical protein	TagL	I
06035	DotU family type IV/VI secretion system protein	TssL	I
06040	Type VI secretion system baseplate subunit	TssK	I
14020	Type VI contractile secretion system sheath small subunit	TssB	II
14025	Type VI secretion system contractile sheath large subunit	TssC	II
14030	Type VI secretion system baseplate subunit	TssK	II
14035	DotU family type IV/VI secretion system protein	TssL	II
14040	OmpA Family protein	TagL	II
14050	Type VI secretion system ATPase	TssH	II
14070	PAAR domain-containing protein	PAAR	II
14075	Type VI secretion protein VskK	TssM	II
14080	Type VI secretion system protein	TssA	II
14085	Type VI secretion system baseplate subunit	TssF	II
14090	Type VI secretion system baseplate subunit	TssG	II
14095	Type VI secretion system lipoprotein	TssJ	II
18715	Membrane protein	TagA	III
18720	Type VI secretion system baseplate subunit	TssE	III
18725	Type VI secretion system lipoprotein	TssJ	III
18730	Type VI secretion system baseplate subunit	TssG	III
18735	Type VI secretion system baseplate subunit	TssF	III
18750	PAAR domain-containing protein	PAAR	III
18800	Type VI secretion system tip protein VgrG	TssI	III
18805	OmpA Family protein	TagL	III
18810	DotU family type IV/VI secretion system protein	TssL	III
18815	Type VI secretion system baseplate subunit	TssK	III
06500	Type VI secretion system tube protein Hcp	TssD	Orfão
15545	Type VI secretion system tube protein Hcp	TssD	Orfão
07300	ATP-dependent chaperone ClpB	TssH	Orfão
08275	ATP-dependent Clp ATP-binding subunit ClpX	TssH	Orfão
11635	ATP-dependent Clp ATP-binding subunit ClpA	TssH	Orfão

Legendas - CP: Componentes putativos

Ainda sobre a Kp52.145, é importante frisar que a anotação do genoma foi atualizada recentemente (FO834906.1 de 07-MAR-2015 / NZ_FO834906.1 de 21-FEB-2017). Nessa reanotação, identificamos em dois *loci*, seis genes (BN49_3360; BN49_3361; BN49_2467; BN49_2468; BN49_2470; BN49_2471) inicialmente

preditos como codificadores de componentes do T6SS reanotados como pseudogenes na versão mais recente (BN49_RS18740, BN49_RS14045, BN49_RS14055 e BN49_RS05960). Compreendidos como cópias desativadas de genes funcionais, os pseudogenes são sequências genômicas não funcionais, mas com similaridade de sequência com genes que codificam proteínas (Zou et al. 2009). Os pseudogenes surgem como cópias de genes que perderam sua função original devido ao acúmulo de mutações, que podem causar a inativação deles ou mudanças nas ORFs com o surgimento de códons de parada prematuros (Kovalenko e Patrushev 2018). Portanto, estes dados indicam que os *loci* II e III podem estar sob processo de pseudo-genificação.

Em nossas análises, identificamos 33 genes que codificam componentes do T6SS (87%) localizados nessas três regiões (*locus* I a III). Além destes genes, encontramos cinco genes órfãos em Kp52.145: três que codificam possíveis TssH (ClpATPases) e dois genes codificando o componente TssD. Um desses *tssD* (BN49_RS06500) já havia sido descrito (Lery et al. 2014).

Na cepa HS11286, identificamos ao todo 29 genes possivelmente relacionados ao T6SS (Tabela 4). Da mesma forma que na Kp52.145, nossa análise detectou na HS11286 genes previamente não relacionados ao T6SS. Em estudo anterior, tinha sido revelada uma região do T6SS contendo 23 genes (KPHS_22970 a 23190) no genoma de *K. pneumoniae* HS11286. Destes, 12 genes codificam os principais componentes do T6SS (Liu et al. 2017). Através da metodologia utilizada neste estudo, encontramos mais 17 genes codificando componentes putativos do T6SS, totalizando 29 genes. De forma diferente que o observado para Kp52.145, em HS11286 os genes do T6SS estão organizados majoritariamente em duas regiões (*loci* I e II). Nos *loci* I e II há 25 genes (86%) agrupados com 15 (51%) e 11(38%) genes respectivamente, enquanto quatro genes órfãos (14%) foram identificados. Os genes órfãos são: três genes que codificam os componentes putativos de TssH (KPHS_39850, KPHS_17930, KPHS_11410) e um gene codifica um componente TssD (KPHS_41670).

Tabela 4: Genes que codificam componentes putativos do T6SS de HS11286.

Locus tag KPHS_	Anotação	CP	Loci
22970	Hypothetical protein	TssB	I
22980	EvpB/family type VI secretion protein	TssC	I
22990	Hypothetical protein	TssK	I
23000	Hypothetical protein	TssL	I
23010	OmpA/MotB domain-containing protein	TagL	I
23020	Hypothetical protein	TssD	I
23030	Putative protease	TssH	I
23040	Rhs element Vgr protein	TssI	I
23120	PaaR repeat-containing protein	PAAR	I
23140	ImcF domain-containing protein	TssM	I
23150	Type VI secretion-associated protein	TssA	I
23160	Type VI secretion-associated protein	TssA	I
23170	Hypothetical protein	TssF	I
23180	Hypothetical protein	TssG	I
23190	Hypothetical protein	TssJ	I
32450	Hypothetical protein	TagA	II
32460	Type VI secretion system lysozyme-like protein	TssE	II
32470	Hypothetical protein	TssJ	II
32480	Hypothetical protein	TssG	II
32490	Hypothetical protein	TssF	II
32500	Hypothetical protein	TssM	II
32520	Hypothetical protein	PAAR	II
32730	Hypothetical protein	TssI	II
32750	Hypothetical protein	TagL	II
32760	Membrane protein	TssL	II
32770	Hypothetical protein	TssK	II
39850	Protein disaggregation chaperone	TssH	Orfão
41670	Type VI secretion system effector	TssD	Orfão
17930	ATP-dependent Clp protease ATP-binding subunit	TssH	Orfão
11410	ATP-dependent protease ATP-binding subunit	TssH	Orfão

Legendas - CP: Componentes putativos

No genoma da cepa NTUH-K2044 encontramos ao todo 30 genes que codificam possíveis membros do T6SS (Tabela 5). Destes, 25 tinham sido descritos em trabalho anterior (Sarris et al. 2011). Nessa cepa, os genes estão organizados em dois grandes *loci* (I e II) contendo 14 (46%) e 11 (36%) genes, respectivamente. Ainda, foram identificados cinco (16%) genes órfãos.

Tabela 5: Genes que codificam componentes putativos do T6SS de NTUH-2044.

Locus tag KP1_RS	Anotação	CP	Loci
11220	Type VI secretion system contractile sheath small subunit	TssB	I
11225	Type VI secretion system contractile sheath large subunit	TssC	I
11230	Type VI secretion system baseplate subunit	TssK	I
11235	DotU family type IV/VI secretion system protein	TssL	I
11240	OmpA family protein	TagL	I
11245	Type VI secretion system tube protein Hcp	TssD	I
11250	Type VI secretion system ATPase	TssH	I
11255	Type VI secretion system tip protein VgrG	TssI	I
11295	Type VI secretion protein VskK	TssM	I
11300	Type VI secretion system protein	TssA	I
11305	Hypothetical protein	PAAR	I
11315	Type VI secretion system baseplate subunit	TssF	I
11320	Type VI secretion system baseplate subunit	TssG	I
11325	Type VI secretion system lipoprotein	TssJ	I
15690	Membrane protein	TagA	II
15695	Type VI secretion system baseplate subunit	TssE	II
15700	Type VI secretion system lipoprotein	TssJ	II
15705	Type VI secretion system baseplate subunit	TssG	II
15710	Type VI secretion system baseplate subunit	TssF	II
15715	Type VI secretion protein VskK	TssM	II
15725	PAAR domain-containing protein	PAAR	II
15775	Type VI secretion system tip protein VgrG	TssI	II
15780	Outer membrane protein	TagL	II
15785	DotU family type IV/VI secretion system protein	TssL	II
15790	Type VI secretion system baseplate subunit	TssK	II
12525	Type VI secretion system tube protein Hcp	TssD	Orfão
19445	ATP-dependent chaperone ClpB	TssH	Orfão
11085	Type VI secretion system contractile sheath small subunit	TssB	Orfão
06015	ATP-dependent Clp protease ATP-binding subunit ClpX	TssH	Orfão
08880	ATP-dependent Clp protease ATP-binding subunit ClpA	TssH	Orfão

Legendas - CP: Componentes putativos

4.2 TssA, TssA1 e TagA em *K. pneumoniae*

Conforme descrito na introdução, a proteína TssA coordena a biogênese do T6SS, a extensão do tubo de TssD e seu revestimento pela bainha formada por TssB-TssC (Brunet et al. 2015b). TssA compartilha um domínio conservado com TssA1 e TagA: o domínio ImpA_N. As análises que fizemos no SecReT6 apontaram a presença de três genes codificadores de componente TssA (presença do domínio ImpA_N) em Kp52.145 e HS11286 e dois genes em NTUH-K2044. Considerando as possíveis diferenças funcionais entre as proteínas TssA, TssA1 e TagA, realizamos

uma análise dos domínios funcionais das sete proteínas previstas como TssA nas três cepas de *K. pneumoniae* (Figura 15).

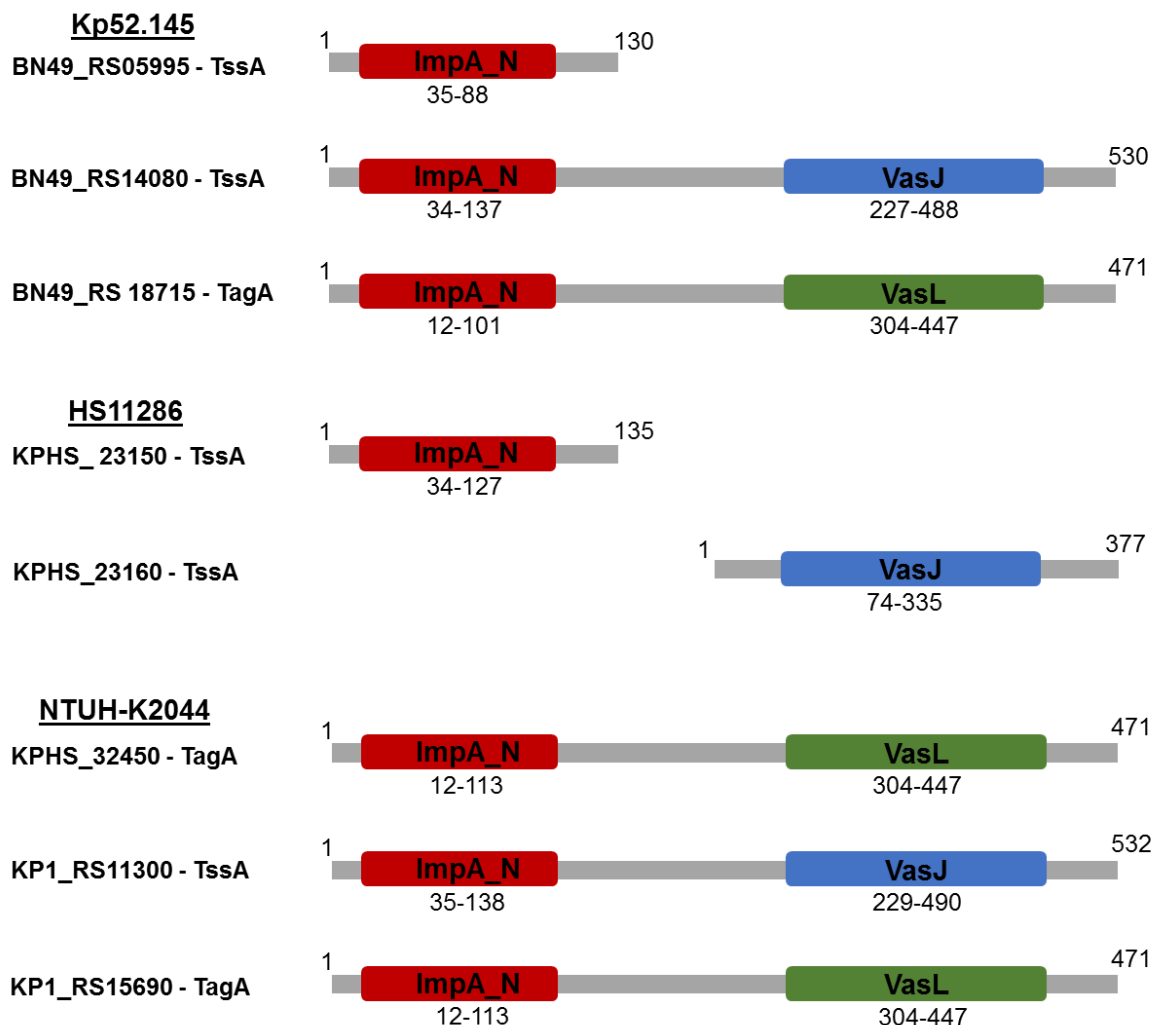


Figura 15: Representação esquemática das seqüências de aminoácidos de TssA e TagA de *K. pneumoniae*, ressaltando os domínios conservados. O domínio ImpA_N (acesso PF06812) é mostrado em vermelho. Nas porções C-terminal, em verde, há um domínio de função desconhecida VasL (acesso PF12486) e, em azul, o domínio VasJ (acesso PF16989). As seqüências de aminoácidos das proteínas TagA nas três cepas de *K. pneumoniae* analisadas compartilham 99,94% de identidade.

Através dessa análise, verificamos que nenhuma cepa analisada codifica TssA1, mas todas possuem TssA e TagA. Os genes identificados como *tssA* (BN49_RS18715, KPHS_32450 e KP1_RS15690), codificam proteínas que possuem o domínio ImpA_N característico da TssA. Entretanto na porção C-terminal há um domínio VasL e não VasJ (característico da TssA), portanto estes genes codificam o componente acessório TagA.

O gene *tagA* está normalmente presente nos agrupamentos de genes do T6SS que codificam o TssA, mas não o TssA1 (Santin et al. 2018). Em nossas análises identificamos genes que codificam TssA e TagA nos *loci* do T6SS nas três cepas de

K. pneumoniae, o que sugere que a extensão do T6SS nestas cepas seja controlada com o auxílio do componente acessório TagA.

4.3 ATPases Clp: componentes TssH identificados em *K. pneumoniae*

As análises no SecReT6 apontaram quatro genes em cada genoma analisado codificando possíveis componentes TssH, todos com sequências pertencentes ao COG0542. ClpV ATPases foram inicialmente descritas como o membro TssH do complexo T6SS, necessárias para a formação de um tubo funcional e reciclagem dos componentes da bainha (Filloux 2009; Kapitein et al. 2013; Förster et al. 2014). Posteriormente, foi demonstrado que *F. tularensis* tem um T6SS funcional que usa a proteína ClpB ATPase ao invés de ClpV (Brodmann et al. 2017). Curiosamente, uma das possíveis proteínas identificadas como possível componente TssH de *K. pneumoniae* corresponde ao gene anotado como *clpB* (KP1_RS19445) da cepa NTUH-K2044. Assim, decidimos curar manualmente a anotação destes genes, através de buscas de similaridade de sequências e análise de domínios, e definir o tipo de ATPase que codificam. As proteínas ClpV nos genomas de *K. pneumoniae* são codificadas pelos genes BN49_RS14050 em Kp52.145, KPHS_23030 em HS11286 e KP1_RS11250 em NTUH-K2044. Estas ClpVs de *K. pneumoniae* são 99% idênticas. A pouca divergência encontrada entre as três ClpV não pertence às sequências correspondentes aos domínios *Walker A* e *Walker B*, caracterizados como essenciais para a atividade ATPásica (Hanson e Whiteheart 2005; Schlieker et al. 2005).

Os outros três genes identificados como *tssH*, de cada cepa, codificam as proteínas ClpB, ClpA e ClpX. De acordo com a anotação do genoma, as proteínas ClpB das cepas Kp52.145, HS11286 e NTUH-K2044 teriam 857, 823 e 857 aminoácidos de comprimento, respectivamente. De acordo com nossa análise, o códon de iniciação para a ClpB da HS11286 foi determinado incorretamente. Nós sugerimos que esta sequência de nucleotídeos seja 102 pb mais longa, codificando uma proteína de 34 aminoácidos a mais e totalizando 857 aminoácidos, o mesmo tamanho das ClpB das demais cepas, como apresentamos na Figura 16.

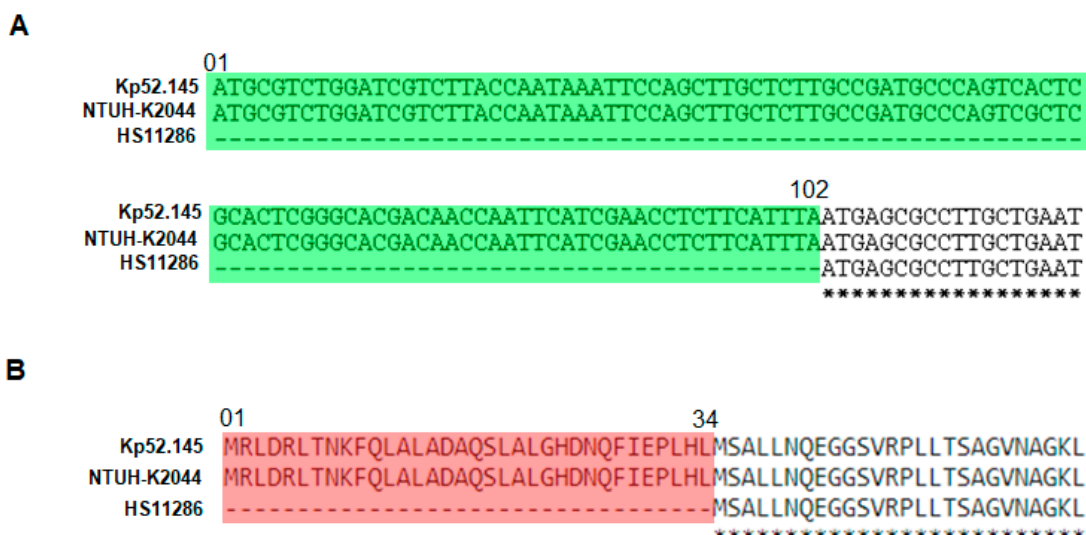


Figura 16: Reanotação da sequência de aminoácidos da ClpB da cepa HS11286, tendo como base às ClpBs de Kp52145 e NTUH-K2044. Na HS11286 a sequência de nucleotídeos a ser incorporada ao gene *clpB* encontra-se entre as posições 01 e 102 (em verde) e a de aminoácidos entre 01 e 102 (em vermelho). * indicam identidade de todos os aminoácidos.

Considerando a reanotação da ClpB na cepa HS11286, as sequências de aminoácidos e nucleotídeos das três cepas de *K. pneumoniae* possuem 100% de identidade (Tabela 6).

Tabela 6: Identidade e similaridade entre as ATPases ClpB, ClpA e ClpX identificadas como componente TssH no genoma de três *K. pneumoniae*.

ATPase	Número de acesso	I%	S%
ClpB	WP_004150973.1(Kp52.145)	YP_005228284.1(HS11286)	100 100
	WP_004150973.1(Kp52.145)	WP_004890358.1(NTUH-K2044)	100 100
ClpA	WP_016532438.1(Kp52.145)	YP_005226093.1(HS11286)	99 99
	WP_016532438.1(Kp52.145)	X WP_002896522.1(NTUH-K2044)	99 99
ClpX	WP_002891807.1(Kp52.145)	YP_005225441.1(HS11286)	100 100
	WP_002891807.1(Kp52.145)	WP_002891807.1(NTUH-K2044)	100 100

Legendas - I%: Percentual de identidade entre as sequências; S%: Percentual de similaridade entre as sequências.

Ft	MDINKFTIKLQEALAE	AQSYAFQKQKATEFTSAHILKALLEQNDVAISILSVQGVNIQNF	60
Kp52.145	MRLDRLTNKFLALADA	AQSLALGHDNQFIEPLHLMSALLNQEGGSRPLLSAGVNAAGL	60
HS11286	MRLDRLTNKFLALADA	AQSLALGHDNQFIEPLHLMSALLNQEGGSRPLLSAGVNAAGL	60
NTUH-K2044	MRLDRLTNKFLALADA	AQSLALGHDNQFIEPLHLMSALLNQEGGSRPLLSAGVNAAGL	60
	* : : : *	* : : : *	
Ft	IKAVNDMVD	SVAVLSGEGNPQVTPSRDLIATLHKMGLANKNGDEFISSEVFLLASLEDK	120
Kp52.145	RTDIEQALSRLPQVEGTG	-GDVQPSQDLVRIINLCKLAQKKKDNFISSELFVLAALLESR	119
HS11286	RTDIEQALSRLPQVEGTG	-GDVQPSQDLVRIINLCKLAQKKKDNFISSELFVLAALLESR	119
NTUH-K2044	RTDIEQALSRLPQVEGTG	-GDVQPSQDLVRIINLCKLAQKKKDNFISSELFVLAALLESR	119
	. : : : .	. : : : .	
Ft	-SLTGLYNKFGITKEKLT	KAVNDYRGGKRVSSQNQEDMKGALDKYTVDLTDLARKGKIDP	179
Kp52.145	GILTDLLKSAGATTANVTQAI	EQMRGGSVNDQGAEDQRQALKKFTVDLTERAEQKGLDP	179
HS11286	GILTDLLKSAGATTANVTQAI	EQMRGGSVNDQGAEDQRQALKKFTVDLTERAEQKGLDP	179
NTUH-K2044	GILTDLLKSAGATTANVTQAI	EQMRGGSVNDQGAEDQRQALKKFTVDLTERAEQKGLDP	179
	*** : : *	*** : : *	
Ft	IIGRDEIRRTIQVLQRRTKNNP	VLIIEPQVGRITAVEGLAQRIVNDEVPEGVKGGKVVLS	239
Kp52.145	VIGRDEEIRRTIQVLQRRTKNNP	VLIIEPQVGRITAVEGLAQRIVNDEVPEGLKGRRVLA	239
HS11286	VIGRDEEIRRTIQVLQRRTKNNP	VLIIEPQVGRITAVEGLAQRIVNDEVPEGLKGRRVLA	239
NTUH-K2044	VIGRDEEIRRTIQVLQRRTKNNP	VLIIEPQVGRITAVEGLAQRIVNDEVPEGLKGRRVLA	239
	:*****	:*****	
Ft	LDMGALLAGAKFRGDFEERL	KSVLKELSKQEGNVILFIDELHTMVGAGKAEQSGMDAGNML	299
Kp52.145	LDMGALVAGAKYRGEFEERL	KGVLTDLKQEGNVILFIDELHTMVGAGKADGAMDAGNML	299
HS11286	LDMGALVAGAKYRGEFEERL	KGVLTDLKQEGNVILFIDELHTMVGAGKADGAMDAGNML	299
NTUH-K2044	LDMGALVAGAKYRGEFEERL	KGVLTDLKQEGNVILFIDELHTMVGAGKADGAMDAGNML	299
	*****	*****	
Ft	KPALARGELKCVGATTLDEY	REYVEKDPALERRFQKVLVDEPTVEDTIAILRGLKERYEL	359
Kp52.145	KPALARGELHCVGATTLDEY	RQYIEKDAALERRFQKVFVAEPSVEDTIAILRGLKERYEL	359
HS11286	KPALARGELHCVGATTLDEY	RQYIEKDAALERRFQKVFVAEPSVEDTIAILRGLKERYEL	359
NTUH-K2044	KPALARGELHCVGATTLDEY	RQYIEKDAALERRFQKVFVAEPSVEDTIAILRGLKERYEL	359
	*****	*****	
Ft	HHGVNITDSAI	VSAATLSHRYITDRQLPKAIDLIDEAASQIRMEIDSKPEFMESLYRRI	419
Kp52.145	HHHVQITDPA	IVAAATLSHRYIADRQLPKAIDLIDEAASSIRMQIDSKPEELDRDRRI	419
HS11286	HHHVQITDPA	IVAAATLSHRYIADRQLPKAIDLIDEAASSIRMQIDSKPEELDRDRRI	419
NTUH-K2044	HHHVQITDPA	IVAAATLSHRYIADRQLPKAIDLIDEAASSIRMQIDSKPEELDRDRRI	419
	*****	*****	
Ft	IQLKMQREQLKKEKDDATK	KRLEILEQEIKGLDSEYKQLEELWKAELKMQOTSKLKEEL	479
Kp52.145	IQLKLEQQALK	KESDEASLKRLDMLNEELADKERQYSVLEEEWKAELKASLSOTQTIKAEEL	479
HS11286	IQLKLEQQALK	KESDEASLKRLDMLNEELADKERQYSVLEEEWKAELKASLSOTQTIKAEEL	479
NTUH-K2044	IQLKLEQQALK	KESDEASLKRLDMLNEELADKERQYSVLEEEWKAELKASLSOTQTIKAEEL	479
	*****	*****	
Ft	EKAKFELEKYQRVODLSK	MAELQYQKIPLELAQIKQIEETEAPSENKLVRTSVTENEIA	539
Kp52.145	EQAKIAIEQARRVGD	LARMSSELQYQKIPLEKQLAAATQS--EGKTMRLLRNKVTDIAEIA	537
HS11286	EQAKIAIEQARRVGD	LARMSSELQYQKIPLEKQLAAATQS--EGKTMRLLRNKVTDIAEIA	537
NTUH-K2044	EQAKIAIEQARRVGD	LARMSSELQYQKIPLEKQLAAATQA--EGKTMRLLRNKVTDIAEIA	537
	*****	*****	
Ft	DVVS	KATGIPVSKMMEGEKDKLLNMFSLHHRVIGQDQAIKAVSNVRRSRGSLSDPNRP	599
Kp52.145	EVLARWGTGIPVSR	MESEERDKLLRMEQELHHRVIGQDEAVEAVSNAIARRSRAGLSDPNRP	597
HS11286	EVLARWGTGIPVSR	MESEERDKLLRMEQELHHRVIGQDEAVEAVSNAIARRSRAGLSDPNRP	597
NTUH-K2044	EVLARWGTGIPVSR	MESEERDKLLRMEQELHHRVIGQDEAVEAVSNAIARRSRAGLSDPNRP	597
	:*****	:*****	
Ft	IGSF	FLGPTGVGKTELTKALAEFLDDEDAMLRVDMSEFMKHSVRLVIGAPPGVVGYE	659
Kp52.145	IGSFLFLGPTGVGKTEL	CKTLANFMFSDDAMVRI DMSEFMKHSVSRVIGAPPGVVGYE	657
HS11286	IGSFLFLGPTGVGKTEL	CKTLANFMFSDDAMVRI DMSEFMKHSVSRVIGAPPGVVGYE	657
NTUH-K2044	IGSFLFLGPTGVGKTEL	CKTLANFMFSDDAMVRI DMSEFMKHSVSRVIGAPPGVVGYE	657
	*****	*****	
Ft	QGGYL	TEAVRRRPYSVILLDEVEKAHADIFNILLQVLDGRLTDGQGRVDFRNTVIMT	719
Kp52.145	EGGYL	TEAVRRRPYSVILLDEVEKAHPDVFNILLQVLDGRLTDGQGRVDFRNTVIMT	717
HS11286	EGGYL	TEAVRRRPYSVILLDEVEKAHPDVFNILLQVLDGRLTDGQGRVDFRNTVIMT	717
NTUH-K2044	EGGYL	TEAVRRRPYSVILLDEVEKAHPDVFNILLQVLDGRLTDGQGRVDFRNTVIMT	717
	*****	*****	
Ft	SNL	GSHRIQEMQG-QDYETVKSVMEMVLSHFRPEFVNRVDDAIVFPLNKEMITEIAKI	778
Kp52.145	SNL	GSDLIQERFGEIDYGHMKDLVLGVSQNFPEFINRIDEVVVFHPLGEKHASIAIQI	777
HS11286	SNL	GSDLIQERFGEIDYGHMKDLVLGVSQNFPEFINRIDEVVVFHPLGEKHASIAIQI	777
NTUH-K2044	SNL	GSDLIQERFGEIDYGHMKDLVLGVSQNFPEFINRIDEVVVFHPLGEKHASIAIQI	777
	*****	*****	
Ft	QIKRLEKRLADLSI	GLEVTTGAMDKLADAGDPVFGARPKRAIQNNLENPLAKLLDGE	838
Kp52.145	QLQRLYKRL	EERGYEVQMSDEALKLLSANGYDPVYGARPKRAIQQIENPLAQQLSGE	837
HS11286	QLQRLYKRL	EERGYEVQMSDEALKLLSANGYDPVYGARPKRAIQQIENPLAQQLSGE	837
NTUH-K2044	QLQRLYKRL	EERGYEVQMSDEALKLLSANGYDPVYGARPKRAIQQIENPLAQQLSGE	837
	*****	*****	
Ft	FRAEDKIVVDIDANNI	IFSK 859	
Kp52.145	LVP	GKTIELVVKDDRIVAVQ- 857	
HS11286	LVP	GKTIELVVKDDRIVAVQ- 857	
NTUH-K2044	LVP	GKTIELVVKDDRIVAVQ- 857	
	: . * : .	: . * : .	

Figura 17: Alinhamento das seqüências de aminoácidos das proteínas ClpB de *F. tularensis* e *K. pneumoniae*. ClpB: Alinhamento múltiplo das proteínas ClpB de *F. tularensis* SCHU (Ft), *K. pneumoniae* Kp52.145, HS11286 e NTUH-K2044. As regiões verde e amarela indicam o domínio Clp_N; Vermelho, azul e roxo são os domínios AAA, AAA2 e ClpB_D2-small, respectivamente. As caixas preta e azul indicam os domínios Walker A e Walker B; ▼ apontam para resíduos do local de ligação de ATP. * indicam identidade de todos os aminoácidos; “:” indica a similaridade de aminoácidos alinhados; “.” indica baixa similaridade de aminoácidos alinhados, de acordo com os parâmetros padrão do ClustalW.

Ainda, observamos que a proteína ClpB de *K. pneumoniae* é conservada em relação à ClpB de *F. tularensis*. O alinhamento das sequências exibe 64% de identidade e 79% de similaridade. Além disso, a ClpB de *K. pneumoniae* apresenta os mesmos domínios conservados que o ortólogo em *F. tularensis*: os domínios Clp_N, ClpB_D2small, AAA e AAA2, bem como o *Walker A*, *Walker B*, sítio de ligação do ATP e arginina, como mostramos na Figura 17, que sugere que eles podem ser capazes de desempenhar papéis semelhantes.

Devido à similaridade entre as sequências de ClpB de *K. pneumoniae* e *F. tularensis*, propomos que as ClpB de *K. pneumoniae* possam funcionar como TssH alternativas em *K. pneumoniae*. No entanto, ainda não foi esclarecido se essas proteínas são de fato necessárias e/ou montadas como parte do T6SS em *K. pneumoniae*.

Avaliando as proteínas ClpA das três cepas de *K. pneumoniae*, constatamos um compartilhamento de 99% de identidade (Tabela 6). Comparadas com a sequência de aminoácido da proteína ClpV, são ~53% similares e 35% idênticas em um alinhamento com 87% de cobertura. Apesar de tais dissimilaridades de sequência, ClpA, ClpB e ClpV contêm os domínios conservados alinhados (Figura 18).

ClpV	MENPASLLRRLNPPCCARAMEG	AASLCQTRAHAEILPEHWLLKLLLEQGEGDLTVLARRYEW	60				
ClpB	-----MRLDRLTNKFLALADA	QSLALGHDNQFIEPLHLMSALLNQEGGSVRPLLSAGV	55				
ClpA	-----MLNQELESLNMA	AFARAREHRHFMTVEHLLALLSNPS--AREALEACS	49				
		*. : : * : . : : : * : ** :					
ClpV	DMDAL	WQDLLSWLDKQPRSVRH-----RPQLSDHTLRLMQEAWLIASLSGEAQIRSVHL	114				
ClpB	NAGKLR	RTDIE----QALSRLPQVEGTGGDVPQSDLVRIINLCKDLQAKK--DNFISSEL	110				
ClpA	DLVAL	RQELEAFIEQTTPLVLPATE-EERDTQPTLSFQRFVLRQAVFHVQSSGRSEVTGANV	108				
		: * : : : : * : * : : . . . : . . . :					
ClpV	LMALVEKQNLIQCDGLWPLLT	LQGQRLERLRP--LLDAQSDERPPAQEAAALQPHGGDV	172				
ClpB	FVLALESRGT	----LTDLL--K-----SAGATTANV---TQAI EQMRG---	145				
ClpA	LVAIFSEQESQ	----AAYLL--RKHEVSRLDVVNFISHGTRKDEP---SQSS--DNS---	154				
		: : : . . ** : : : : : : : : . . :					
ClpV	EFVGRPAGSELNADGLNPALQ	NALDKFTLDVTAKARDGLIDPVFGRDTEIRQMVDILSRR	232				
ClpB	-----GESVN	-DQGAEDQEQALKKFTVDLTERAEQKLDVPIGRDEEIRRTIQVLQRR	197				
ClpA	-----GSQPG	-NEEQAGGEERMENTINLNQLARVGGIDPLIGREKELGRAIQVLCRR	206				
		* . . . : : : : * . * * : : * : * : : * * *					
		↑↑↑↑↑↑↑					
ClpV	RKNNP	ILV	BEPGVGKTA	ALVEGLALRIAEGNV	PDALKPVSVRTLDLGLLQAGAGVKGEFEQ	292	
ClpB	TKNNP	VL	I	BEPGVGKTA	IVEGLAQRIVNGEVPELKGRRVLALDMGALVAGAKYRGFEF	257	
ClpA	RKNNP	LLV	BESGVGKTA	IAEGLAWRIVQGDVPEVMADCTIYSLDIGSLLAGTKYRDFEK	266		
				****. * * * * : . * * * * : * : * * : : : * * * * * : * : * * :			
				↓			
ClpV	RLKNI	IEAVQQSPSP	LL	FIDE	AHTIIGAGNQG--GGADAANLLKPALARGE	LRTIAATTW	351
ClpB	RLKGV	LDLSKQEGNV	IL	FIDE	LHTMVGAGKAD--GAMDAGNMLKPALARGE	LHCVGATTL	316
ClpA	RFKAL	LKQLE--QDTNS	IL	FIDE	LHTIIGAGAAGGQVDAANLIKPLSSGKIRIVG	325	
				↓			
				↓			
ClpV	SEYKQYFERDAALERRFQM	VKVDEPDDDTACLMLRGLKSRVADHHGVHITDDAVRAAVTL	411				
ClpB	DEYRQYIEKDAALERRFQ	KVFAEPSVEDTIAILRGLKERYELHHVQITDPAIVAAATL	376				
ClpA	QEFSNIFEKDRALARRFQ	KIDITEPSVEETVQIINGLKPKYEAHHDVRYTAKAVRAAVEL	385				
		. * : : * * * * * : : * * . : : : . * * * * : * * * : * * * * *					
ClpV	SRRYLTGRQLPDKAVD	LDDTASARLRMSLDTVPEPLTRMKAQLTALAMEKQALLEDIALG	471				
ClpB	SHRYIADRLQPDKAID	LIDEAASSIRMQIDSKPEELDRDRRIIQLKLEQQAALKKESDEA	436				
ClpA	AVKYINDRHLPDKAID	VIDEAGARARLMPVSKRKK-----	420				
		: : * : . * * * * * : * * : * : : : :					
ClpV	NSARGDRLAAIEQEEI	RLILALDLETQYQQE-----LQLTEALLACRRDIS	518				
ClpB	SLKRLDMLN	---EELADKERQYSVLEEWEKAEKASLSGTQTIKAELEQAKIAIEQARRVG	493				
ClpA	-----	-----	420				
ClpV	RQAE-----	INDLQITALIA---VQQGNPLGLD	VDTVIVADWIGVPLSSLMK	566			
ClpB	DLARMSELQYGKI	PELEKQLAAATQSEKIMRLLRNKVTDAEIAEVLARWIGIPVSRMME	553				
ClpA	-----	-----TVNVADIESVVIARIARIPEKSVSR	444				
		* : : * * : * * : . . :					
ClpV	DEQTELLSLEESL	GKRVVQEAALSAIARRLRAAKTGLTPENGPQGVFLLVGPSGTGKTE	626				
ClpB	SERDKLLRMEQEL	HHRVIGQDEAVEAVSNAIRRSRAGLSDPNRPIGSFLFGPTGVGKTE	613				
ClpA	SDRDITLKNL	GDRLKMLVFGQDKAIEALTEAIKMARAGLGHEHKPVGSLFAGPTGVGKTE	504				
		. : : * : : * * * * * : * : * : . : : * * : * * * * * : * * * * * * * * *					
ClpV	KALALADALFGGEKAL	ITINLSEYQEPHTVSQKGSPPGVYVGGGILTEAVRKRFPYSV	686				
ClpB	LCKTLANFMFDS	DAMVRIIDMSEFMEKHSVSRVGGAPPVYVYEEGGYLTAVRRRFPYSV	673				
ClpA	VIVQLAKA	---LGIELLRFDMSEYMERHTVSRVIGAPPVYVGFQDQGLLTDVAVIKHFAV	561				
		** : : : * * * * * : * * * * * : * * * * * : * * * * * : * * * * *					
ClpV	VLLDEVEKAHRD	VMNLFYQVFD	RGVMDGEGREIDFRNTVILMTANLGSDDLMLQLLDEQP	746			
ClpB	ILLDEVEKAHP	DVFNILLQVLD	DGRLTDGQRTVDFRNTVIMTSLGSDLIQERFELD	733			
ClpA	LLLDEIEKAHP	DVFNILLQVMD	NGTLTDNNGRKADFRNVVMTNAGVRETERKSIGLI	621			
		: * * * * * * * * : * * * * * : * * * * * * * * *					
ClpV	EASESDLHELLRP	VLRGHFQALLARFQVI--YRPLPAGALRAIVGMKLGQVSQRLACHY	805				
ClpB	---YGHMKDL	VLGVSQNFREFINRIDEVVVFHPLGKHHIASIAQIQLQRLVKRLE-ER	789				
ClpA	---QQNSPDAMDE	IKKIFTFEFRNLDNI IWF DHLSTTVIHQVVDKFI	VELQVQLD-QK	677			
		: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *					
ClpV	GITTTLS	ESLFDALTEACL	LPDTGARNVDSLLNQILPALSQQLLSHMAAGQKPRQVTLG	865			
ClpB	GVEVQMSDEAL	KLLSANGYDPVYGARPLKRAIQQQIENPLAQQILSGELVPGKTI	ELV-	848			
ClpA	GVSLEVSQ	EARDWLATKGYDRAMGARPMGRVIQDNLKKPLANELLFGSLV	DGGQVTVSL-	736			
		* : * : . . * : * * * * : : : : * : * : * : :					
ClpV	YHEEEGVVMA	FDGEGTISDE-----	884				
ClpB	-KDDR-IVAVQ	-----	857				
ClpA	-DSDK-NITLYDFQ	SAPKHKPEAAH	759				
						

Figura 18: Alinhamento múltiplo das proteínas ClpV, ClpB e ClpA de *K. pneumoniae* Kp52.145. A região verde indica o domínio Clp_N; Vermelho, azul e roxo são os domínios AAA, AAA2 e ClpB_D2-small, respectivamente. As caixas preta e azul indicam os domínios Walker A e Walker B; ▼ apontam para resíduos do local de ligação de ATP. * indicam identidade de todos os aminoácidos; ":" indica a similaridade de aminoácidos alinhados; "." indica baixa similaridade de aminoácidos alinhados, de acordo com os parâmetros padrão do ClustalW.

As sequências de ClpX de *K. pneumoniae* são idênticas entre as três cepas, no entanto a ClpX tem 424 aminoácidos de comprimento, enquanto ClpV tem 884. Se ClpA e ClpX poderiam desempenhar o papel de TssH como membro do complexo T6SS, há necessidade de mais estudos para que seja elucidado.

4.4 Contexto genético dos genes órfãos

Tendo sido demonstrado a expressão de genes órfãos que codificam componentes do T6SS em condições de infecção e em isolados clínicos (Bent et al. 2016), seria razoável considerar a importância destes genes para a funcionalidade do T6SS nas bactérias. A partir de análises filogenéticas, foi sugerido que a aquisição de genes *tssD* órfãos por *Pantoea ananatis* seria oriunda de diferentes ancestrais ou proveniente da duplicação de genes associados ao T6SS (Shyntum et al. 2014). As três cepas de *K. pneumoniae*, deste trabalho possuem genes órfãos relacionados ao T6SS (Tabelas 3, 4 e 5), todos codificando possíveis componentes essenciais para o funcionamento do complexo, como TssD e TssH (Mougous et al. 2006; Bönemann et al. 2009; Silverman et al. 2013; Ma et al. 2017). Ao analisarmos a vizinhança dos genes codificadores dos componentes TssD órfãos nas cepas Kp52.145 e HS11286 (BN49_06500 e KPHS_41670), percebemos que as anotações dos genes eram similares.

Essas regiões contêm genes que codificam proteínas hipotéticas, transportadores de ferro / heme da família ABC (BN49_RS06505 a BN49_RS06520) e um receptor dependente de TonB (BN49_RS06525). Em HS11286, a vizinhança do gene órfão *tssD* (KPHS_41670) apresenta genes que codificam um sistema de transporte de heme ATP (KPHS_41660 a KPHS_41640), uma metil transferase dependente de S-adenosil metionina (KPHS_41630), uma porina de membrana externa para enterobactina férrica e colicinas B e D (KPHS_41620). Embora esses genes relacionados ao ferro em Kp52.145 e HS11286 tenham nomes diferentes, eles são ortólogos e seus produtos apresentam sequências idênticas. Assim, esses genes órfãos *tssD*, em ambas as cepas, são codificados na mesma região genômica. Alinhando esses genes observamos que as sequências de nucleotídeos apresentam 99% de identidade, o que aponta uma origem evolutiva comum dessa região.

Outro ponto interessante a ser ressaltado é que os genes vizinhos a esses *tssD* mencionados codificam estruturas relacionadas ao transporte de íons Fe^{3+} , como a proteína TonB (Qiu e Misra 2019). A proteína TonB é um componente essencial na transdução de energia de um sistema transmembrana citoplasmático de transporte de

íons Fe em bactérias Gram-negativas (Dong et al. 2019). Curiosamente, foi demonstrado que em resposta ao estresse oxidativo e para contrapor defesas de imunidade de células hospedeiras, o T6SS de *Y. pseudotuberculosis* e de *B. thailandensis* secretam efetores envolvidos na captação de íons. A importação de íons, nestas espécies, é mediada pelas proteínas da família ABC e/ou TonB (Wang et al. 2015; Si et al. 2017a). Os íons metálicos transportados por mecanismos dependentes do T6SS em *Y. pseudotuberculosis*, *B. thailandensis* e *P. aeruginosa* são Zn^{2+} , Mn^{2+} e Fe^{3+} (Wang et al. 2015; Lin et al. 2017; Si et al. 2017a). Dessa forma, acreditamos que estes transportadores de íons codificados nas cepas Kp52.145 e HS11286, adjacentes a genes do T6SS, possam apresentar função relacionada ao T6SS. Até o momento, nenhum estudo tinha sugerido o papel do T6SS de *K. pneumoniae* na captação de íons metálicos. Contudo, com base na arquitetura genômica, apontamos esta possibilidade.

Por outro lado, a internalização de íons mediada pelo T6SS depende da secreção de proteínas efetoras do T6SS, que se ligam aos íons e são importadas pelos receptores ABC e TonB: em *Y. pseudotuberculosis*, o efector é YezP (Wang et al. 2015); em *B. thailandensis*, são os efetores TseM (Si et al. 2017b) e TseZ (Si et al. 2017a); e em *P. aeruginosa* é o efector TseF (Lin et al. 2017). Até o momento não identificamos genes homólogos aos que codificam esses efetores nos genomas de *K. pneumoniae* analisados. Desta forma, é possível supor, caso ocorra o transporte de íons via T6SS em *K. pneumoniae*, seria possível que ele ocorra mediante algum outro efector ainda não identificado.

Por outro lado, na vizinhança do gene putativo *tssH/clpB* (BN49_RS07300 na Kp52.145, KPHS_39859 na HS11286 e KP1_RS19445 na NTUH-K2044), encontramos o gene *pgeF* - codificando um fator de edição de peptidoglicano. A PgeF contribui para a manutenção da composição da cadeia peptídica das moléculas de peptidoglicanas em *E. coli*, contribuindo assim para a integridade dessa estrutura bacteriana (Parveen e Reddy 2017). Além disso encontramos que os genes órfãos estão acompanhados por genes com atividades paralelas e relacionadas às atividades do T6SS. Na vizinhança de *tssH/clpX* (BN49_RS08275, KPHS_11410 e KP1_RS06015), encontramos um gene que codifica uma outra ATPase, ClpP, e uma Endopeptidase, LonA, que em *V. cholerae* ativa a expressão de genes do T6SS (Rogers et al. 2016). Na vizinhança do *tssH/clpA* (BN49_RS11635, KPHS_17930 e KP1_RS0880), encontramos genes que codificam uma ClpS ATPase e uma proteína de transporte de macrolídeos, conforme podemos acompanhar na Figura 19.

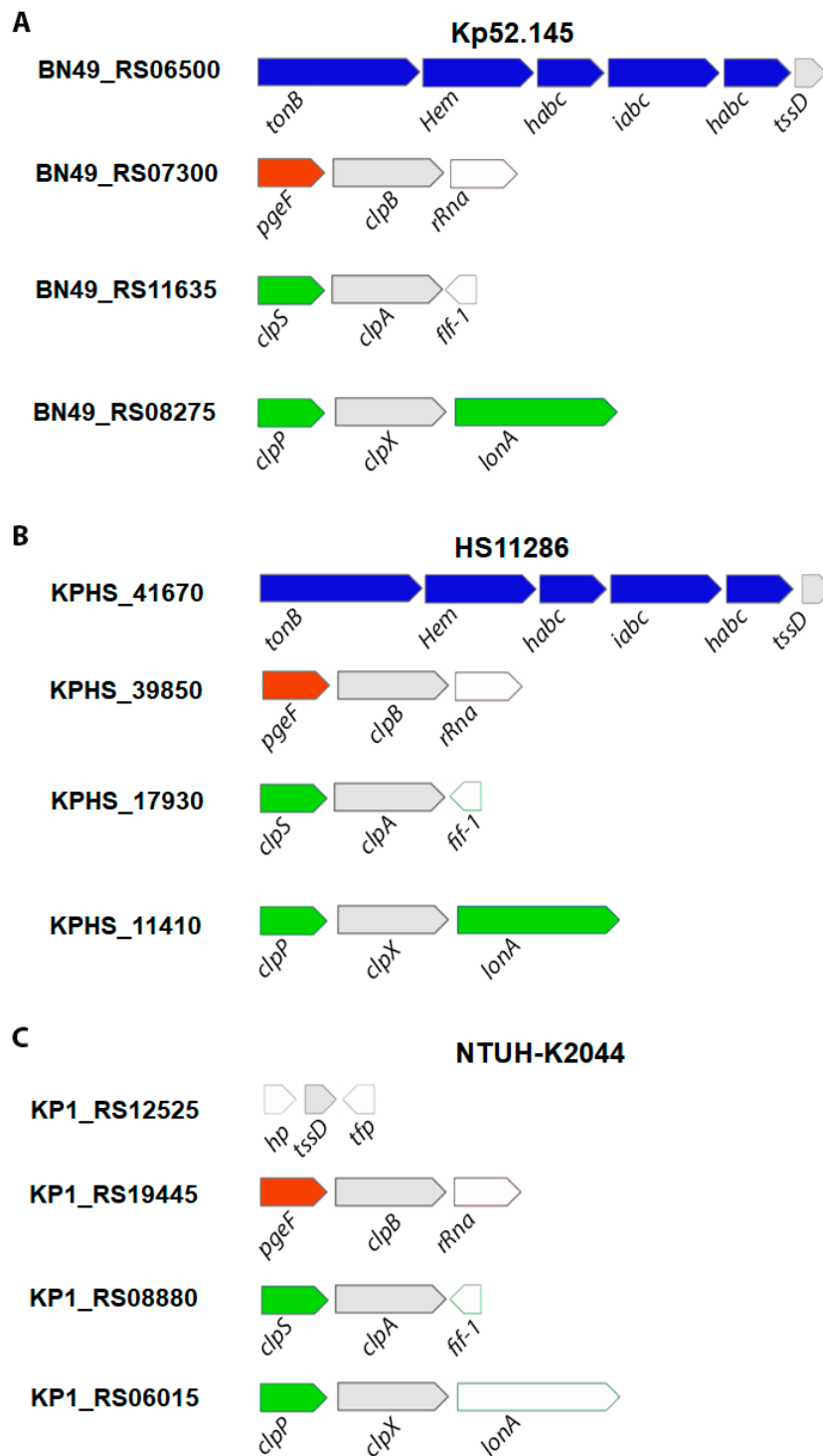


Figura 19: Análise do contexto genômico de genes órfãos do T6SS em Kp52.145, HS11286 e NTUH-K2044. Genes que codificam transportadores relacionados ao ferro são adjacentes ao *tssD* nos genomas do *K. pneumoniae*. Kp52.145 (A), HS11286 (B) e NTUH-K2044 (C). Ao lado de cada grupo de genes, estão os *locus_tag* dos componentes do T6SS. Os genes relacionados a T6SS são coloridos em cinza; os genes que codificam proteínas relacionadas à absorção de ferro estão em azul; gene relacionado à integridade da camada de peptidoglicanas em vermelho; genes que codificam proteases e peptidases em verde e os genes que codificam outras funções são representados em branco. *tonB* – *TonB-dependent receptor*; *Hem* – *hemin-degrading factor*; *habc* - *heme ABC transporter ATP-binding protein*; *iabc* – *iron ABC transporter permease*; *pgeF* – *peptideoglycan editing fator PgeF*; *rRNA* – *16S ribosomal RNA*; *fff-1* – *translation initiation factor IF-1*; *lonA* – *endopeptidase La*; *hp* – *hypothetical protein*.

O fato dos genes do T6SS poderem ser coregulados, indica que suas ações sejam paralelas como em *V. cholerae* (Metzger et al. 2016). Outros experimentos, como a verificação da liberação de Fe/heme a partir de células hospedeiras sob ataque do T6SS, precisam ser realizados para verificar se esses produtos gênicos podem estar relacionados com o T6SS e ser usados para competição bacteriana e/ou invasão tecidual.

4.5 Sintenia entre *loci* do T6SS em *K. pneumoniae*

Os resultados aqui apresentados mostram que os genes do T6SS em *K. pneumoniae*, codificados no genoma das cepas Kp52.145, HS11286 e NTUH-K2044, são de algum modo conservados entre si, em termos de semelhança de sequências. Além disso, eles são semelhantes às regiões codificadoras de T6SS em *E. coli* E042 e *P. ananatis* (Figura 20).

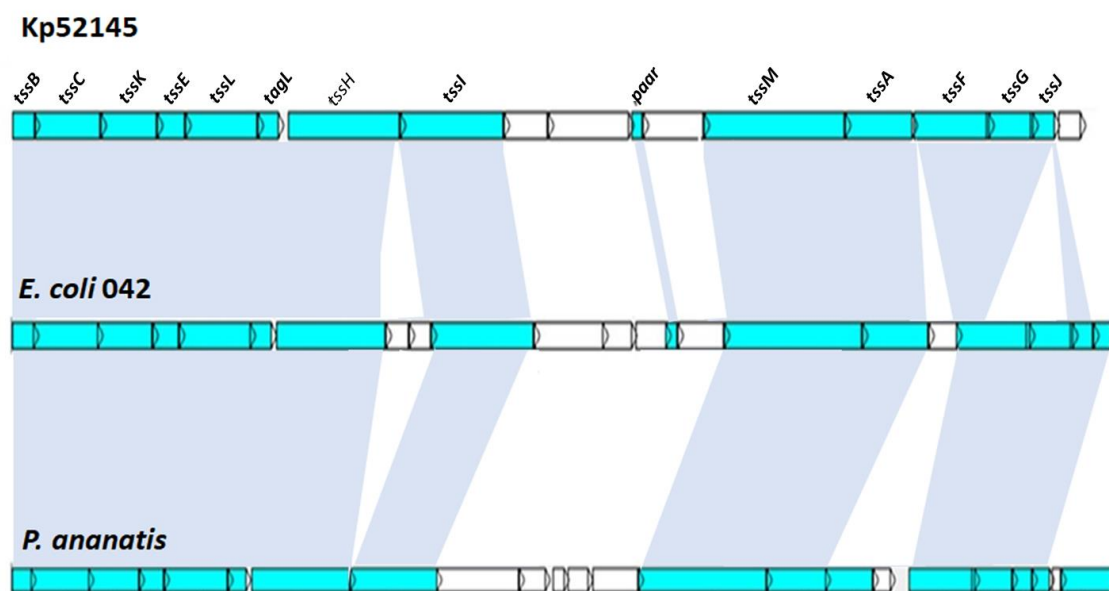


Figura 20: Regiões que codificam genes relacionados com o T6SS em *K. pneumoniae* Kp52.145, *E. coli* E042 e *P. ananatis* LMG106. Os genes relacionados com T6SS estão representados em setas azuis, genes não relacionados ao T6SS estão em branco. Regiões de sintenia entre as sequências são ressaltadas por blocos em azul claro.

Conforme já mencionado, os genes que codificam os componentes do T6SS são frequentemente agrupados em regiões denominadas ilhas de patogenicidade (Cascales 2008; Chen et al. 2015). Esse tipo de arquitetura genômica pode ser reflexo de um evento evolutivo de aquisição desses blocos de genes e pode conferir uma vantagem funcional. Os genes bacterianos necessários para o mesmo processo funcional são frequentemente agrupados na mesma região genômica e submetidos aos mesmos processos de regulação transcricional (Goldman et al. 2015; Barbosa e Lery 2019).

Além de analisar a sequência de cada gene individualmente, analisamos as regiões genômicas que codificam os genes do T6SS em Kp52.145 (Kp52_R1; Kp52_R2; Kp52_R3), HS11286 (HS_R1; HS_R2) e NTUH-K2044 (NT_R1; NT_R2). A comparação das sequências dessas grandes regiões do T6SS das três cepas, evidenciou sintenia entre algumas dessas regiões. De forma geral, encontramos dois *loci* sintênicos presentes nas três cepas. Um deles codifica os genes *tssK-tssL-tagL-tssI* e *paar-tssM-tssF-tssG-tssJ-tssE-tagA* e está representado na Figura 21A. O outro contém os genes *tssB-tssC-tssK-tssL-tagL-tssD-tssH-tssI* e *tssM-tssA-tssF-tssG-tssJ*, conforme mostrado na Figura 21B.

Cada uma dessas regiões varia em tamanho entre 24 e 30 kb. Os genes que codificam componentes do T6SS são codificados no início e no final dessas regiões (marcados em azul na Figura 21), enquanto as regiões centrais contêm genes que codificam proteínas com outras funções (representados em branco).

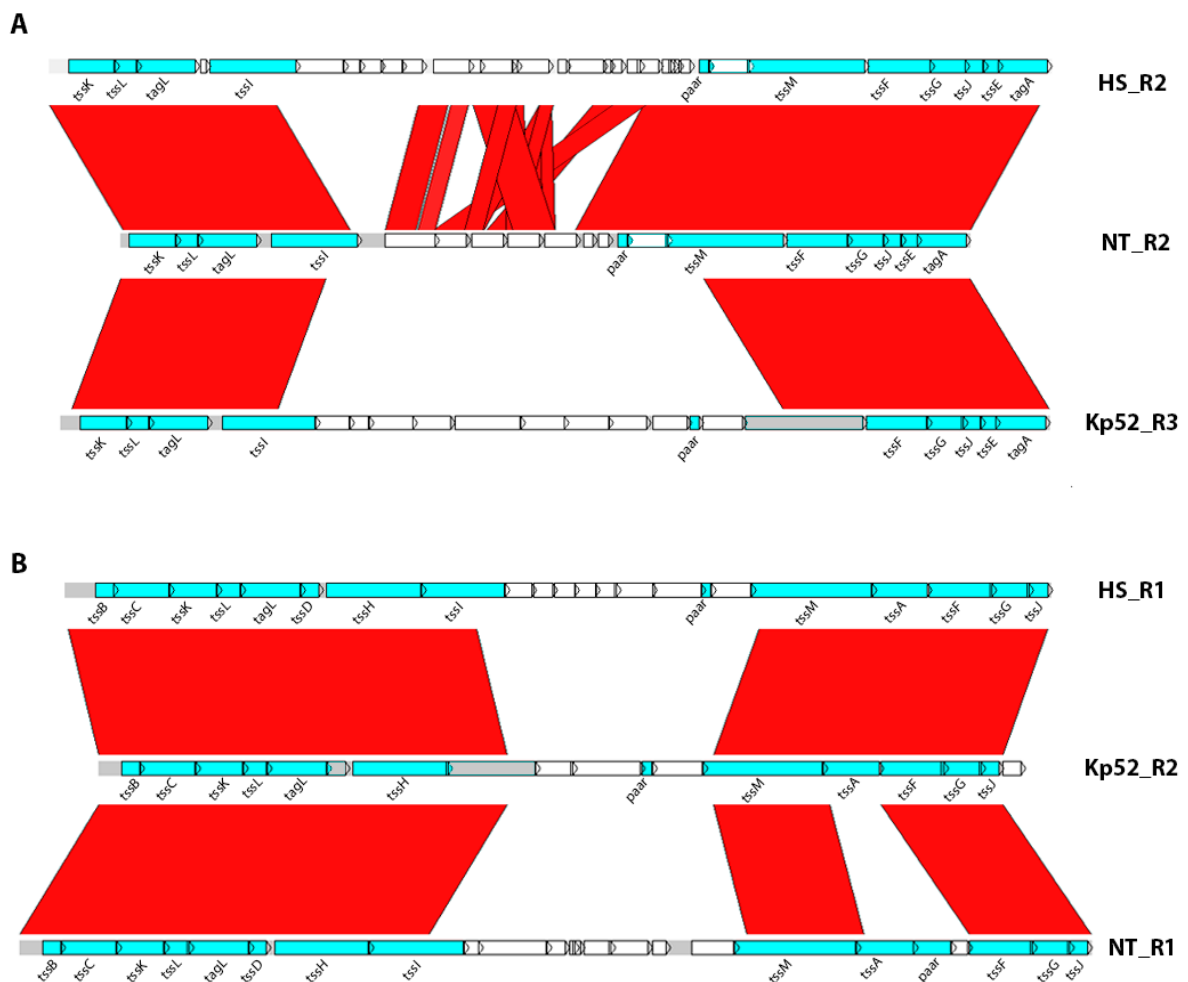


Figura 21: Análise da sintenia entre as regiões que codificam genes relacionados com o T6SS em *K. pneumoniae* Kp52.145 (Kp52), NTUH-K2044 (NT) e HS11286 (HS). As regiões genômicas foram denominadas R1, R2 e R3, de acordo com suas localizações genômicas. Os genes relacionados com T6SS estão representados em setas azuis, genes não relacionados ao T6SS estão em branco. Pseudogenes são exibidos em cinza. Regiões de sintenia (> 86% identidade) entre as sequências são ressaltadas por blocos vermelhos.

Uma característica interessante em ambas as regiões é a presença de uma “região variável”, no meio dessas regiões. Estas “regiões variáveis” são menos conservadas e específicas de cada cepa, codificando principalmente proteínas de função desconhecida ou transposases. As regiões HS_R2 e NT_R2 compartilham regiões curtas de similaridade, variando de 87 a 96% de identidade. A inserção observada em Kp52_R3 inclui genes que codificam fosfolipases, lipoproteínas Sel-1 e uma proteína PAAR. Essa região foi previamente caracterizada como implicada na virulência de *K. pneumoniae* (Lery et al. 2014).

Esse tipo de arquitetura, onde inserções variáveis em *loci* conservados codificando genes do T6SS, tem sido descrito em algumas espécies de bactérias. Em algumas inserções foram encontradas genes que codificam proteínas de função desconhecida, mas também a presença de genes codificadores da transglicosilase lítica, reguladores transcricionais do tipo LysM e proteínas efetoras e de imunidade (Sarris et al. 2011; Shyntum et al. 2014; Journet e Cascales 2016; Liu et al. 2017).

No *locus* II que codifica os componentes do T6SS na Kp52.145, identificamos um gene que codifica uma sequência de aminoácidos com o domínio DUF3274 e α / β - hidrolase (BN49_RS14065) e no *locus* I, identificamos uma sequência com o domínio DUF4150 (BN49_RS06020). Em *P. aeruginosa* PAO1, o produto do gene PA0260 possui domínios DUF3274 e α / β - hidrolase sendo relacionado como membro da família Tle3, cuja produção - direcionada para o periplasma - é tóxica para *E. coli* (Wood et al. 2019). O domínio DUF4150 está associado, em *A. tumefaciens*, ao efector Tde2, sendo necessário para adaptar ou conectar a proteína à TssI (Ma et al. 2014).

Toxinas antibacterianas são codificadas adjacentes a proteínas de imunidade (anti-toxina) que neutralizam o efector em células doadoras e irmãs para evitar autoeliminação. Recentemente genes presentes nas regiões variáveis do T6SS em Kp52.145, foram nomeados de *kpe1-7* (*Klebsiella pneumoniae effector 1-7*) e *kpi1-4* (*Klebsiella pneumoniae immunity 1-4*). O gene BN49_RS06020 foi denominado *kpe3*, vizinho ao gene *kpe4* e a uma proteína de imunidade *kpi1* (Storey, A. McNally 2019). Em Kp52.145, já havia sido identificado que a proteína da família PLD, poderia ser considerado uma Tle5 (Lery et al. 2014). No *locus* III do T6SS, já foram identificados três genes que codificam PLDs (BN49_RS29985, BN49_RS29980 e BN49_RS18775). Nas regiões adjacentes a estes genes são encontrados cinco genes que codificam lipoproteínas com domínio Sel1. Estas lipoproteínas são essenciais em *Legionella pneumophila* para invasão de células hospedeiras (Newton et al. 2007) e

mais recentemente foi demonstrado que a Sel1E de Kp52.145 atua inibindo a toxina TssI4 em *K. pneumoniae* Kp52.145, atuando como uma proteína de imunidade (Storey, A. McNally 2019).

Para aprofundar a caracterização dessas regiões sinténicas, avaliamos o conteúdo de GC (% de bases Guanina e Citosina numa sequência nucleotídica). O conteúdo de GC nos genomas bacterianos é característico de cada espécie, variando entre 20 e 70% (Bohlin et al. 2018). Regiões adquiridas por mecanismos de transferência horizontal apresentam conteúdo de GC característico da espécie doadora, e assim, podem ser identificadas por este tipo de análise. Para esta análise, avaliamos os trechos conservados e variáveis de cada região sinténica, de forma independente. De forma geral, observamos que o conteúdo GC dessas regiões é diferente da composição média do genoma de cada cepa, como podemos visualizar na Tabela 7.

Tabela 7: Conteúdo de GC das regiões genômicas do T6SS em Kp52.145, HS11286 e NTUH-K2044.

Cepa	Locí	Região	Local (<i>locus_tag</i>)	GC%
Kp52.145			Genoma	57,4
	Locus I	<i>tssE – tssI</i>	05940 – 06025	56,85
		RV	06020 – 06005	44,14
		<i>tssM – tssK</i>	06000 – 06040	53,37
	Locus II	<i>tssB – psg</i>	14020 – 14055	60,85
		RV	14055 – 29060	47,51
		<i>tssM – tssJ</i>	14075 – 14095	59,32
	Locus III	<i>tagA – psg</i>	18715 – 18740	61,31
		RV	18740 – 29985	54,91
		<i>tssI – tssK</i>	18800 – 18815	60,26
HS11286			Genoma	57,48
	Locus I	<i>tssB – tssI</i>	22970 – 23040	61,57
		RV	23050 – 23130	58,46
		<i>tssM – tssJ</i>	23140 – 23190	60,17
	Locus II	<i>tssK – tssI</i>	32770 – 32730	60,35
		RV	32510 – 32720	43,91
<i>tssM – tagA</i>		32500 – 32450	61,47	
NTUH_K2044			Genoma	57,68
	Locus I	<i>tssB-tssI</i>	11220 - 11255	60,86
		RV	11260 - 11290	41,0
		<i>tssM – tssJ</i>	11295 - 11325	56,1
	Locus II	<i>tssK – tssI</i>	15790 - 15775	60,61
		RV	15770 - 15720	44,31
<i>tssM – tagA</i>		15715 - 15690	61,53	

Legenda: RV – Região variável; *psg*– pseudogene.

As regiões contendo os genes do *core* do T6SS possuem conteúdo GC maior que a média do percentual GC dos genomas completos das respectivas cepas - chegando a ser maior que 61%. Ao contrário, as regiões variáveis apresentam baixo conteúdo de GC, variando entre 41% e 54%. Esses dados reforçam que as regiões variáveis tenham sido adquiridas em eventos distintos da aquisição das ilhas do T6SS e, possivelmente, provenientes de organismos diferentes.

4.6 Regiões promotoras

4.6.1 Sítios de ligação de fatores sigma σ

Visando a caracterização do mecanismo de regulação transcricional, precisamos identificar sequências promotoras da expressão gênica. Sequências promotoras são regiões que contém sítios para ligação de proteínas, como a RNA polimerase e reguladores transcricionais, que controlam a posição e o momento do início da transcrição dos genes sob controle deste promotor (Gan et al. 2009, 2012). Para obter informações sobre mecanismos de regulação transcricional do T6SS em *K. pneumoniae*, inicialmente identificamos as sequências de ligação dos fatores σ^{54} (-12 e -24) e σ^{70} (-10 e -35) a montante das CDSs que codificam os genes relacionados ao T6SS. Para que o processo de transcrição gênica seja iniciado é necessário o reconhecimento de elementos dos promotores, localizados a montante do início de transcrição, pela holo RNA polimerase (Murakami e Darst 2003). Um fator preponderante na ligação da RNA polimerase ao DNA é o fator σ , atuando no reconhecimento do promotor. Em cada bactéria, diversos fatores σ estão disponíveis e, cada um deles, regula a expressão de um conjunto de genes diferentes. Em *E. coli*, o fator sigma primário necessário para iniciar a transcrição, na maioria dos genes é chamado σ^{70} , este mesmo fator está associado à transcrição de genes constitutivos (Goldman et al. 2015; Narayanan et al. 2018). Até o presente momento não há relatos experimentais na literatura de genes do T6SS sendo diretamente regulados pelo fator σ^{70} . Entretanto, por ser um fator que está associado à regulação global em bactérias, acreditamos que ele possa estar envolvido na regulação da transcrição do T6SS. Por outro lado, em *V. cholerae* foi descrito que o fator σ^{54} atua na regulação do gene *tssD* e, em *P. aeruginosa* PAO1, este fator estaria participando da regulação de operons do T6SS. Por meio de ensaios de gene-repórter foi mostrado que agrupamentos gênicos do T6SS em *V. cholerae*, *A. hydrophila*, *Pectobacterium atrosepticum* e *Marinomonas* sp. MWYL1 tem sua expressão controlada por σ^{54} e *bEBP* (Bernard

et al. 2011; Kitaoka et al. 2011; Sana et al. 2013). Assim, neste estudo focamos nas regiões promotoras dependentes de σ^{70} e σ^{54} .

No total, 17 sítios de início de transcrição dependentes de σ^{70} foram encontrados em Kp52.145, 12 em HS11286 e 13 em NTUH-K2044. Para o fator σ^{54} foram encontrados 17 promotores em Kp52.145, 12 em HS11286 e 15 em NTUH-K2044.

Tendo predito sítios putativos para ligação de fatores σ^{70} e sítios putativos para o fator σ^{54} , anotamos as sequências a 250 pb a montante dos sítios de início de transcrição preditos, como possíveis sequências promotoras. No total anotamos 34, 24 e 28 sequências promotoras nas cepas Kp52.145, HS11286 e NTUH-K2044, respectivamente (Apêndice D). Utilizamos essas sequências para pesquisar possíveis sítios de ligação de reguladores transcricionais.

4.6.2 Reguladores transcricionais codificados nos genomas de *K. pneumoniae*

Neste estudo, utilizamos o banco de dados do Virtual Footprint para obter informações sobre sítios de ligação de reguladores transcricionais ao DNA. Esse banco de dados compreende informações sobre 116 reguladores de oito espécies. Considerando que nem todos esses reguladores estão presentes em *K. pneumoniae*, precisamos inicialmente identificar quais dos reguladores que tem seus sítios descritos no Virtual Footprint são codificados nos genomas de *K. pneumoniae*. Essa análise realizada pelo BLASTp, resultou na identificação de 44 reguladores a serem usados (Tabela 8).

Tabela 8: Reguladores transcricionais avaliados neste estudo e informações sobre as matrizes de peso-posição correspondentes, disponíveis no Virtual Footprint.

MPP	Pontuação	Desvio Padrão	Tamanho
AraC	13,59	2,01	29
ArcA	6,30	0,81	10
ArgR	9,12	0,88	14
CpxR	14,04	0,09	16
Crp	8,79	0,26	22
CspA	10,00	0,00	5
CysB	9,68	0,73	31
CytR	7,61	0,31	12
DeoR	19,61	1,00	16
FadR	5,52	1,09	7
FhlA	10,18	0,37	7
Fis	2,87	0,56	15
Fnr	7,48	1,17	14
FruR	12,39	0,83	12
Fur (18mer)	28,50	0,00	18
GalR	18,12	0,30	13
GalS	19,14	0,75	14
GcvA	10,00	0,00	5
GlnG	18,53	0,64	17
GlpR	6,23	0,25	20
GntR	10,35	1,86	16
H-NS	5,56	0,65	10
IHF	6,23	1,24	16
IciA	5,03	0,57	12
IlvY	28,50	0,00	27
LexA	11,45	0,99	16
Lrp (SELEX)	11,16	0,73	25
Mall	28,00	0,00	15
MalT	8,02	0,66	6
MarR	32,50	0,00	23
MetJ (SELEX)	17,00	1,78	16
MetR	8,88	0,42	7
Mlc (SELEX)	18,53	3,01	21
ModE	27,65	0,64	23
Nac	9,88	0,62	11
NagC	15,02	0,71	16
NagC (SELEX)	21,11	1,29	23
NarL	3,77	0,52	7
NharR	12,09	0,79	11
OmpR C box	12,36	0,49	10
OmpR F box	13,47	0,49	10
OmpR	8,22	0,52	7
OxyR	3,17	0,58	11
OxyR (SELEX)	18,17	4,92	46
PdhR	6,42	0,59	6
PhoP	18,42	1,07	17
PurR	13,14	0,60	10
RcsAB	5,02	0,78	8

Legenda: MPP – Matriz de Peso Posição.

4.6.3 Identificação de sítios de ligação de reguladores transcricionais

Numa análise inicial, dos três genomas de *K. pneumoniae*, realizamos buscas por sítios putativos nos promotores dependentes de σ^{70} para 36 reguladores transcricionais na Kp52.145, 20 na HS11286 e 38 na NTUH-K2044. Na cepa Kp52.145, 106 sítios de ligação para 13 reguladores transcricionais foram preditos: Crp, CytR, FhIA, Fis, Fnr, GcvA, H-NS, MalT, MetR, Mlc, OmpR, OxyR e RcsAB. Em HS11286, encontramos 72 sítios de ligação para nove reguladores transcricionais: Fis, GcvA, H-NS, Lrp, NarL, OmpR, OxyR, PdhR e RcsAB. Na NTUH-K2044, encontramos 114 sítios de ligação para 17 reguladores transcricionais: Crp, FadR, FhIA, Fis, Fnr, GcvA, GlpR, H-NS, Lrp, MalT, MetJ, MetR, Mlc, OmpR, OxyR, PdhR e RcsAB (Barbosa e Lery 2019).

Analisando sítios de reguladores transcricionais nos promotores dependentes de σ^{70} , observamos que seis reguladores transcricionais possuem sítios em promotores de genes do T6SS das três cepas: Fis, OxyR, H-NS, OmpR, GcvA e RcsAB. A posição desses sítios está indicada na Figura 22. De todos os promotores analisados, ~85% possuem sítios de ligação de Fis e ~75% sítios de ligação de OxyR.

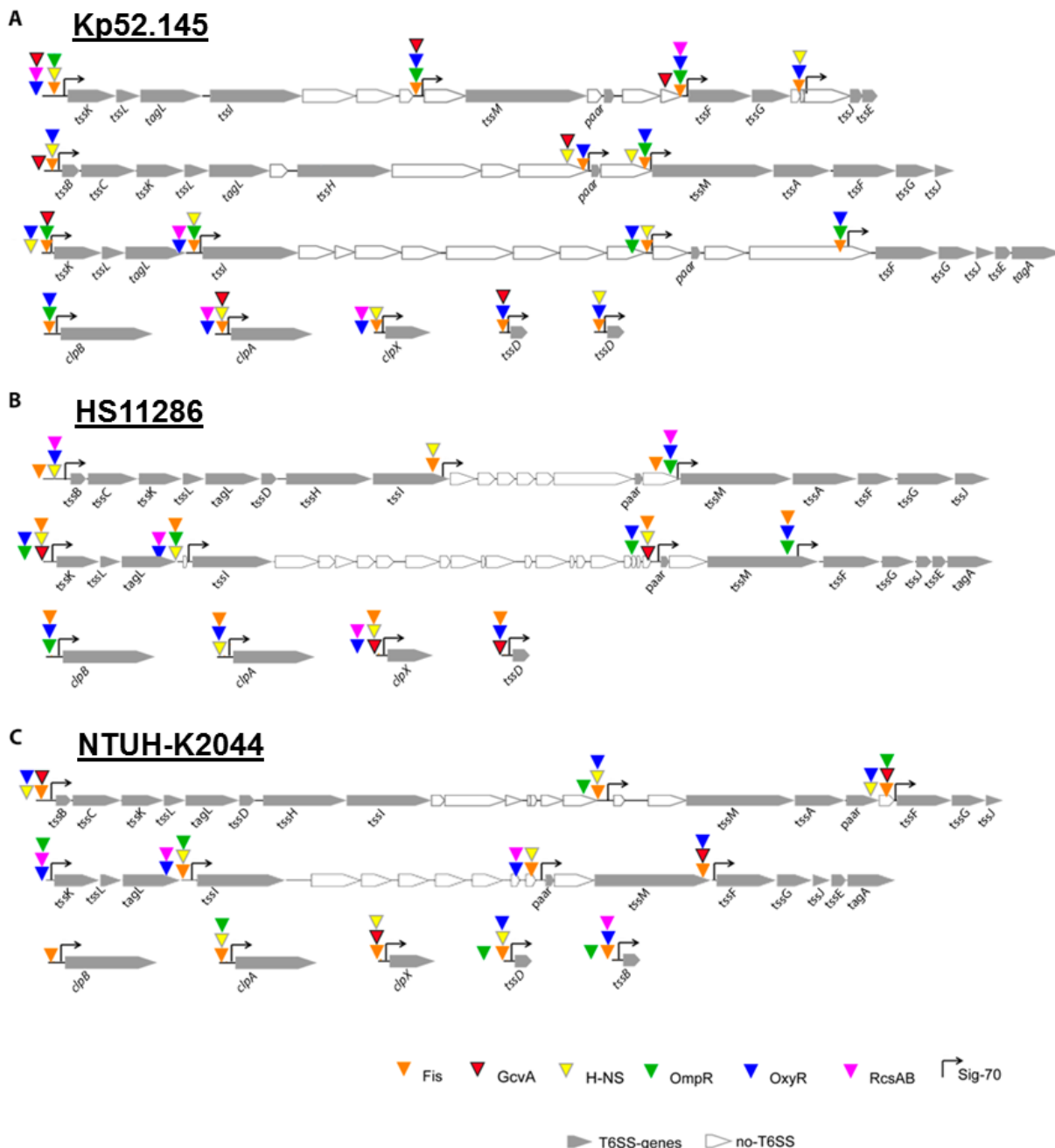


Figura 22: Organização de genes de T6SS, sítios de início da transcrição nos promotores dependentes de σ^{70} e sítios putativos de ligação de reguladores transcricionais em genomas de *K. pneumoniae* Kp52.145 (a), HS11286 (b) e NTUH-K2044 (c). Os genes de T6SS são representados como setas cinzas. Triângulos coloridos apontam a posição dos seis reguladores transcricionais com mais sítios de ligação previstos (Barbosa e Lery 2019).

Até agora, o T6SS de *K. pneumoniae* havia sido associado à atividade antibacteriana, invasão celular e colonização *in vitro* (Liu et al. 2017; Hsieh et al. 2018).

A análise do contexto genômico descrita aqui apontou que vários transportadores relacionados ao ferro estão codificados em torno dos genes do T6SS. Esta observação levanta a hipótese de que o T6SS de *K. pneumoniae* poderia desempenhar um papel na absorção de íons. Por exemplo, o T6SS de *Y. pseudotuberculosis* transporta Zn^{2+} para combater múltiplos estresses e imunidade do hospedeiro (Wang et al. 2015). *B. thailandensis* utiliza o T6SS para absorver íons Mn^{2+} , para resistir ao estresse oxidativo e competir com outras bactérias (Si et al. 2017b). E mais recentemente em *K. pneumoniae* Kp52.145, foi sugerido que na presença de quelante de ferro, ocorre regulação positiva da expressão do *locus* II do T6SS (Storey, A. McNally 2019).

Este T6SS de *B. thailandensis* é regulado por OxyR, um regulador transcricional de resposta ao estresse oxidativo. As espécies reativas de oxigênio (EROs) podem danificar as células bacterianas, portanto, as bactérias desintoxicam as EROs ao produzir enzimas antioxidantes para prevenir lesões no DNA e permitir a formação de compostos que sequestram íons metálicos. Um desses mecanismos envolve catalases. Curiosamente, o OxyR regula a expressão de KatN, uma catalase secretada de maneira dependente de T6SS pela *E. coli* enterohemorrágica (Wan et al. 2017). OxyR é um dos reguladores que induzem a resposta do estresse oxidativo bacteriano. Além disso, foi demonstrado que o OxyR de *K. pneumoniae* regula a formação de biofilme, genes fimbriais, resistência a antibióticos e adesão a células epiteliais (Hennequin e Forestier 2009; Rajamohan et al. 2013). Curiosamente, previmos sítios de ligação de OxyR conservados na maioria das regiões promotoras de σ^{70} do T6SS de *K. pneumoniae*, em todas as três cepas analisadas. Em conjunto, esses dados sugerem que o T6SS de *K. pneumoniae* pode ser ativado sob condições de estresse oxidativo e pode ajudar as bactérias a lidar com a desintoxicação de EROs.

Outra constatação interessante foi que os locais de ligação de H-NS putativos estão amplamente presentes e conservados nos promotores do T6SS de *K. pneumoniae*. Foi demonstrado anteriormente que o silenciamento por H-NS de um *locus* T6SS limita a morte interbacteriana de *S. enterica* (Brunet et al. 2015a). H-NS é uma proteína de estruturação do nucleóide, com efeitos globais no silenciamento da expressão gênica (Fang e Rimsky 2008). Sua atividade depende da temperatura e da osmolaridade. Tem sido demonstrado que a H-NS de *K. pneumoniae* reprime a expressão de importantes fatores de virulência, como o *pili* tipo 3 e a cápsula (Ares et al. 2016, 2017). Recentemente, Hsieh e cols. (2018) (Hsieh et al. 2018) mostraram

que H-NS se liga ao promotor de *tssD* e silencia a expressão de *tssD* (KP1_RS11245) em NTUH-K2044. Assim, nossa previsão foi confirmada por esses dados experimentais recentes.

Locais de ligação putativos para quatro reguladores adicionais foram previstos nas sequências promotoras σ^{70} dos genes T6SS de todas as três cepas: RcsAB, GcvA, Fis e OmpR. O RcsAB é um sistema regulatório não usual que se liga a uma caixa 'rcsAB' e modula o gene *galF* em *K. pneumoniae*, afetando a expressão e a virulência da cápsula (Peng et al. 2018; Su et al. 2018). O GcvA é o regulador transcricional do sistema de clivagem da glicina, envolvido no metabolismo dos aminoácidos. GcvA ainda não foi estudado em *K. pneumoniae*, no entanto, é necessário para o metabolismo de *F. tularensis* e sua virulência (Brown et al. 2014). Fis (fator de estimulação por inversão) é um regulador transcricional que responde às mudanças no ambiente nutricional em enterobactérias (Mallik et al. 2004). OmpR é o regulador de resposta de um sistema de dois componentes com o sensor kinase EnvZ. OmpR liga-se à região promotora do T6SS de *Y. pseudotuberculosis*, envolvido na sobrevivência bacteriana em condições de alta osmolaridade, presença do desoxicolato de sódio e homeostase do pH (Gueguen et al. 2013; Zhang et al. 2013). Tem sido demonstrado que em *K. pneumoniae* a OmpR regula a via de sinalização do c-di-GMP, a expressão da fímbria do tipo 3 e a formação de biofilme em resposta a estresses osmóticos (Lin et al. 2018b). Em conjunto, estes resultados compreendem dados bastante interessantes, sugerindo que o T6SS de *K. pneumoniae* pode ser regulado em resposta a sinais ambientais que são de fato detectados pelas bactérias dentro do hospedeiro humano: temperatura (H-NS), limitação nutricional (Fis), estresse oxidativo (OxyR) e osmolaridade (RscAB e OmpR) (Tabela 9). Surpreendentemente, os reguladores H-NS, OxyR e OmpR são expressos pela *K. pneumoniae* BAA2146 durante a infecção *in vitro* de macrófagos (Bent et al. 2016).

Tabela 9: Características de reguladores transcricionais preditos como possíveis reguladores do T6SS de *K. pneumoniae*.

Regulador	Atividade	Ambiente de ativação
ArcA	Ativador / Repressor	Anaerobiose (Boulette e Payne 2007)
FhIA	Ativador	Anaerobiose (Rossmann et al. 1991)
Fis	Ativador / Repressor	Resposta estrigente (Walker et al. 1999)
GcvA	Ativador	Disponibilidade de AA (Brown et al. 2014)
H-NS	Repressor	Osmolaridade (Hurme e Rhen 1998)
IHF	Ativador	-
OmpR	Ativador	Osmolaridade
OxyR	Ativador	Estresse oxidativo (Storz e Imlay 1999)
RcsAB	Ativador	Estresse osmótico

Legenda: AA - Aminoácidos

Estes dados foram publicados na revista BMC Genomics, em 2019 (Barbosa e Lery 2019). Após essa publicação dos resultados iniciais, ampliamos as buscas por sítios de ligação putativos de reguladores transcricionais para os reguladores identificados nos três genomas. Inicialmente, em Kp52.145, identificamos 349 possíveis sítios de ligação, para 31 matrizes (29 reguladores), nas sequências promotoras de σ^{70} e 310 sítios para 28 matrizes (29 reguladores) nas sequências promotoras dependentes de σ^{54} . Na HS11286, identificamos 235 possíveis sítios de ligação, para 29 matrizes (27 reguladores), nas sequências promotoras dependentes de σ^{70} e 224 sítios para 28 matrizes (26 reguladores) nas sequências promotoras dependentes de σ^{54} . Na cepa NTUH-K2044, identificamos 261 possíveis sítios de ligação, para 30 matrizes (27 reguladores), nas sequências promotoras de σ^{70} e 307 sítios para 31 matrizes (27 reguladores) nas sequências promotoras dependentes de σ^{54} .

Contudo, sabemos que a predição de sítios de ligação pode gerar alto percentual de falsos positivos, devido às sequências serem curtas. Para minimizar este problema, utilizamos o *score* somado ao desvio padrão de cada matriz de peso posição, para filtrar os resultados encontrados. Ao utilizarmos esse valor como valor de corte, muitos dos sítios identificados inicialmente não apresentaram pontuação

significativa. Assim, dos 235 e 224 sítios putativos identificados na Kp52.145, nos promotores de σ^{70} e σ^{54} , passamos a trabalhar com 82 sítios, tanto nos promotores de σ^{70} , quanto nos de σ^{54} . Na HS11286, recuperamos 64 e 59 sítios preditos e na NTUH-K2044 recuperamos 56 e 77 sítios preditos (Apêndice L).

Ao todo, foram identificados sítios para 20 possíveis reguladores transcricionais, sendo que 15 reguladores são comuns às três cepas (Tabela 10).

Tabela 10: Número de sítios de ligação putativos dos reguladores transcricionais identificados nos promotores de σ^{70} e σ^{54} em *K. pneumoniae*.

MPP	Kp52.145		HS11286		NTUH-K2044	
	S70	S54	S70	S54	S70	S54
AraC		1				
ArcA	2	6	4	4	1	4
ArgR	2	1	1	1	1	1
CysB				1		
CytR	3	7	1	7		6
FhIA	1	3		2		4
Fis	5	2	9	2	4	5
Fnr	1	2	2	1	1	
GcvA	8	9	3	6	6	9
GlpR	2	2	2	1	4	2
GntR					1	1
H-NS	14	11	8	7	11	8
IHF	6	5	5	4	2	4
IciA	11	13	9	10	8	11
Lrp (SELEX)	1	1		1		1
Nac	6	3	3	1	5	2
NagC (SELEX)			1		1	1
OmpR F box		2		1		2
OmpR	1	1	1	1		
OxyR	6	3	6	2	3	5
RcsAB	13	11	9	8	8	12
Total	82	82	64	61	56	78

Legenda: MPP – Matriz de Peso Posição.

Pelas análises realizadas neste trabalho, um dos promotores que contém sítio putativo para OxyR é o do gene órfão *tssD* (BN49_RS06500 e KPHS_41670) vizinho aos genes que codificam proteínas relacionadas ao transporte de íons, nas cepas Kp52.145 e HS11286. Em *B. thailandensis*, o aumento dos níveis de espécies reativas de oxigênio intracelular pode atuar como mensageiro para ativar o regulador OxyR, sensível a alterações do potencial redox. Uma vez ativado, o regulador OxyR induz a expressão de uma ampla gama de sistemas antioxidantes, incluindo T6SS4 de *B. thailandensis* e *Y. pseudotuberculosis* (Wang et al. 2015; Si et al. 2017a). Nestas análises dos promotores de σ^{70} , encontramos 15 sítios para o regulador transcricional

OxyR, sendo que na Kp52.145 identificamos que estes sítios estão concentrados no terceiro *locus* do T6SS, bem como no promotor de *tssD* (BN49_RS0500) órfão, do promotor de uma ATPase ClpX e do promotor do gene *mItE*, conforme mostra a Figura 23.

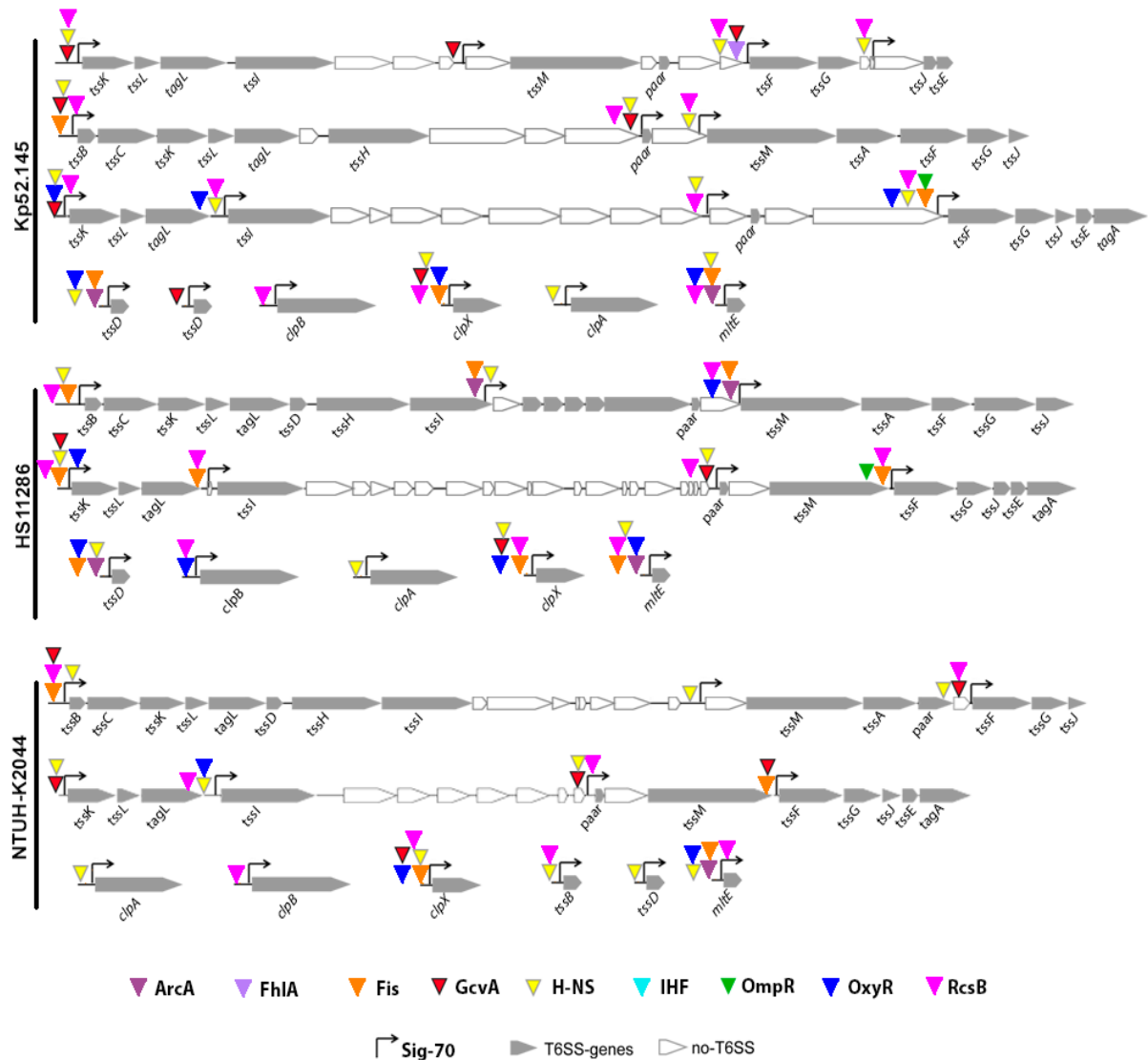


Figura 23: Sítios de início da transcrição dependentes de σ^{70} e locais putativos de ligação de reguladores transcricionais em genomas de *K. pneumoniae* Kp52.145 (a), HS11286 (b) e NTUH-K2044 (c). Os genes de T6SS são representados como setas cinzas. Triângulos coloridos apontam a posição dos seis reguladores transcricionais com mais sítios de ligação previstos (Barbosa e Lery 2019).

Considerando que Fis e H-NS se ligam a regiões ricas em AT, seria possível ocorrer sobreposição nos seus perfis e sítios de ligação ao DNA (Grainger et al. 2006). Para verificar se isto pode acontecer, alinhemos todos os 27 sítios de Fis preditos em *K. pneumoniae*, bem como todos os 36 sítios de ligação de H-NS. Comparando a conservação das sequências de ligação de cada um desses reguladores, podemos observar que os sítios preditos apresentam perfis diferentes (Figura 24).

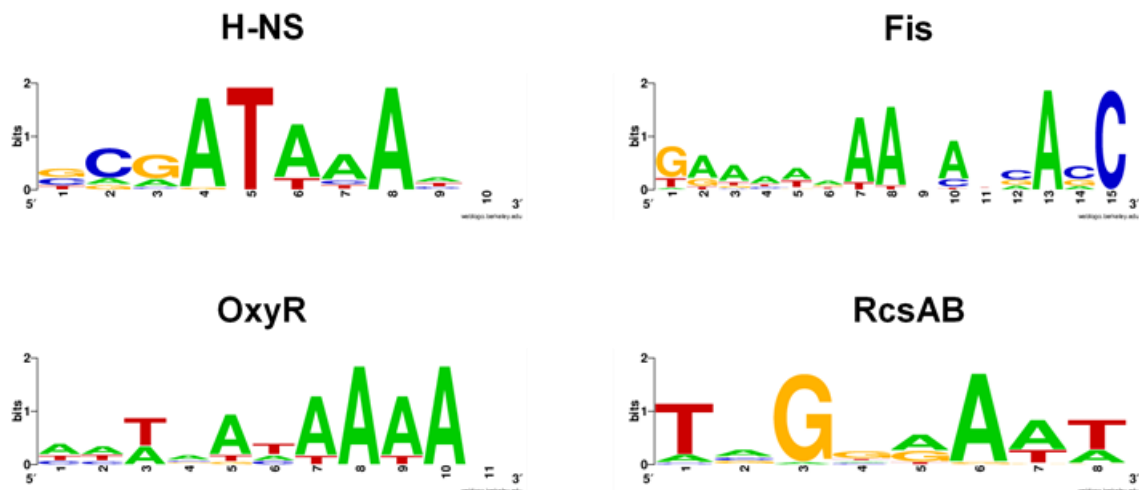


Figura 24: Logotipos dos sítios de ligação preditos para os reguladores transcricionais H-NS, Fis, OxyR e RcsAB nos promotores σ^{70} de *K. pneumoniae* Kp.52.145, HS11286 e NTUH-K2044.

Observando esta figura, é perceptível que nas representações dos sítios putativos de H-NS e Fis o nucleotídeo adenina é predominante, o que explicaria a sobreposição de alguns destes sítios nas sequências promotoras em nossos resultados. Destacamos ainda que a representação do regulador OxyR também possui significativa presença de adenina.

Enquanto nos promotores de σ^{70} , identificamos 35 sítios para o regulador transcricional OxyR e 22 sítios para o regulador OmpR, ao analisar os sítios destes reguladores nos promotores dependentes de σ^{54} , observamos apenas dois sítios para o regulador OmpR, e apenas 10 para o OxyR. Isso pode implicar em menor influência das condições de estresse oxidativo sobre a expressão dos genes do T6SS. Assim como nos promotores de σ^{70} , encontramos diversos sítios para os reguladores Fis, GcvA, H-NS e RcsAB nos promotores de σ^{54} (Figura 25).

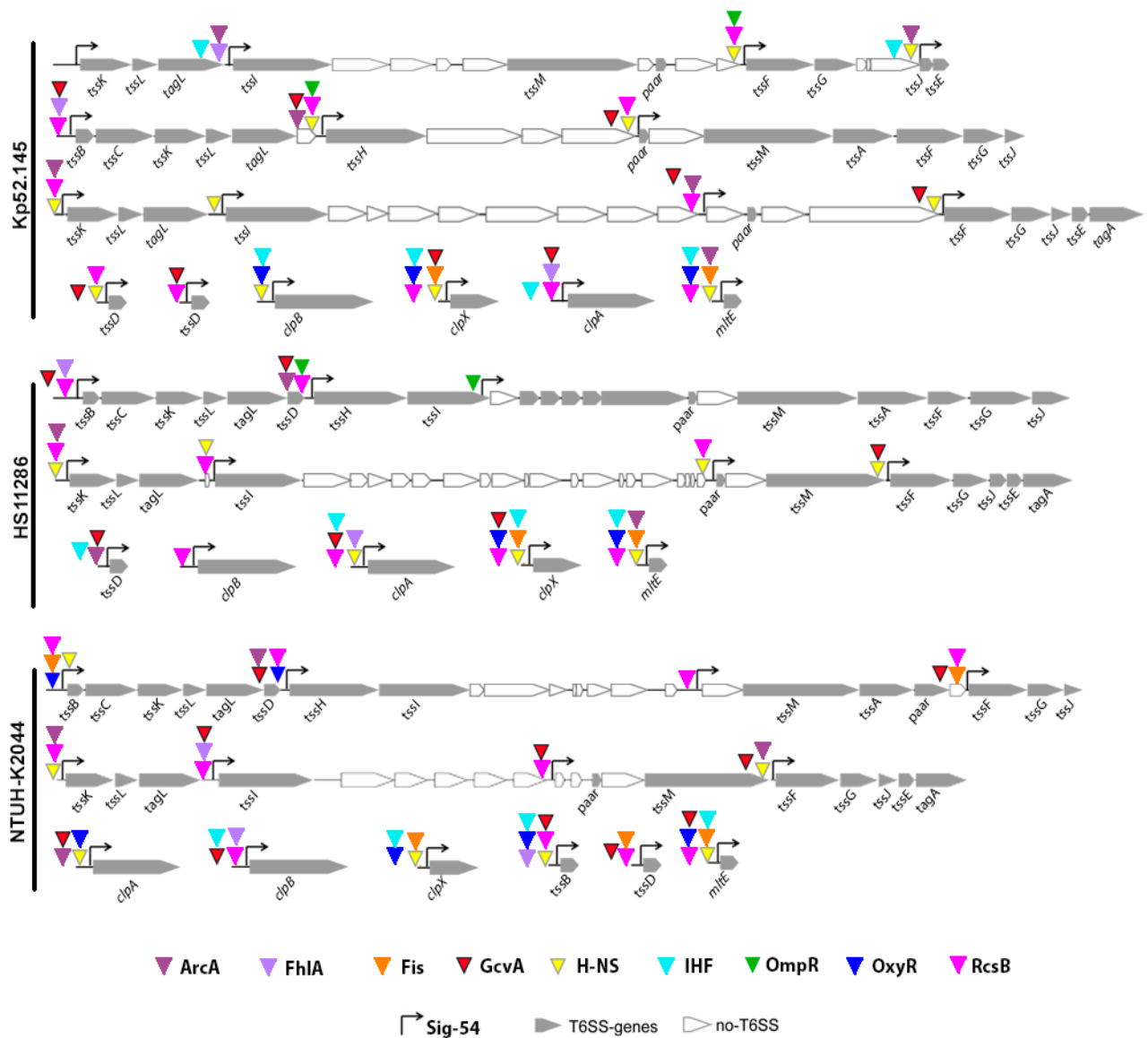


Figura 25: Sítios de início da transcrição dependentes de σ^{54} e locais putativos de ligação de reguladores transcricionais em genomas de *K. pneumoniae* Kp52.145 (a), HS11286 (b) e NTUH-K2044 (c). Os genes de T6SS são representados como setas cinzas. Triângulos coloridos apontam a posição dos seis reguladores transcricionais com mais sítios de ligação previstos (Barbosa e Lery 2019).

Nestes promotores dependentes de σ^{54} , encontramos também um número significativo de sítios para os reguladores ArcA, FhIA e IHF. Em contraste com o início da transcrição mediada por σ^{70} , a transcrição dependente de σ^{54} necessita de um ativador que acopla a energia gerada a partir da hidrólise de ATP, esses ativadores são conhecidos como *bEBPs* (Schumacher et al. 2004; Rappas et al. 2007). Neste contexto, a proteína IHF (*Integration Host Factor*) auxilia a formação de uma volta de 180° no DNA, aproximando a *bEBP* da RNA polimerase (Bush e Dixon 2012). Em nossas análises, identificamos nos promotores de σ^{54} , 13 sítios de ligação putativos para o fator IHF. O regulador FhIA (*Formate hydrogen lyase Activator*) também é regulado por alterações na concentração de oxigênio, é um dos reguladores

transcricionais que são *bEBPs* (Bush e Dixon 2012). O regulador FhIA é um ativador que exibe atividade de ATPase (Hopper e Böck 1995) e de ativador da transcrição dos genes que codificam o sistema de eliminação de espécies reativas de oxigênio em *E. coli* (Leonhartsberger et al. 2000). A presença dos sítios de ligação de IHF e FhIA nos promotores de σ^{54} , seria um indício que componentes do T6SS de *K. pneumoniae* podem ser regulados quando em condições de anaerobiose.

Diferente do regulador OxyR, o fator transcricional ArcA (*Aerobic respiratory control*), um sistema de dois componentes, responde à baixa disponibilidade de oxigênio. O sistema ArcAB, atua como um regulador global controlando a expressão de genes em bactérias patogênicas como *E. coli* e *S. typhimurium* (Green et al. 2014). Em *K. pneumoniae* mutadas para o gene *arcA*, foi observada atenuação na colonização do trato gastrointestinal (Boll et al. 2012) e, assim como o regulador OxyR, ArcA está associado à resistência bacteriana a espécies reativas de oxigênio em *S. enterica* sorovar Enteritidis SE2472 (Boulette e Payne 2007). A presença destes sítios indica que em ambientes de anaerobiose, em *K. pneumoniae*, poderia ter a ativação da transcrição de componentes do T6SS.

É importante destacar que, até o momento, analisamos as sequências promotoras dependentes de σ^{54} e σ^{70} . Provavelmente, outros promotores sigma estão envolvidos na regulação transcricional do T6SS. Hsieh e cols, 2018 (Hsieh et al. 2018) detectaram que, em NTUH-K2044, os genes *tssB-tssC-tssK-tssL-tagL-tssD* formam uma única unidade de transcrição e que os genes *tssH* e *tssI* são transcritos independentemente. Em nossa análise, detectamos o promotor dependente de σ^{70} a montante de *tssB*, mas não *tssH* e *tssI*. Entretanto, a montante do gene *tssH*, identificamos um promotor dependente de σ^{54} com sítios para os reguladores ArcA, GcvA, OmpR e RcsAB.

Neste trabalho identificamos mediante a predição de sítios de reguladores transcricionais, que sinais ambientais encontrados por *K. pneumoniae* no trato gastrointestinal ou quando infectam mucosas, como anaerobiose, estresse osmótico e oxidativo, podem estar relacionados com ativação do T6SS em *K. pneumoniae*. Essa ativação ocorreria pela interação de fatores transcricionais (ArcA, FhIA, H-NS, OmpR, OxyR e RcsAB) a sequências promotoras dos genes que codificam os componentes do T6SS (Figura 26).

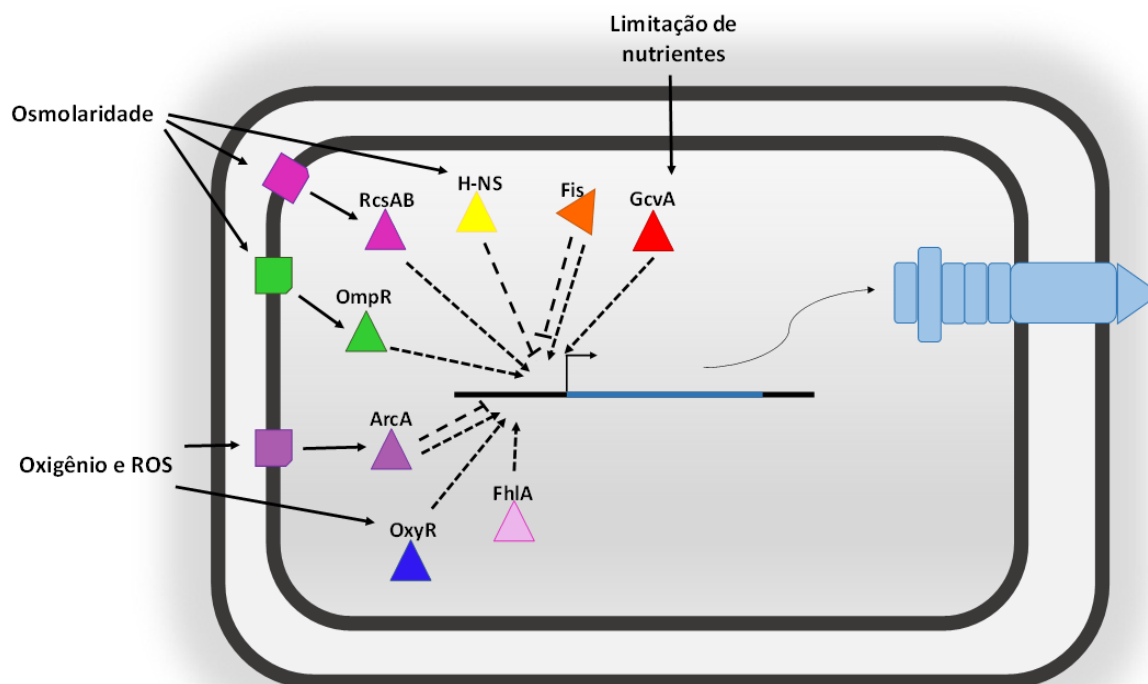


Figura 26: Panorama geral de estímulos específicos e reguladores transcripcionais que podem regular a transcrição do T6SS em *K. pneumoniae*. Representado em formato de quadrado estão os sensores de quinase dos sistemas de dois componentes, os triângulos coloridos representam os reguladores transcripcionais que irão modular a expressão dos genes do T6SS.

De forma muito interessante e complementar a ao nosso estudo, em julho de 2019 foi publicado um trabalho que avaliou a expressão de genes do T6SS na cepa Kp52.145 (Storey, A. McNally 2019). Utilizando o promotor do gene *tssB* (BN49_RS14020) e um gene repórter, os autores concluíram que quanto menor a concentração de oxigênio dissolvido no meio, maior a atividade do T6SS. Eles observaram também que o aumento da força iônica e osmolaridade pela adição de KCl, ramnose ou sacarose diminui a atividade do T6SS. Por outro lado, o aumento da concentração de NaCl no meio, aumentou a atividade do sistema. Ainda, estes autores mostraram que H-NS atua como repressor gênico do T6SS na Kp52.145. Portanto, esses dados recentes concordam e reforçam as previsões realizadas pelo nosso grupo.

De acordo com trabalhos anteriores, nos quais também demonstram que o fator σ^{54} é importante para a transcrição de genes do T6SS (Kitaoka et al. 2011; Sana et al. 2013), identificamos neste trabalho além dos promotores putativos dependentes de σ^{54} , promotores putativos dependentes de σ^{70} , o que torna estes promotores promissores para ensaios futuros com o intuito de verificar a exata posição de ligação dos reguladores transcripcionais aqui preditos e sua relação com o T6SS em *K.*

pneumoniae. Curiosamente, Storey e cols, 2019 mostraram ainda que o T6SS de *K. pneumoniae* é dependente não somente do fator σ^{54} , como também de σ^{38} . Essa análise não foi feita no nosso estudo, mas seria interessante futuramente analisar os sítios de σ^{38} nas 3 cepas analisadas neste estudo e verificar a conservação deste mecanismo regulatório.

5 PERSPECTIVAS

Com os resultados aqui apresentados, temos como perspectivas validar as predições computacionais apresentadas. Estas validações incluem ensaios de ligação dos reguladores transcricionais às sequências promotoras e a verificação da expressão dos genes que codificam os reguladores transcricionais e os componentes do T6SS em diferentes condições como: anaerobiose, estresse oxidativo e alterações da osmolaridade.

6 CONCLUSÕES

Apresentamos uma análise genômica de três cepas de *K. pneumoniae* e fornecemos novas informações sobre o papel e a regulação do T6SS:

- 1) A anotação dos componentes do T6SS foi padronizada entre as cepas,
- 2) Identificamos que a organização destes genes ocorre em blocos em sintonia nas três cepas de *K. pneumoniae*,
- 3) Considerando que as sequências da proteína ClpB de *K. pneumoniae* são bastante conservadas em relação à ClpB de *F. tularensis*, sugerimos que elas podem agir como TssH,
- 4) Foi hipotetizado que o T6SS em *K. pneumoniae* pode desempenhar um papel na captação de íons,
- 5) Nas 44 sequências promotoras de σ^{54} e 42 sequências promotoras de σ^{70} , encontramos diversos sítios de ligação para os reguladores transcricionais ArcA, FhIA, Fis, GcvA, H-NS, IHF, OmpR, OxyR e RcsAB,
- 6) Correlacionamos os reguladores transcricionais com sítios de ligação identificados nos promotores com os estímulos ambientais que possam estimular ou bloquear a atividade de transcrição do T6SS, como alterações na anaerobiose, osmolaridade, resposta estrigente e estresse oxidativo.

7 REFERÊNCIAS BIBLIOGRÁFICAS

- Aguirre JD, Culotta VC. Battles with Iron: Manganese in Oxidative Stress Protection. *J Biol Chem*. 20 de abril de 2012;287(17):13541–8.
- Alcoforado Diniz J, Liu Y-C, Coulthurst SJ. Molecular weaponry: diverse effectors delivered by the Type VI secretion system. *Cell Microbiol*. dezembro de 2015;17(12):1742–51.
- Anderson MC, Vonaesch P, Saffarian A, Marteyn BS, Sansonetti PJ. *Shigella sonnei* Encodes a Functional T6SS Used for Interbacterial Competition and Niche Occupancy. *Cell Host Microbe*. junho de 2017;21(6):769-776.e3.
- Ares MA, Fernández-Vázquez JL, Pacheco S, Martínez-Santos VI, Jarillo-Quijada MD, Torres J, et al. Additional regulatory activities of MrkH for the transcriptional expression of the *Klebsiella pneumoniae* mrk genes: Antagonist of H-NS and repressor. Cascales E, organizador. *PLoS One*. 9 de março de 2017;12(3):e0173285.
- Ares MA, Fernández-Vázquez JL, Rosales-Reyes R, Jarillo-Quijada MD, von Bargen K, Torres J, et al. H-NS Nucleoid Protein Controls Virulence Features of *Klebsiella pneumoniae* by Regulating the Expression of Type 3 Pili and the Capsule Polysaccharide. *Front Cell Infect Microbiol*. 9 de fevereiro de 2016;6(February).
- Aschtgen M-S, Thomas MS, Cascales E. Anchoring the type VI secretion system to the peptidoglycan: TssL, TagL, TagP... what else? *Virulence*. 27 de novembro de 2010;1(6):535–40.
- Aubert DF, Flannagan RS, Valvano MA. A Novel Sensor Kinase-Response Regulator Hybrid Controls Biofilm Formation and Type VI Secretion System Activity in *Burkholderia cenocepacia*. *Infect Immun*. 1 de maio de 2008;76(5):1979–91.
- Ballister ER, Lai AH, Zuckermann RN, Cheng Y, Mougous JD. In vitro self-assembly of tailorable nanotubes from a simple protein building block. *Proc Natl Acad Sci*. 11 de março de 2008;105(10):3733–8.
- Barbosa VAA, Lery LMS. Insights into *Klebsiella pneumoniae* type VI secretion system transcriptional regulation. *BMC Genomics*. 18 de dezembro de 2019;20(1):506.
- Barker JR, Chong A, Wehrly TD, Yu J, Rodriguez SA, Liu J, et al. The *Francisella tularensis* pathogenicity island encodes a secretion system that is required for phagosome escape and virulence. *Mol Microbiol*. dezembro de 2009;74(6):1459–70.
- Basler M, Mekalanos JJ. Type 6 Secretion Dynamics Within and Between Bacterial Cells. *Science (80-)*. 17 de agosto de 2012;337(6096):815–815.

- Basler M, Pilhofer M, Henderson GP, Jensen GJ, Mekalanos JJ. Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature*. 26 de março de 2012;483(7388):182–6.
- Bent ZW, Poorey K, LaBauve AE, Hamblin R, Williams KP, Meagher RJ. A Rapid Spin Column-Based Method to Enrich Pathogen Transcripts from Eukaryotic Host Cells Prior to Sequencing. Chang Y-F, organizador. *PLoS One*. 21 de dezembro de 2016;11(12):e0168788.
- Bernard CS, Brunet YR, Gavioli M, Lloubes R, Cascales E. Regulation of Type VI Secretion Gene Clusters by 54 and Cognate Enhancer Binding Proteins. *J Bacteriol*. 1 de maio de 2011;193(9):2158–67.
- Bingle LE, Bailey CM, Pallen MJ. Type VI secretion: a beginner's guide. *Curr Opin Microbiol*. fevereiro de 2008;11(1):3–8.
- Bleves S, Sana TG, Voulhoux R. The target cell genus does not matter. *Trends Microbiol*. junho de 2014;22(6):304–6.
- Blondel CJ, Jiménez JC, Contreras I, Santiviago CA. Comparative genomic analysis uncovers 3 novel loci encoding type six secretion systems differentially distributed in *Salmonella* serotypes. *BMC Genomics*. 2009;10(1):354.
- Bobrov AG, Kirillina O, Fetherston JD, Miller MC, Burlison JA, Perry RD. The *Yersinia pestis* siderophore, yersiniabactin, and the ZnuABC system both contribute to zinc acquisition and the development of lethal septicaemic plague in mice. *Mol Microbiol*. agosto de 2014;93(4):759–75.
- Bohlin J, Eldholm V, Brynildsrud O, Petterson JHO, Alfsnes K. Modeling of the GC content of the substituted bases in bacterial core genomes. *BMC Genomics*. 6 de dezembro de 2018;19(1):589.
- Boll EJ, Nielsen LN, Krogfelt KA, Struve C. Novel screening assay for in vivo selection of *Klebsiella pneumoniae* genes promoting gastrointestinal colonisation. *BMC Microbiol*. 2012;12(1):201.
- Bondage DD, Lin J-S, Ma L-S, Kuo C-H, Lai E-M. VgrG C terminus confers the type VI effector transport specificity and is required for binding with PAAR and adaptor–effector complex. *Proc Natl Acad Sci*. 5 de julho de 2016;113(27):E3931–40.
- Bönemann G, Pietrosiuk A, Diemand A, Zentgraf H, Mogk A. Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. *EMBO J*. 18 de fevereiro de 2009;28(4):315–25.
- Borukhov S, Nudler E. RNA polymerase holoenzyme: structure, function and biological implications. *Curr Opin Microbiol*. abril de 2003;6(2):93–100.

- Boulette ML, Payne SM. Anaerobic Regulation of *Shigella flexneri* Virulence: ArcA Regulates fur and Iron Acquisition Genes. J Bacteriol. 1 de outubro de 2007;189(19):6957–67.
- Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? BMC Genomics. 2009;10(1):104.
- Broberg CA, Palacios M, Miller VL. *Klebsiella*: a long way to go towards understanding this enigmatic jet-setter. F1000Prime Rep. 1 de agosto de 2014;6(August).
- Brodmann M, Dreier RF, Broz P, Basler M. *Francisella* requires dynamic type VI secretion system and ClpB to deliver effectors for phagosomal escape. Nat Commun. 16 de agosto de 2017;8(1):15853.
- Brown MJ, Russo BC, O'Dee DM, Schmitt DM, Nau GJ. The contribution of the glycine cleavage system to the pathogenesis of *Francisella tularensis*. Microbes Infect. abril de 2014;16(4):300–9.
- Bruchmann S, Muthukumarasamy U, Pohl S, Preusse M, Bielecka A, Nicolai T, et al. Deep transcriptome profiling of clinical *Klebsiella pneumoniae* isolates reveals strain and sequence type-specific adaptation. Environ Microbiol. novembro de 2015;17(11):4690–710.
- de Bruin OM, Ludu JS, Nano FE. The *Francisella* pathogenicity island protein IgIA localizes to the bacterial cytoplasm and is needed for intracellular growth. BMC Microbiol. 17 de janeiro de 2007;7:1.
- Brunet YR, Bernard CS, Gavioli M, Lloubès R, Cascales E. An Epigenetic Switch Involving Overlapping Fur and DNA Methylation Optimizes Expression of a Type VI Secretion Gene Cluster. Hughes D, organizador. PLoS Genet. 28 de julho de 2011;7(7):e1002205.
- Brunet YR, Henin J, Celia H, Cascales E. Type VI secretion and bacteriophage tail tubes share a common assembly pathway. EMBO Rep. 1 de março de 2014;15(3):315–21.
- Brunet YR, Khodr A, Logger L, Aussel L, Mignot T, Rimsky S, et al. H-NS Silencing of the *Salmonella* Pathogenicity Island 6-Encoded Type VI Secretion System Limits *Salmonella enterica* Serovar Typhimurium Interbacterial Killing. Bäumlér AJ, organizador. Infect Immun. julho de 2015a;83(7):2738–50.
- Brunet YR, Zoued A, Boyer F, Douzi B, Cascales E. The Type VI Secretion TssEFGK-VgrG Phage-Like Baseplate Is Recruited to the TssJLM Membrane Complex via Multiple Contacts and Serves As Assembly Platform for Tail Tube/Sheath

- Polymerization. Viollier PH, organizador. PLOS Genet. 13 de outubro de 2015b;11(10):e1005545.
- Bush M, Dixon R. The Role of Bacterial Enhancer Binding Proteins as Specialized Activators of 54-Dependent Transcription. Microbiol Mol Biol Rev. 1 de setembro de 2012;76(3):497–529.
- Carpenter BM, Whitmire JM, Merrell DS. This Is Not Your Mother's Repressor: the Complex Role of Fur in Pathogenesis. Infect Immun. 1 de julho de 2009;77(7):2590–601.
- Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics. 15 de fevereiro de 2012;28(4):464–9.
- Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics. 15 de agosto de 2005;21(16):3422–3.
- Cascales E. The type VI secretion toolkit. EMBO Rep. 11 de agosto de 2008;9(8):735–41.
- Chakraborty S, Sivaraman J, Leung KY, Mok Y-K. Two-component PhoB-PhoR Regulatory System and Ferric Uptake Regulator Sense Phosphate and Iron to Control Virulence Genes in Type III and VI Secretion Systems of *Edwardsiella tarda*. J Biol Chem. 11 de novembro de 2011;286(45):39417–30.
- Chen F, Zhang W, Schwarz S, Zhu Y, Li R, Hua X, et al. Genetic characterization of an MDR/virulence genomic element carrying two T6SS gene clusters in a clinical *Klebsiella pneumoniae* isolate of swine origin. J Antimicrob Chemother. 1 de junho de 2019;74(6):1539–44.
- Chen L, Zou Y, She P, Wu Y. Composition, function, and regulation of T6SS in *Pseudomonas aeruginosa*. Microbiol Res. março de 2015;172:19–25.
- Cherrak Y, Flaugnatti N, Durand E, Journet L, Cascales E. Structure and Activity of the Type VI Secretion System. Microbiol Spectr. 5 de julho de 2019;7(4):453–72.
- Cherrak Y, Rapisarda C, Pellarin R, Bouvier G, Bardiaux B, Allain F, et al. Biogenesis and structure of a type VI secretion baseplate. Nat Microbiol. 15 de dezembro de 2018;3(12):1404–16.
- Cossart P. Bacterial Invasion: The Paradigms of Enteroinvasive Pathogens. Science (80-). 9 de abril de 2004;304(5668):242–8.
- Dong TG, Ho BT, Yoder-Himes DR, Mekalanos JJ. Identification of T6SS-dependent effector and immunity proteins by Tn-seq in *Vibrio cholerae*. Proc Natl Acad Sci.

- 12 de fevereiro de 2013;110(7):2623–8.
- Dong Y, Geng J, Liu J, Pang M, Awan F, Lu C, et al. Roles of three TonB systems in the iron utilization and virulence of the *Aeromonas hydrophila* Chinese epidemic strain NJ-35. *Appl Microbiol Biotechnol*. 2019;103(10):4203–15.
- Durand E, Nguyen VS, Zoued A, Logger L, Péhau-Arnaudet G, Aschtgen M-S, et al. Biogenesis and structure of a type VI secretion membrane core complex. *Nature*. 22 de julho de 2015;523(7562):555–60.
- Dutta P, Jijumon AS, Mazumder M, Dileep D, Mukhopadhyay AK, Gourinath S, et al. Presence of actin binding motif in VgrG-1 toxin of *Vibrio cholerae* reveals the molecular mechanism of actin cross-linking. *Int J Biol Macromol*. julho de 2019;133:775–85.
- El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. *Nucleic Acids Res*. 8 de janeiro de 2019;47(D1):D427–32.
- English G, Byron O, Cianfanelli FR, Prescott AR, Coulthurst SJ. Biochemical analysis of TssK, a core component of the bacterial Type VI secretion system, reveals distinct oligomeric states of TssK and identifies a TssK–TssFG subcomplex. *Biochem J*. 15 de julho de 2014;461(2):291–304.
- Fang C-T, Chuang Y-P, Shun C-T, Chang S-C, Wang J-T. A Novel Virulence Gene in *Klebsiella pneumoniae* Strains Causing Primary Liver Abscess and Septic Metastatic Complications. *J Exp Med*. 1 de março de 2004;199(5):697–705.
- Fang FC, Rimsky S. New insights into transcriptional regulation by H-NS. *Curr Opin Microbiol*. abril de 2008;11(2):113–20.
- Filloux A. The type VI secretion system: a tubular story. *EMBO J*. 18 de fevereiro de 2009;28(4):309–10.
- Flaugnatti N, Le TTH, Canaan S, Aschtgen M-S, Nguyen VS, Blangy S, et al. A phospholipase A 1 antibacterial Type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery. *Mol Microbiol*. março de 2016;99(6):1099–118.
- Förster A, Planamente S, Manoli E, Lossi NS, Freemont PS, Filloux A. Coevolution of the ATPase ClpV, the Sheath Proteins TssB and TssC, and the Accessory Protein TagJ/HsiE1 Distinguishes Type VI Secretion Classes. *J Biol Chem*. 21 de novembro de 2014;289(47):33032–43.
- Fouts DE, Tyler HL, DeBoy RT, Daugherty S, Ren Q, Badger JH, et al. Complete Genome Sequence of the N₂-Fixing Broad Host Range Endophyte *Klebsiella*

- pneumoniae* 342 and Virulence Predictions Verified in Mice. Guttman DS, organizador. PLoS Genet. 25 de julho de 2008;4(7):e1000141.
- Gan Y, Guan J, Zhou S. A pattern-based nearest neighbor search approach for promoter prediction using DNA structural profiles. Bioinformatics. 15 de agosto de 2009;25(16):2006–12.
- Gan Y, Guan J, Zhou S. A comparison study on feature selection of DNA structural properties for promoter prediction. BMC Bioinformatics. 2012;13(1):4.
- Goldman SR, Nair NU, Wells CD, Nickels BE, Hochschild A. The primary σ factor in *Escherichia coli* can access the transcription elongation complex from solution in vivo. Elife. 15 de setembro de 2015;4:1–17.
- Gordon L, Chervonenkis AY, Gammerman AJ, Shahmuradov IA, Solovyev V V. Sequence alignment kernel for recognition of promoter regions. Bioinformatics. 12 de outubro de 2003;19(15):1964–71.
- Grainger DC, Hurd D, Goldberg MD, Busby SJW. Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. Nucleic Acids Res. setembro de 2006;34(16):4642–52.
- Green J, Rolfe MD, Smith LJ. Transcriptional regulation of bacterial virulence gene expression by molecular oxygen and nitric oxide. Virulence. 17 de novembro de 2014;5(8):794–809.
- Grote A, Klein J, Retter I, Haddad I, Behling S, Bunk B, et al. PRODORIC (release 2009): a database and tool platform for the analysis of gene regulation in prokaryotes. Nucleic Acids Res. 1 de janeiro de 2009;37(Database):D61–5.
- Gruber TM, Gross CA. MULTIPLE SIGMA SUBUNITS AND THE PARTITIONING OF BACTERIAL TRANSCRIPTION SPACE. Annu Rev Microbiol. 2003;57:441–66.
- Gueguen E, Durand E, Zhang XY, D'Amalric Q, Journet L, Cascales E. Expression of a *Yersinia pseudotuberculosis* Type VI Secretion System Is Responsive to Envelope Stresses through the OmpR Transcriptional Activator. Beloin C, organizador. PLoS One. 19 de junho de 2013;8(6):e66615.
- Hachani A, Wood TE, Filloux A. Type VI secretion and anti-host effectors. Curr Opin Microbiol. 2016;29:81–93.
- Hanson PI, Whiteheart SW. AAA+ proteins: have engine, will work. Nat Rev Mol Cell Biol. julho de 2005;6(7):519–29.
- Haugen SP, Ross W, Gourse RL. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. Nat Rev Microbiol. 3 de julho de 2008;6(7):507–19.

- Hennequin C, Forestier C. *oxyR*, a LysR-Type Regulator Involved in *Klebsiella pneumoniae* Mucosal and Abiotic Colonization. *Infect Immun*. 1 de dezembro de 2009;77(12):5449–57.
- Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc Natl Acad Sci*. 7 de julho de 2015;112(27):E3574–81.
- Hood RD, Singh P, Hsu F, Güvener T, Carl MA, Trinidad RRS, et al. A Type VI Secretion System of *Pseudomonas aeruginosa* Targets a Toxin to Bacteria. *Cell Host Microbe*. janeiro de 2010;7(1):25–37.
- Hopper S, Böck A. Effector-mediated stimulation of ATPase activity by the sigma 54-dependent transcriptional activator FHLA from *Escherichia coli*. *J Bacteriol*. maio de 1995;177(10):2798–803.
- Hsieh P-F, Lu Y-R, Lin T-L, Lai L-Y, Wang J-T. *Klebsiella pneumoniae* Type VI Secretion System Contributes to Bacterial Competition, Cell Invasion, Type-1 Fimbriae Expression, and In Vivo Colonization. *J Infect Dis*. 29 de janeiro de 2018;219(4):637–47.
- Hsu F, Schwarz S, Mougous JD. TagR promotes PpkA-catalysed type VI secretion activation in *Pseudomonas aeruginosa*. *Mol Microbiol*. 2009;72(5):1111–25.
- Hurme R, Rhen M. Temperature sensing in bacterial gene regulation - what it all boils down to. *Mol Microbiol*. outubro de 1998;30(1):1–6.
- Ishihama A, Shimada T, Yamazaki Y. Transcription profile of *Escherichia coli*: genomic SELEX search for regulatory targets of transcription factors. *Nucleic Acids Res*. 18 de março de 2016;44(5):2058–74.
- Ishikawa T, Rompikuntal PK, Lindmark B, Milton DL, Wai SN. Quorum Sensing Regulation of the Two *hcp* Alleles in *Vibrio cholerae* O1 Strains. Bereswill S, organizador. *PLoS One*. 24 de agosto de 2009;4(8):e6734.
- Jones C, Allsopp L, Horlick J, Kulasekara H, Filloux A. Subinhibitory Concentration of Kanamycin Induces the *Pseudomonas aeruginosa* type VI Secretion System. Forestier C, organizador. *PLoS One*. 8 de novembro de 2013;8(11):e81132.
- Journet L, Cascales E. The Type VI Secretion System in *Escherichia coli* and Related Species. *EcoSal Plus*. 11 de janeiro de 2016;7(1).
- Kanamaru S. Structural similarity of tailed phages and pathogenic bacterial secretion systems. *Proc Natl Acad Sci*. 2009;
- Kapitein N, Bönemann G, Pietrosiuk A, Seyffer F, Hausser I, Locker JK, et al. ClpV

- recycles VipA/VipB tubules and prevents non-productive tubule formation to ensure efficient type VI protein secretion. *Mol Microbiol.* março de 2013;87(5):1013–28.
- Kitaoka M, Miyata ST, Brooks TM, Unterweger D, Pukatzki S. VasH Is a Transcriptional Regulator of the Type VI Secretion System Functional in Endemic and Pandemic *Vibrio cholerae*. *J Bacteriol.* 1 de dezembro de 2011;193(23):6471–82.
- Kovalenko TF, Patrushev LI. Pseudogenes as Functionally Significant Elements of the Genome. *Biochem.* 15 de novembro de 2018;83(11):1332–49.
- Kube S, Kapitein N, Zimniak T, Herzog F, Mogk A, Wendler P. Structure of the VipA/B Type VI Secretion Complex Suggests a Contraction-State-Specific Recycling Mechanism. *Cell Rep.* julho de 2014;8(1):20–30.
- LaCourse KD, Peterson SB, Kulasekara HD, Radey MC, Kim J, Mougous JD. Conditional toxicity and synergy drive diversity among antibacterial effectors. *Nat Microbiol.* 19 de abril de 2018;3(4):440–6.
- Lawlor MS, Hsu J, Rick PD, Miller VL. Identification of *Klebsiella pneumoniae* virulence determinants using an intranasal infection model. *Mol Microbiol.* novembro de 2005;58(4):1054–73.
- Lazzaro M, Feldman MF, García Véscovi E. A Transcriptional Regulatory Mechanism Finely Tunes the Firing of Type VI Secretion System in Response to Bacterial Enemies. Laub MT, organizador. *MBio.* 6 de setembro de 2017;8(4):1–17.
- Lee DJ, Minchin SD, Busby SJW. Activating Transcription in Bacteria. *Annu Rev Microbiol.* 13 de outubro de 2012;66(1):125–52.
- Leiman PG, Basler M, Ramagopal UA, Bonanno JB, Sauder JM, Pukatzki S, et al. Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc Natl Acad Sci.* 17 de março de 2009;106(11):4154–9.
- Leonhartsberger S, Ehrenreich A, Böck A. Analysis of the domain structure and the DNA binding site of the transcriptional activator FhIA. *Eur J Biochem.* junho de 2000;267(12):3672–84.
- Lertpiriyapong K, Gamazon ER, Feng Y, Park DS, Pang J, Botka G, et al. *Campylobacter jejuni* Type VI Secretion System: Roles in Adaptation to Deoxycholic Acid, Host Cell Adherence, Invasion, and In Vivo Colonization. Bereswill S, organizador. *PLoS One.* 27 de agosto de 2012;7(8):e42842.
- Lery LMS, Frangeul L, Tomas A, Passet V, Almeida AS, Bialek-Davenet S, et al. Comparative analysis of *Klebsiella pneumoniae* genomes identifies a

- phospholipase D family protein as a novel virulence factor. *BMC Biol.* 29 de dezembro de 2014;12(1):41.
- Li J, Tai C, Deng Z, Zhong W, He Y, Ou H-Y. VRprofile: gene-cluster-detection-based profiling of virulence and antibiotic resistance traits encoded within genome sequences of pathogenic bacteria. *Brief Bioinform.* 10 de janeiro de 2017;19(4):bbw141.
- Li J, Yao Y, Xu HH, Hao L, Deng Z, Rajakumar K, et al. SecReT6: a web-based resource for type VI secretion systems found in bacteria. *Environ Microbiol.* julho de 2015;17(7):2196–202.
- Lin C-T, Lin T-H, Wu C-C, Wan L, Huang C-F, Peng H-L. CRP-Cyclic AMP Regulates the Expression of Type 3 Fimbriae via Cyclic di-GMP in *Klebsiella pneumoniae*. He Y-W, organizador. *PLoS One.* 15 de setembro de 2016;11(9):e0162884.
- Lin J, Pissaridou P, Wu H, Tsai M-D, Filloux A, Lai E-M. TagF-mediated repression of bacterial type VI secretion systems involves a direct interaction with the cytoplasmic protein Fha. *J Biol Chem.* 8 de junho de 2018a;293(23):8829–42.
- Lin J, Zhang W, Cheng J, Yang X, Zhu K, Wang Y, et al. A *Pseudomonas* T6SS effector recruits PQS-containing outer membrane vesicles for iron acquisition. *Nat Commun.* 28 de abril de 2017;8(1):14888.
- Lin T-H, Chen Y, Kuo J, Lai Y, Wu C, Huang C, et al. Phosphorylated OmpR Is Required for Type 3 Fimbriae Expression in *Klebsiella pneumoniae* Under Hypertonic Conditions. *Front Microbiol.* 12 de outubro de 2018b;9(October):1–13.
- Lisher JP, Giedroc DP. Manganese acquisition and homeostasis at the host-pathogen interface. *Front Cell Infect Microbiol.* 2013;3(December):1–15.
- Liu L, Ye M, Li X, Li J, Deng Z, Yao Y-F, et al. Identification and Characterization of an Antibacterial Type VI Secretion System in the Carbapenem-Resistant Strain *Klebsiella pneumoniae* HS11286. *Front Cell Infect Microbiol.* 2017;
- Liu P, Li P, Jiang X, Bi D, Xie Y, Tai C, et al. Complete Genome Sequence of *Klebsiella pneumoniae* subsp. *pneumoniae* HS11286, a Multidrug-Resistant Strain Isolated from Human Sputum. *J Bacteriol.* 1 de abril de 2012;194(7):1841–2.
- Lossi NS, Manoli E, Simpson P, Jones C, Hui K, Dajani R, et al. The archetype *Pseudomonas aeruginosa* proteins TssB and TagJ form a novel subcomplex in the bacterial type VI secretion system. *Mol Microbiol.* outubro de 2012;86(2):437–56.
- Ma AT, McAuley S, Pukatzki S, Mekalanos JJ. Translocation of a *Vibrio cholerae* Type VI Secretion Effector Requires Bacterial Endocytosis by Host Cells. *Cell Host*

- Microbe. março de 2009;5(3):234–43.
- Ma J, Pan Z, Huang J, Sun M, Lu C, Yao H. The Hcp proteins fused with diverse extended-toxin domains represent a novel pattern of antibacterial effectors in type VI secretion systems. *Virulence*. 3 de outubro de 2017;8(7):1189–202.
- Ma L-S, Hachani A, Lin J-S, Filloux A, Lai E-M. *Agrobacterium tumefaciens* Deploys a Superfamily of Type VI Secretion DNase Effectors as Weapons for Interbacterial Competition In Planta. *Cell Host Microbe*. julho de 2014;16(1):94–104.
- Mallik P, Pratt TS, Beach MB, Bradley MD, Undamatla J, Osuna R. Growth Phase-Dependent Regulation and Stringent Control of *fis* Are Conserved Processes in Enteric Bacteria and Involve a Single Promoter (*fis P*) in *Escherichia coli*. *J Bacteriol*. 1 de janeiro de 2004;186(1):122–35.
- Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, et al. CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res*. 28 de novembro de 2012;41(D1):D348–52.
- Martin RM, Bachman MA. Colonization, Infection, and the Accessory Genome of *Klebsiella pneumoniae*. *Front Cell Infect Microbiol*. 22 de janeiro de 2018;8(January):1–15.
- Metzger LC, Stutzmann S, Scignari T, Van der Henst C, Matthey N, Blokesch M. Independent Regulation of Type VI Secretion in *Vibrio cholerae* by TfoX and TfoY. *Cell Rep*. maio de 2016;15(5):951–8.
- Miyata ST, Unterweger D, Rudko SP, Pukatzki S. Dual Expression Profile of Type VI Secretion System Immunity Genes Protects Pandemic *Vibrio cholerae*. Mougous JD, organizador. *PLoS Pathog*. 5 de dezembro de 2013;9(12):e1003752.
- Moscoso JA, Mikkelsen H, Heeb S, Williams P, Filloux A. The *Pseudomonas aeruginosa* sensor RetS switches Type III and Type VI secretion via c-di-GMP signalling. *Environ Microbiol*. 2011;13(12):3128–38.
- Mougous JD, Cuff ME, Raunser S, Shen A, Zhou M, Gifford CA, et al. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science*. 9 de junho de 2006;312(5779):1526–30.
- Munch R. PRODORIC: prokaryotic database of gene regulation. *Nucleic Acids Res*. 1 de janeiro de 2003;31(1):266–9.
- Munoz-Price LS, Poirel L, Bonomo RA, Schwaber MJ, Daikos GL, Cormican M, et al. Clinical epidemiology of the global expansion of *Klebsiella pneumoniae* carbapenemases. *Lancet Infect Dis*. setembro de 2013;13(9):785–96.
- Murakami KS, Darst SA. Bacterial RNA polymerases: the whole story. *Curr Opin Struct*

- Biol. fevereiro de 2003;13(1):31–9.
- Nano FE, Schmerk C. The Francisella Pathogenicity Island. Ann N Y Acad Sci. 29 de março de 2007;1105(1):122–37.
- Narayanan A, Vago FS, Li K, Qayyum MZ, Yernool D, Jiang W, et al. Cryo-EM structure of *Escherichia coli* σ 70 RNA polymerase and promoter DNA complex revealed a role of σ non-conserved region during the open complex formation. J Biol Chem. 11 de maio de 2018;293(19):7367–75.
- Nassif X, Sansonetti PJ. Correlation of the virulence of *Klebsiella pneumoniae* K1 and K2 with the presence of a plasmid encoding aerobactin. Infect Immun. dezembro de 1986;54(3):603–8.
- Newton HJ, Sansom FM, Dao J, McAlister AD, Sloan J, Cianciotto NP, et al. Sel1 Repeat Protein LpnE Is a *Legionella pneumophila* Virulence Determinant That Influences Vacuolar Trafficking. Infect Immun. 1 de dezembro de 2007;75(12):5575–85.
- Ng W-L, Bassler BL. Bacterial Quorum-Sensing Network Architectures. Annu Rev Genet. dezembro de 2009;43(1):197–222.
- Paczosa MK, Meccas J. *Klebsiella pneumoniae*: Going on the Offense with a Strong Defense. Microbiol Mol Biol Rev. 15 de setembro de 2016;80(3):629–61.
- Parveen S, Reddy M. Identification of YfiH (PgeF) as a factor contributing to the maintenance of bacterial peptidoglycan composition. Mol Microbiol. setembro de 2017;105(5):705–20.
- Peng D, Li X, Liu P, Zhou X, Luo M, Su K, et al. Transcriptional regulation of galF by RcsAB affects capsular polysaccharide formation in *Klebsiella pneumoniae* NTUH-K2044. Microbiol Res. novembro de 2018;216:70–8.
- Pereira PS, de Araujo CFM, Seki LM, Zahner V, Carvalho-Assef APD, Asensi MD. Update of the molecular epidemiology of KPC-2-producing *Klebsiella pneumoniae* in Brazil: spread of clonal complex 11 (ST11, ST437 and ST340). J Antimicrob Chemother. 1 de fevereiro de 2013;68(2):312–6.
- Pieper R, Huang S-T, Robinson JM, Clark DJ, Alami H, Parmar PP, et al. Temperature and growth phase influence the outer-membrane proteome and the expression of a type VI secretion system in *Yersinia pestis*. Microbiology. 1 de fevereiro de 2009;155(2):498–512.
- Pietrosiuk A, Lenherr ED, Falk S, Bönemann G, Kopp J, Zentgraf H, et al. Molecular Basis for the Unique Role of the AAA + Chaperone ClpV in Type VI Protein Secretion. J Biol Chem. 26 de agosto de 2011;286(34):30010–21.

- Podschun R, Ullmann U. *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. Clin Microbiol Rev. outubro de 1998;11(4):589–603.
- Pukatzki S, Ma AT, Revel AT, Sturtevant D, Mekalanos JJ. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. Proc Natl Acad Sci. 25 de setembro de 2007;104(39):15508–13.
- Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, et al. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. Proc Natl Acad Sci. 31 de janeiro de 2006;103(5):1528–33.
- Qiu N, Misra R. Overcoming iron deficiency of an Escherichia coli tonB mutant by increasing outer membrane permeability. J Bacteriol. 24 de junho de 2019;(June).
- Rajamohan G, Venkataramaiah M, Srinivasan VB, Chauhan NK, Mondal A. Role of oxyR KP, a novel LysR-family transcriptional regulator, in antimicrobial resistance and virulence in *Klebsiella pneumoniae*. Microbiology. 1 de julho de 2013;159(7):1301–14.
- Rapisarda C, Cherrak Y, Kooger R, Schmidt V, Pellarin R, Logger L, et al. In situ and high-resolution cryo-EM structure of a bacterial type VI secretion system membrane complex. EMBO J. 15 de maio de 2019;38(10):e100886.
- Rappas M, Bose D, Zhang X. Bacterial enhancer-binding proteins: unlocking σ 54-dependent gene transcription. Curr Opin Struct Biol. fevereiro de 2007;17(1):110–6.
- Rogers A, Townsley L, Gallego-Hernandez AL, Beyhan S, Kwuan L, Yildiz FH. The LonA Protease Regulates Biofilm Formation, Motility, Virulence, and the Type VI Secretion System in *Vibrio cholerae*. DiRita VJ, organizador. J Bacteriol. 15 de março de 2016;198(6):973–85.
- Rossmann R, Sawers G, Böck A. Mechanism of regulation of the formate-hydrogenlyase pathway by oxygen, nitrate, and pH: definition of the formate regulon. Mol Microbiol. novembro de 1991;5(11):2807–14.
- Russell AB, Hood RD, Bui NK, LeRoux M, Vollmer W, Mougous JD. Type VI secretion delivers bacteriolytic effectors to target cells. Nature. 20 de julho de 2011;475(7356):343–7.
- Russell AB, Peterson SB, Mougous JD. Type VI secretion system effectors: poisons with a purpose. Nat Rev Microbiol. 2 de fevereiro de 2014;12(2):137–48.
- Russell AB, Singh P, Brittnacher M, Bui NK, Hood RD, Carl MA, et al. A Widespread

- Bacterial Type VI Secretion Effector Superfamily Identified Using a Heuristic Approach. *Cell Host Microbe*. maio de 2012;11(5):538–49.
- Salomon D, Gonzalez H, Updegraff BL, Orth K. *Vibrio parahaemolyticus* Type VI Secretion System 1 Is Activated in Marine Conditions to Target Bacteria, and Is Differentially Regulated from System 2. II RMR, organizador. *PLoS One*. 16 de abril de 2013;8(4):e61086.
- Sana TG, Baumann C, Merdes A, Soscia C, Rattei T, Hachani A, et al. Internalization of *Pseudomonas aeruginosa* Strain PAO1 into Epithelial Cells Is Promoted by Interaction of a T6SS Effector with the Microtubule Network. Frank D, Rubin EJ, organizadores. *MBio*. 2 de junho de 2015;6(3):1510–4.
- Sana TG, Hachani A, Bucior I, Soscia C, Garvis S, Termine E, et al. The Second Type VI Secretion System of *Pseudomonas aeruginosa* Strain PAO1 Is Regulated by Quorum Sensing and Fur and Modulates Internalization in Epithelial Cells. *J Biol Chem*. 3 de agosto de 2012;287(32):27095–105.
- Sana TG, Soscia C, Tonglet CM, Garvis S, Bleves S. Divergent Control of Two Type VI Secretion Systems by RpoN in *Pseudomonas aeruginosa*. Beloin C, organizador. *PLoS One*. 21 de outubro de 2013;8(10):e76030.
- Santin YG, Camy CE, Zoued A, Doan T, Aschtgen M-S, Cascales E. Role and Recruitment of the TagL Peptidoglycan-Binding Protein during Type VI Secretion System Biogenesis. Silhavy TJ, organizador. *J Bacteriol*. 25 de março de 2019;201(12).
- Santin YG, Doan T, Lebrun R, Espinosa L, Journet L, Cascales E. In vivo TssA proximity labelling during type VI secretion biogenesis reveals TagA as a protein that stops and holds the sheath. *Nat Microbiol*. 1 de novembro de 2018;3(11):1304–13.
- Sarris PF, Zoumadakis C, Panopoulos NJ, Scoulica E V. Infection , Genetics and Evolution Distribution of the putative type VI secretion system core genes in *Klebsiella* spp . "Infection, Genet Evol. 2011;11(1):157–66.
- Sauer RT, Bolon DN, Burton BM, Burton RE, Flynn JM, Grant RA, et al. Sculpting the Proteome with AAA+ Proteases and Disassembly Machines. *Cell*. outubro de 2004;119(1):9–18.
- Schell MA, Ulrich RL, Ribot WJ, Brueggemann EE, Hines HB, Chen D, et al. Type VI secretion is a major virulence determinant in *Burkholderia mallei*. *Mol Microbiol*. junho de 2007;64(6):1466–85.
- Schlieker C, Zentgraf H, Dersch P, Mogk A. ClpV, a unique Hsp100/Clp member of

- pathogenic proteobacteria. *Biol Chem*. 1 de janeiro de 2005;386(11):1115–27.
- Schumacher J, Joly N, Rappas M, Zhang X, Buck M. Structures and organisation of AAA+ enhancer binding proteins in transcriptional activation. *J Struct Biol*. outubro de 2006;156(1):190–9.
- Schumacher J, Zhang X, Jones S, Bordes P, Buck M. ATP-dependent Transcriptional Activation by Bacterial PspF AAA+Protein. *J Mol Biol*. maio de 2004;338(5):863–75.
- Shalom G, Shaw JG, Thomas MS. In vivo expression technology identifies a type VI secretion system locus in *Burkholderia pseudomallei* that is induced upon invasion of macrophages. *Microbiology*. 1 de agosto de 2007;153(8):2689–99.
- Shang G, Liu X, Lu D, Zhang J, Li N, Zhu C, et al. Structural insight into how *Pseudomonas aeruginosa* peptidoglycanhydrolase Tse1 and its immunity protein Tsi1 function. *Biochem J*. 1 de dezembro de 2012;448(2):201–11.
- Sheng L, Gu D, Wang Q, Liu Q, Zhang Y. Quorum sensing and alternative sigma factor RpoN regulate type VI secretion system I (T6SSVA1) in fish pathogen *Vibrio alginolyticus*. *Arch Microbiol*. 16 de maio de 2012;194(5):379–90.
- Shneider MM, Buth SA, Ho BT, Basler M, Mekalanos JJ, Leiman PG. PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature*. 7 de agosto de 2013;500(7462):350–3.
- Shrivastava S, Mande SS. Identification and Functional Characterization of Gene Components of Type VI Secretion System in Bacterial Genomes. Hannenhalli S, organizador. *PLoS One*. 13 de agosto de 2008;3(8):e2955.
- Shyntum D, Venter S, Moleleki L, Toth I, Coutinho T. Comparative genomics of type VI secretion systems in strains of *Pantoea ananatis* from different environments. *BMC Genomics*. 2014;15(1):163.
- Si M, Wang Y, Zhang B, Zhao C, Kang Y, Bai H, et al. The Type VI Secretion System Engages a Redox-Regulated Dual-Functional Heme Transporter for Zinc Acquisition. *Cell Rep*. julho de 2017a;20(4):949–59.
- Si M, Zhao C, Burkinshaw B, Zhang B, Wei D, Wang Y, et al. Manganese scavenging and oxidative stress response mediated by type VI secretion system in *Burkholderia thailandensis*. *Proc Natl Acad Sci*. 14 de março de 2017b;114(11):E2233–42.
- Silverman JM, Agnello DM, Zheng H, Andrews BT, Li M, Catalano CE, et al. Haemolysin Coregulated Protein Is an Exported Receptor and Chaperone of Type VI Secretion Substrates. *Mol Cell*. setembro de 2013;51(5):584–93.

- Silverman JM, Brunet YR, Cascales E, Mougous JD. Structure and Regulation of the Type VI Secretion System. *Annu Rev Microbiol.* 13 de outubro de 2012;66(1):453–72.
- V. Solovyev AS. V. Solovyev, A Salamov (2011) Automatic Annotation of Microbial Genomes and Metagenomic Sequences. In *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies* (Ed. R.W. Li), Nova Science Publishers, p.61-78. *Nov Sci Publ.* 2011;(February):61–78.
- Storey, A. McNally MÅ. *Klebsiella pneumoniae* type VI secretion system-mediated microbial competition is PhoPQ controlled and reactive oxygen species dependent. julho de 2019;
- Storz G, Imlay JA. Oxidative stress. *Curr Opin Microbiol.* abril de 1999;2(2):188–94.
- Su K, Zhou X, Luo M, Xu X, Liu P, Li X, et al. Genome-wide identification of genes regulated by RcsA, RcsB, and RcsAB phosphorelay regulators in *Klebsiella pneumoniae* NTUH-K2044. *Microb Pathog.* outubro de 2018;123:36–41.
- Suarez G, Sierra JC, Erova TE, Sha J, Horneman AJ, Chopra AK. A Type VI Secretion System Effector Protein, VgrG1, from *Aeromonas hydrophila* That Induces Host Cell Toxicity by ADP Ribosylation of Actin. *J Bacteriol.* 1 de janeiro de 2010a;192(1):155–68.
- Suarez G, Sierra JC, Kirtley ML, Chopra AK. Role of Hcp, a type 6 secretion system effector, of *Aeromonas hydrophila* in modulating activation of host immune cells. *Microbiology.* 2010b;156(12):3678–88.
- Tomás A, Lery L, Regueiro V, Pérez-Gutiérrez C, Martínez V, Moranta D, et al. Functional Genomic Screen Identifies *Klebsiella pneumoniae* Factors Implicated in Blocking Nuclear Factor κ B (NF- κ B) Signaling. *J Biol Chem.* 3 de julho de 2015;290(27):16678–97.
- Trunk K, Coulthurst SJ, Quinn J. A New Front in Microbial Warfare—Delivery of Antifungal Effectors by the Type VI Secretion System. *J Fungi.* 14 de junho de 2019;5(2):50.
- Walker KA, Atkins CL, Osuna R. Functional determinants of the *Escherichia coli* fis promoter: roles of -35, -10, and transcription initiation regions in the response to stringent control and growth phase-dependent regulation. *J Bacteriol.* fevereiro de 1999;181(4):1269–80.
- Wan B, Zhang Q, Ni J, Li S, Wen D, Li J, et al. Type VI secretion system contributes to Enterohemorrhagic *Escherichia coli* virulence by secreting catalase against host reactive oxygen species (ROS). *PLoS Pathog.* 2017;

- Wang T, Si M, Song Y, Zhu W, Gao F, Wang Y, et al. Type VI Secretion System Transports Zn²⁺ to Combat Multiple Stresses and Host Immunity. Skaar EP, organizador. PLOS Pathog. 2 de julho de 2015;11(7):e1005020.
- Waters CM, Bassler BL. QUORUM SENSING: Cell-to-Cell Communication in Bacteria. Annu Rev Cell Dev Biol. novembro de 2005;21(1):319–46.
- Weber BS, Hennon SW, Wright MS, Scott NE, de Berardinis V, Foster LJ, et al. Genetic Dissection of the Type VI Secretion System in *Acinetobacter* and Identification of a Novel Peptidoglycan Hydrolase, TagX, Required for Its Biogenesis. MBio. 2 de novembro de 2016;7(5):1–17.
- Whiteley L, Meffert T, Haug M, Weidenmaier C, Hopf V, Bitschar K, et al. Entry, Intracellular Survival, and Multinucleated-Giant-Cell-Forming Activity of *Burkholderia pseudomallei* in Human Primary Phagocytic and Nonphagocytic Cells. Roy CR, organizador. Infect Immun. outubro de 2017;85(10):1–12.
- Wood TE, Howard SA, Wettstadt S, Filloux A. PAAR proteins act as the ‘sorting hat’ of the type VI secretion system. Microbiology. 5 de agosto de 2019;
- Wu KM, Li NH, Yan JJ, Tsao N, Liao TL, Tsai HC, et al. Genome sequencing and comparative analysis of *Klebsiella pneumoniae* NTUH-K2044, a strain causing liver abscess and meningitis. J Bacteriol. 2009;191(14):4492–501.
- Yang X, Long M, Shen X. Effector–Immunity Pairs Provide the T6SS Nanomachine its Offensive and Defensive Capabilities. Molecules. 26 de abril de 2018;23(5):1009.
- Yang Y, Siu LK, Yeh K, Fung C, Huang S, Hung H, et al. Recurrent *Klebsiella pneumoniae* Liver Abscess: Clinical and Microbiological Characteristics □. 2009;47(10):3336–9.
- Zhang W, Wang Y, Song Y, Wang T, Xu S, Peng Z, et al. A type VI secretion system regulated by OmpR in *Yersinia pseudotuberculosis* functions to maintain intracellular pH homeostasis. Environ Microbiol. fevereiro de 2013;15(2):557–69.
- Zheng M. Activation of the OxyR Transcription Factor by Reversible Disulfide Bond Formation. Science (80-). 13 de março de 1998;279(5357):1718–22.
- Zhou Y, Tao J, Yu H, Ni J, Zeng L, Teng Q, et al. Hcp Family Proteins Secreted via the Type VI Secretion System Coordinately Regulate *Escherichia coli* K1 Interaction with Human Brain Microvascular Endothelial Cells. Bliska JB, organizador. Infect Immun. março de 2012;80(3):1243–51.
- Zou C, Lehti-Shiu MD, Thibaud-Nissen F, Prakash T, Buell CR, Shiu S-H. Evolutionary and Expression Signatures of Pseudogenes in Arabidopsis and Rice. Plant Physiol. setembro de 2009;151(1):3–15.

Zoued A, Brunet YR, Durand E, Aschtgen MS, Logger L, Douzi B, et al. Architecture and assembly of the Type VI secretion system. *Biochim Biophys Acta - Mol Cell Res.* 2014;1843(8):1664–73.

Zoued A, Durand E, Santin YG, Journet L, Roussel A, Cambillau C, et al. TssA: The cap protein of the Type VI secretion system tail. *BioEssays.* outubro de 2017;39(10):1600262.

8 APÊNDICES

APÊNDICE A - Genes putativos do T6SS em Kp52.145

<i>Locus-tag</i> antigo (Gene)	<i>Locus-Tag</i> Atual (Gene)	Coordenadas dos genes, Orientação (+/-), Tamanho (pb)	Proteína Acc (GI)	T6SS componente	SecreT6 e- value	COG/PFAM - VR Profile
BN49_0927	BN49_RS05940	4488461..4488886 [+] 426	597513597	TssE	2.3e-24	COG3518
BN49_0928	BN49_RS05945	4488140..4488457 [+] 318	597513598	TssJ	2e-16	COG3521
BN49_0933	BN49_RS05965	4485521..4486603 [+] 1083	597513603	TssG	3.5e-78	COG3520
BN49_0934	BN49_RS05970	4483797..4485557 [+] 1761	597513604	TssF	2.6e-173	COG3519
BN49_0937	BN49_RS05990	4481491..4481751 [+] 261	597513607	PAAR		pfam05488
BN49_0938	BN49_RS05995	4481026..4481418 [+] 393	597513608	TssA	9.4e-08	COG3515
BN49_0939	BN49_RS06000	4477668..4481033 [+] 3366	597513609	TssM	2.3e-141	COG3523
BN49_0944	BN49_RS06025	4470624..4473179 [+] 2556	597513614	Tssl	2.7e-119	COG3501
BN49_0945	BN49_RS06030	4468705..4470408 [+] 1704	597513615	TagL		COG2885
BN49_0946	BN49_RS06035	4468037..4468702 [+] 666	597513616	TssL	1.7e-45	COG3455
BN49_0947	BN49_RS06040	4466700..4468040 [+] 1341	597513617	TssK	1.3e-124	COG3522
BN49_2462	BN49_RS14020	2873864..2874355 [-] 492	597515126	TssB	4.2e-55	COG3516
BN49_2463	BN49_RS14025	2872283..2873821 [-] 1539	597515127	TssC	9.3e-212	COG3517
BN49_2464	BN49_RS14030	2870930..2872270 [-] 1341	597515128	TssK	2e-139	COG3522
BN49_2465	BN49_RS14035	2870244..2870933 [-] 690	597515129	TssL	2.4e-34	COG3455
	BN49_RS14040	2868541..2870247 [-] 1707		TagL		COG2885
BN49_2469	BN49_RS14050	2865126..2867780 [-] 2655	597515133	TssH	4.4e-276	COG0542
	BN49_RS14070	2579321..2579584 [+] 264		PAAR		pfam05488
BN49_2475	BN49_RS14075	2854480..2857875 [-] 3396	597515139	TssM	3.1e-169	COG3523
BN49_2476	BN49_RS14080	2852891..2854483 [-] 1593	597515140	TssA	2.8e-66	COG3515
BN49_2477	BN49_RS14085	2851058..2852812 [-] 1755	597515141	TssF	1.2e-179	COG3519
BN49_2478	BN49_RS14090	2850009..2851094 [-] 1086	597515142	TssG	2.9e-78	COG3520
BN49_2479	BN49_RS14095	2849489..2850031 [-] 543	597515143	TssJ	4.3e-32	COG3521
BN49_3355	BN49_RS18715	1959816..1961231 [+] 1416	597516014	TssA	4.7e-36	COG3515
BN49_3356	BN49_RS18720	1959353..1959796 [+] 444	597516015	TssE	8.4e-32	COG3518

BN49_3357	BN49_RS18725	1958814..1959350 [+] 537	597516016	TssJ	8.4e-34	COG3521
BN49_3358	BN49_RS18730	1957787..1958833 [+] 1047	597516017	TssG	7.0e-69	COG3520
BN49_3359	BN49_RS18735	1956024..1957787 [+] 1764	597516018	TssF	3.7e-169	COG3519
BN49_3363	BN49_RS18750	3487766..3488029 [-] 264	597516022	paar		pfam05488
BN49_3373	BN49_RS18800	1937205..1939904 [+] 2700	597516032	Tssl	2.6e-113	COG3501
BN49_3374	BN49_RS18805	1935049..1936746 [+] 1698	597516033	TagL		COG2885
BN49_3375	BN49_RS18810	1934392..1935045 [+] 654	597516034	TssL	1.2e-39	COG3455
BN49_3376	BN49_RS18815	1933055..1934395 [+] 1341	597516035	TssK	1.2e-120	COG3522

APÊNDICE B - Genes putativos do T6SS em HS11286

<i>Locus-Tag</i> Atual (Gene)	Coordenadas dos genes, Orientação (+/-), Tamanho (pb)	Proteína Acc (GI)	T6SS componente	SecreT6 e- value	COG/PFAM - VR Profile
KPHS_22970	2326237..2326728 [+] 492	378978456	TssB	5.6e-55	COG3516
KPHS_22980	2326771..2328315 [+] 1545	378978457	TssC	2.1e-214	COG3517
KPHS_22990	2328325..2329668 [+] 1344	378978458	TssK	1.5e-139	COG3522
KPHS_23000	2329665..2330354 [+] 690	378978459	TssL	7.5e-34	COG3455
KPHS_23010	2330351..2332057 [+] 1707	378978460			COG2885
KPHS_23020	2332062..2332553 [+] 492	378978461	TssD	4.2e-48	COG3157
KPHS_23030	2332818..2335472 [+] 2655	378978462	TssH	2.5e-277	COG0542
KPHS_23040	2335474..2337843 [+] 2370	378978463	Tssl	1.3e-123	COG3501
KPHS_23120	2343463..2343720 [+] 258	378978471	PAAR		pfam05488
KPHS_23140	2344840..2348265 [+] 3426	378978473	TssM	3.8e-167	COG3523
KPHS_23150	2348265..2348672 [+] 408	378978474	TssA	2.4e-39	COG3515
KPHS_23170	2349937..2351691 [+] 1755	378978476	TssF	1.3e-179	COG3519
KPHS_23180	2351655..2352740 [+] 1086	378978477	TssG	2.1e-78	COG3520
KPHS_23190	2352718..2353260 [+] 543	378978478	TssJ	4.3e-32	COG3521
KPHS_32770	3262332..3263672 [-] 1341	378979436	TssK	2.0e-120	COG3522
KPHS_32760	3261682..3262335 [-] 654	378979435	TssL	1.3e-40	COG3455
KPHS_32750	3259981..3261678 [-] 1698	378979434	TagL		COG2885

KPHS_32730	3257035..3259521 [-] 2487	378979432	Tssl	3.3e-112	COG3501
KPHS_32520	3255656..3257011 [-] 1356	378979411	PAAR		pfam05488
KPHS_32500	3240404..3243814 [-] 3411	378979409	TssM	4.2e-134	COG3523
KPHS_32490	3238507..3240270 [-] 1764	378979408	TssF	1.7e-169	COG3519
KPHS_32480	3237461..3238477 [-] 1017	378979407	TssG	9.3e-69	COG3520
KPHS_32470	3236944..3237432 [-] 489	378979406	TssJ	1.6e-32	COG3521
KPHS_32460	3236498..3236941 [-] 444	378979405	TssE	7.3e-32	COG3518
KPHS_32450	3235063..3236478 [-] 1416	378979404	TssA	7.8e-37	COG3515

APÊNDICE C – Genes putativos do T6SS em NTUH

Locus-tag antigo (Gene)	Locus-Tag Atual (Gene)	Coordenadas dos genes, Orientação (+/-), Tamanho (pb)	Proteína Acc (GI)	T6SS componente	Secret6 e-value	COG/PFAM - VR Profile
KP1_2392	KP1_RS11220	2298854..2299351 [+] 498	238894408	TssB	2.6e-55	COG3516
KP1_2393	KP1_RS11225	2299394..2300938 [+] 1545	238894409	TssC	1.3e-214	COG3517
KP1_2395	KP1_RS11230	2300951..2302291 [+] 1341	238894410	TssK	1.3e-139	COG3522
KP1_2396	KP1_RS11235	2302288..2302977 [+] 690	238894411	TssL	2.4e-34	COG3455
KP1_2397	KP1_RS11240	2302974..2304680 [+] 1707	238894412	TagL		COG2885
KP1_2398	KP1_RS11245	2304685..2305176 [+] 492	238894413	TssD	4.2e-48	COG3157
KP1_2399	KP1_RS11250	2305440..2308094 [+] 2655	238894414	TssH	8.3e-276	COG0542
KP1_2400	KP1_RS11255	2308087..2310786 [+] 2700	238894415	Tssl	9.2e-124	COG3501
KP1_2413	KP1_RS11295	2318471..2321926 [+] 3456	238894427	TssM	2.5e-171	COG3523
KP1_2414	KP1_RS11300	2321926..2323524 [+] 1599	238894428	TssA	1.1e-64	COG3515
KP1_2415	KP1_RS11305	2323555..2324610 [+] 1056	238894429	PAAR		pfam05488
KP1_2419	KP1_RS11315	2325154..2326908 [+] 1755	238894431	TssF	8.6e-180	COG3519
KP1_2420	KP1_RS11320	2326872..2327957 [+] 1086	238894432	TssG	2.9e-78	COG3520
KP1_2421	KP1_RS11325	2327935..2328477 [+] 543	238894433	TssJ	4.3e-32	COG3521
KP1_3387	KP1_RS15790	3213532..3214872 [-] 1341	238895328	TssK	1.4e-121	COG3522
KP1_3386	KP1_RS15785	3212882..3213535 [-] 654	238895327	TssL	1.2e-39	COG3455
KP1_3385	KP1_RS15780	3211181..3212878 [-] 1698	238895326	TagL		COG2885

KP1_3383	KP1_RS15775	3208236..3210722 [-] 2487	238895324	TssI	2.1e-111	COG3501
KP1_3372	KP1_RS15725	3200332..3200598 [-] 267	238895314	PAAR		pfam05488
KP1_3370	KP1_RS15715	3195777..3199187 [-] 3411	238895312	TssM	3.5e-133	COG3523
KP1_3369	KP1_RS15710	3193880..3195643 [-] 1764	238895311	TssF	1.7e-169	COG3519
KP1_3368	KP1_RS15705	3192834..3193880 [-] 1047	238895310	TssG	5.5e-69	COG3520
KP1_3367	KP1_RS15700	3192317..3192853 [-] 537	238895309	TssJ	9,00E-34	COG3521
KP1_3366	KP1_RS15695	3191871..3192314 [-] 444	238895308	TssE	8.4e-32	COG3518
KP1_3364	KP1_RS15690	3190436..3191851 [-] 1416	238895307	TssA	7.8e-37	COG3515

APÊNDICE D – Sequências promotoras $\sigma 70$

Sequência promotora Kp52.145

>SP1 complementar (972892..973142)

tatcttttcccttctgtacttcactaatctcgcagtagatgggtcgaactttttgccaccaa
tcattcacagcaaaaatatcattttattaaagaagcttttattttattccgaaaaagtggcggt
gcgatattattacgcgatcgtttatcggtattttatgcttatacttagccataaccaagcag
attgcccttgatgggtcaataaattgagaaacttcatataaataaaaactatcttaggatacga
gcA

>SP2 complement (962359..962609)

tcaaacgggtgggaagctgtttatttctgagttgaatgaagttggctttaatgaattcatttt
caaaaatcatacttcaggcgctatcaatgatgaaagatggaagctacatgtgggcggaaact
ggtttttatagatcgcaaggtaaataatgactgggtgaagttgagcatatttacattcttaaat
agctcactgctaataccttccaccgattgatgacttatttttattaaatggataaactttat
ggC

>SP3 complement (955166..955416)

tattttatcttcaataataagaggtatgaggtgccgatccgtagggagtattcggggaataa
atatcattattgggtatttagaggggttcgggtgaaaaagccggatactggcatgaaagatttc
ccgcctcatattcatacagaaagtatctgaataagtaaaaaattaaataaacgcaaccgtaa
aaaacatgatttaattattatattgataattgctcgcgatataatttttcgtaaaataata
tcA

>SP4 complement (951972..952222)

tgacggcactaaagaagtggcatcggcgattttctacaagcagatcaaaacgctgcatctcg
attttgatcgcgcagcgcgctgaatacagatgcgcaagacacaccgctttccacgatgatc
catgtctggcaactgaaaacacgtgagcgttttgacgaggccaactagactggccccctgaa
tctccagacaaccaatatcacttaataaagtgatagtccttaataactagtttttagactagtc
atT

>SP5 2564194..2564444

gtaattgattgcggttggttacgctaaatggcataatgagcgaattatattgccccgggtcgtc
tctatggtttcttttattcatccgcttaatgtttttagcccgccttatagtatccattgcaa
cattcctttggtacattatttattgaatttatcattgacggctgttacacattttaaatatt
gatattcatgggtgcttttattttcgtgggtgaaaggggggtgtttatgaataaatataattatt
agG

>SP6 2579045..2579295

ggataaagaatttatggctgattttacggcgaagctgactggactcagggggaggatccgt
atctcttactggaataacctgacaatgtgcctgagccgatataattagtcgtgaaactggt
actgataaatttaataaggagaaatataaattaccatgtacaggtctgtcaataaggcata
aacaagtattattaagacattcagtttaataaacgtgatgaatatcaatatatatatttaag
gtC

>SP7 2580868..2581118

attgcattttcctggtctggtatgggtgatattttcatttcttctttagtagtatctttttcttt
tatagccattcttacgcgcaatgtaattttcaaaactgctgtatcctcgtttttgtaaggctg
tcagagagcccaaaaaggaaagggttggaatgagaaaaaacatacgattggtaggttttggtc
tcttgattttggttatcttgttcctgattgttttattaaatttctcagataaataaaatccct
gcT

>SP8 complement (3506156..3506406)

tctctctctcgttcgccctgacacccgccccacagcggagcgggaaaccaaatacggggtaa
tgcagcgcctggcaatggccgggtggggcggggcgctcctccgggtatctggcagttgccgg
gggagattctggactcagcgcctactgagggatatacgcattgcgggtgatatttttattattat
ttataatttaaggtaaatataaatattcctgacggagggattaattcgcctaattaatattcaa
caG

>SP9 complement (3501877..3502127)

gttgaacaacaatctgcctgctgaccctgtgcgctacgggtccctgataaacaggaaacag
cgatggggcggtagcgtgcctgccgccagtcaatcttgtcggctaagcgaagcctctgcgggg
atgaagctcgggaaaggggatggataaaggagatcagtatggatagcgcctaaagacagtgcc
agagaaatgctggcggcgataaacggtgaaggctccggtttctggctgctttttttatttct
caG

>SP10 complement (3489079..3489329)

cagcgtggcggctccgtaaaggactgaacgaacgtgccatccgtttcttccgcgagggagac
atcataccggttgataaattaaatcttgtccgcctgcgctcctgctcgccttggcgataaatg
gctgcggggagagaggcaggtgctctggcgactgacgggcgaggcataaatctcctcttacg
gaggttatcgctggcattaatacctataaatttaaggacaaatcatatgcgtaaattttta
atC

>SP11 complement (3483648..3483898)

atttaaccgaaacgcgcgactacggcagcctgggtggcggcggggctgggcccaggagtggatt
tgggcaaacgctgtttgttcccccgatggaacaggcctggcagcaggtgctgaccccggcgg
cggaaagtctcaacgccagtgccgcagcgcgggtggaggactggaatagcgccttttggc
ggtcgcctacccctttaaaaataaccagcagtgaaagtctctctgcccgcctgctggcaaaatacct
cgA

>SP12 complement (1071605..1071855)

gctgctccaccagggacatctggctgcccagggcacgccgcaggagggtgatccagcagccgg
ttatccatcgctggtacggggcggatgtccgcctcgtccagcaccggacaacgcgggtccc
caggtttatctggcgccataatccccgcgcggcgcaactggaatccggatcgcgccgcgctt
acgctttctggcaataattgctgcatatattattgaggggtcatttatatttaaaatctat
acC

>SP13 2861175..2861425

tactatcaaggccatcgttttataagtatattgaaaggccagatgaacataagaaatattgtg
ccggaacaatgaatgatctaactctgtcatcacaggcgttcccttattttaaataccgatgattat
ttctgatgatgaacttgctgaattacttattaaattaatgaggtgaatggatggatttcttg
aaaagatgggttaaagaagaattcattctttttttatggggatgggggtattgtattattaat
caT

>SP14 1219422..1219672

gcgaaaaatattttgctgatatacctaactggcgcggcagcgtctggcaaacgtcggcggtt
gagcagattttcggcggcgatcgctgcacgctaagcgaaaaggatgattttttctcttaccg
gcgcgacaagaccaccggtcgtatggcaagtttcatttggtgatataacctagagaatcaa
gacgatccagaaggcaaacctattttttcacatagttcgggtctttaaccttgaataattgag
ggA

>SP15 1414997..1415247

ttgagcgtgacaccgagcgtgaccgcttctgtctgctgcggaagccgtggagtacgggctg
gtcgcactctattttgacccatcgtaattgatgcccaggcgcaaggggtgccgctttacgctt
catccttctcgccacgtctgtgcgagcgggtggcttttattcggccgcttgtctggcaggct

tgcggccggtggcttcgctaccggctcgatgtagcgcggaaggatTTTgtgtatataactaatg
aaG

>SP16 2084253..2084503

tattgtagatttTgtgatgagcataacctctggTcatcgacaggcctccgcctgTtggcaaca
gcaaccgatgatgaacactgacgatgagtaagagagattggctggacttcgagcacctggtt
gacgacgaagttcgtgatgcgattaaaccgccatctatgtataaagtgatattggtcaatga
tgactacactccgatggagtttTgttattgacgtgttacaaaaattcttttcttatgatgttg
aaC

>SP17 complement (3545783..3546032)

acgatccggttgcggccgggcaactgggccagctggTtaatatcggccagccgTtcatgttc
agtcccggcgaaacgctgctgccggcgggaatgacctgctggTgggccacctcgtgtccgg
cctgctgcagacgctcgacgccgggctgccgcctgctggTtgatgcaataaccggatggc
gcgaccagataaaaactgagacatgataattccttgctatgaatgaaagggttatgatgccgc
ctG

Sequência Promotora HS11286

>SP1 2325856..2326106

ggcaatcggatggaagccagatattttccctctgagtaattgattgcgTtTgttacgctaaat
ttgcataatgagcgaattatattgccccggTcgtctctatggTttcctttattcatccgctt
aatgTtttagcccgccttatggTatccattgcaaacattcctttggTgcattatttattgaa
tttatcattgactgTtTgtttcatattttaaatattgatattcatggTgcttttattttcgtg
gtG

>SP2 2337311..2337561

atcttaaacagccgacattgctgatgtcggccccgcaggggattgccgcagTcacgcggaa
acaaccctcctgcacagcggtaaggggctgtatctgcagagcctgggagaggTcaacatcac
cacggcacagcgtgtTctctgaatgccagccaggctatttcaactgctggcgcacaggagg
ggatgaggctggTctccgccaaggTccgctgcaagtagagTcgcacggcgatattctatcg
ctG

>SP3 2344593..2344843

cagaacgaactgagcggccccccggaagcctgtgcgctgttttttcagccgggactggaagc
gctggcgcacagcggatgggatattaacctgtatcgccaggatgcctgTtggggagacatcg
gcgagatggaaggcttgaccgttctgtcgtggcggcgatctacgccgcgactaccaacaa
ccgtgtggctggctagcccgcatcccttgaataccctcgcaattggaattgtaaaacctga
tgG

>SP4 complement (3263988..3264238)

ctctctctcgTtTgcctgataccgcctacagcggagcgggaaaccaaatacggggTaat
gcagcgcctggcaatggccgggtgggggcggggcgtcctccgggcatctggcagTtgccgg
gggagactctggactcagcgtactgaggggatatcgattgcggTgatatttttattattat
ttatatttaaggTaaatataaatattcctgacggagggattaattcgttaattaattca
caG

>SP5 complement (3259706..3259956)

gaacaacaatctgcctgcctgaccctgtgcgctacgggtccctgataaacaggaaacagcga
tgggcggtagcgtgcctgcagccagTcaatcttTgtcggctaagcgaagcctctgcggggatg
aagctcgggaaaggggatggataaaggagaacagTatggatagcgtaaagacagTgccaga
gaaatactggcggcgataaacgttgaaggctccggTttctggctgctttttttatttttctca
gtC

>SP6 complement (3245229..3245479)
gtctggcaattggcggttcgaagtgaaaaattattaatgtaggggatatcacctgggggaaga
agtatgtattaagtaaagagaggtgtgcagatattccagatatttttatgtttattattc
ttcctttgacaggagtcgccggtttttatctgttattaaatgttattgcctcactctttagg
ttgaaatagatgcgattgagatattctaattaataattactgtatcgctgatataaaaagga
aaT

>SP7 complement (3241047..3241296)
ccgaaacgcgcgactacggcagcctggtggcggcggggctggggccaggagtggagtggattt
gggcaaacgctgtttgttcgcccgatggaacaggcctggcagcaggtgctgaccccgggcggc
ggaaagtctcaacgcccagtgggcgcagcgcggtggtggaggactggaatagcgccttttggcg
gtcgcctacccttttaaaaataaccagcagtggaagtctctctgcccgtgctggcaaaatacctc
aA

>SP8 4196847..4197097
gctgcttcaccagggacatctggtcgcccagggcacgccgcaggaggtgatccagcagccgg
ttatccatcgctggtacggggcggatgtccgcctcgtccagcaccggacaacgcgggtcccc
caggtttatctggcgccataatccccgcgcggcgcaactggaatccggatcgcgcccgcgctt
acgctttctggcaataattgctgcatattatattgaggggtcattttatatttaaatctat
acC

>SP9 complement (4037488..4037738)
gcgaaaaatattttgctgatatctaccaactggcgcggcagcgtctggcaaacgctcggcggtt
gagctgattttcggcggcgcgatcgcctgcacgctaagcgaaaaggatgattttttctcttaccg
gcgcgacaagaccaccggctcgtatggcaagtttcatttggtgatataacctagagaatcaa
gacgatccagaaggcaaacctattttttcacatagttcgggtctttaaccttgaataattgag
ggA

>SP10 1846533..1846783
tattgtagatttgtgatgagcataacctctggtcatcgacaggcctccgcctggtggcaaca
gcaaccgatgatgaacactgacgatgagtaagagagattggctggacttcgagcacctggtt
gacgacgaagttcgtgatgcatgaaaccgccatctatgtataaagtgatattggtcaatga
tgactacactccgatggagtttggtattgacgtgttacaaaaattcttttcttatgatgttg
aaC

>SP11 1197200..1197450
ttgagcgtgacaccgagcgtgaccgcttctgtctgctgcggaagccgtggagtacgggctg
gtcgactctattttgaccatcgtaattgatgcccaaggcgcaagggtgccgctttacgctt
catccttctcgccacgtctgtgagcggttggcttttattcggccgcttgtctggcaggct
tgcgggccgggtggcttcgctaccggctcgatgtagcgcgaaggattttgtgtatataactaatg
aaG

>SP12 complement (3300071..3300320)
acgatccggttgccggccgggcaactggggccagctgggttaatatcggccagccgttcatgttc
agtcccggcgaaacgctgctgccggcggggaatgacctgctggtggggccacctcgtgtccgg
cctgctgcagacgctcgacgccgcggtgcccgcctgctggttgatgcaatacccggatggc
gcgaccagataaaaactgagacatgataattccttgctatgaaatgaaagggttatgatgccgc
ctG

Promoter sequences NTUH-K2044

>SP1 2298514..2298764

gtaatttattgCGttgttacgctaaatttgCataatgagCGaattatattgccccggTcgtc
tctatggtttcctttattcatccgcttaatgttttagcccgcttatagTatccattgCaaa
cattcctttggTgcattatTTattgaatttatcattgacggctgttacacattTTaaatatt
gatattcatggTgcttttattttcgtggTgaaagggggTgtttatgaataaatataattatt
agG

>SP2

gggggTaaaagaatcagttgaaataaacagtagctttaccaatgacgaaggcactgttttat
ttgcttataaattggatgtattggTgggatagaccagttactgtcaactatTTcgcgtata
aaatggattgaaatatgCccTacgagggggaagagaatatcatagttgaaaacggtaggtgtg
gtggagaaaaccctccgattcctgattataattaaaaaagacaagtctctgtTaaattacat
gcT

>SP3 2324858..2325108

gattaggaaagatgcagcgcTcgacattggTtgagggtTggTggTcaatacgaatataacg
ctgcaggtagcggTggTcaaaatTTtattcTTtctTcaaaagagggcgTgtgctacgcacTcg
atTTatataaaagatgattTTTTtaagggcgcgattaaaaatggTtcaagcaataaatgtTt
aaaaaaaaataaaggcattTtaattataagtattaattcaggtgtTataaaggTtattTTTTta
gtT

>SP4 complement (3215046..3215296)

cgctactgagggTatatcgaattgcggTgatattTTtattattattatTTtaaggTaaata
taaatttctcTgacggagggttaattcgctaatTaatattcaacaggaaataataaaagg
gtcgcTtatggaatattcTTtctacatgtgatgaccgagagcaataaagctacctTgggtTg
aattcagatattatttagggccaggcgTtacctccccctgactctTTcgtTTtatgataaac
atC

>SP5 complement (3210909..3211159)

gtTgaacaacaatctgcctgcctgacctTgtgCGctacgggtTcctgataaacaggaaacag
cgatgggCGgtagcgtgcctgcccgccagTcaatctTgtcggcaaaagcggggcctctgcgggg
atgacgcgcaggaaggggatggataaaggagatcattatggatagcGctaaaggactgtggc
agagaaatgctggcggcgataaacgTtgaaggctccggTtctggctgctTTTTtattTct
caG

>SP6 complement (3200893..3201143)

gatgaaaataaaacggctTggTtattTcctaatccgctgcctTTtagTctTgaaccagTaat
gacgcagcgatggatgcgcgaacgctTtgattTcccataggttatggcgaacgcaaaatga
tagggagcaatagTcgtcatattagcGaggtataccactgtTgccacctaatacagaaaatg
tctgTTTTtattTgcttataattctgattattTgtTgttagTgtattTTTTtattTgtatag
agC

>SP7 complement (3195649..3195899)

gaaggaccgctggcGctgctggcGctgcgaaactTtacgctaccggagacgatattcagcgt
gCGcGctctgctgaacgtgtTccgctgaccgacgatatacccggcgaagagggatactgaa
cTTtctccctcactgattggcgtTaaacctgcaacgccaggcctcTTcGctcatggccgccc
gacggccaactaatctgaccgcGcagccatgctgcGtGtctccctgtagacgaataatggaa
taT

>SP8 1821713..1821963

tattgtagattTgtgatgagcataacctctggTcatcGacaggcctccgctgtTggcaaca
gcaaccgatgatgaacactgacgatgagtaagagagattggctggactTcGagcacctggTt

gacgacgaagttcgtgatgcgattaaaccgcatctatgtataaagtgatattgggtcaatga
tgactacactccgatggagtttgttattgacgtgttacaaaaattcttttcttatgatgttg
aaC

>SP9 complement (4008562..4008812)

taccaactggcgcggcagcgtctggcaaacgtcggcgttgagcagattttcggcgggcgatcg
ctgcacgctaagcgaaaaggatgatttttctcttaccggcgcgacaagaccaccgggtcgta
tggcaagtttcatttggtgatataacctagagaatcaagacgatccagaaggcaaacttat
ttttcgcatagttcaggtctttaaccttgaataattgaggggatgacctcatttaactctccag
taG

>SP10 1225506..1225756

ttgagcgtgacaccgagcgtgaccgcttcctgtctgctgcggaagccgtggagtacgggctg
gtcgactctattttgacctatcgtaattgatgcccaaggcgcaaggggtgccgctttacgctt
catccttctcgccacgtctgtgagcgggtggcttttattcggccgcttgtctggcaggct
tgcgggccgggtggcttcgctaccggctcgatgtagcgcgaaggattttgtgtatataactaatg
aaG

>SP11 2268956..2269206

gctacagcgataaatgtgattgctccatcgcgacgctggcgtaaagtttattcagtgatgca
ttttatttttcgttgccagggatcaggtgtatatatttatcttaccattatttatcaaaat
aatcattgggtgggtgatataggtcccggattgaatattcatgatgctttatcttttgaatt
gaattggtgacttgggtgaataaattaaattcttaacctgttaagttgatcggtaactctactt
ttG

>SP12 complement (2568080..2568330)

ctgacaatttgaaaatgcataattcattttattacaatgcgttatacgctaaccatctggag
atgatgtactttctcatcgaacctgtcatcacaggcgttccttattttaaaccgatgattat
ttctgatgatgaacttgctgaattacttattaaattaatgaggtgaatggatggatttcttg
aaaagatgggttaaagaagaattcattcttttttatggggatggggattgtattattaat
caT

>SP13 complement (3251228..3251477)

acgatccgggtgcgccgggcaactgggcccagctgggttaatatcggccagccggtcatgttc
agtcccggcgaaaacgctgctgccggcggggaatgacctgctgggtgggccacctcgtgtccgg
cctgctgcagacgctcgacgccgcgggctgccgcctgctgggtgatgcaatacccggatggc
gcgaccagataaaaactgagacatgataattccttgctatgaatgaaagggttatgatgccgc
ctG

APÊNDICE E - Sequências promotoras σ^{54}

>PS1

gcaaaaccggataaaaagtttcaaaagggaaattttgatgtctgcaaaggaacatatttcttc
atctgcggtatgtcaacaatgatgggatcaggcggcatttatcagtcgttgtttaaaaaa
tacatctgacggtagctatccatactccactttcctgtctggcgaaatctgcgaggatgtg
gttatgcatcagccagacgtccagcacgtaagtcacgctggcacgggtctcaggtttctcaa
at

>PS2

gtcccggcgattggccgcgagctttagtcagaccaagggcggtgactgcgacagtcctgc
actttgtcttgtagcagcggctctggccgcaatagagtcaatttaggtgtcctcaaagtcgg
tagattggtagcggcgtccttcggaagacgaaaaaccgctgcccaagctatacgcgagcgc
acttatagccaccatggccatggctcgttggcgataagtgttaaatcggactacatttggac
ta

>PS3

attcggggaataaataatcattattgggtattagagggttcggtgaaaaagccggatactgg
catgaaagatttcccgcctcatattcatacagaaagtatctgaataagtaaaaaattaata
aacgcaaccgtaaaaaacatgatttaattattatattgataattgctcgcgatataat
tcgtaaaataataatcaacatcgggttttacatctgaagacttgcactatatttacttttaag
aa

>PS4

gataaagcgtgactacataagtatcatgcccaccagacgggttaacggcagcaaagaacc
ttgcagaggcgttcgagcattataacgaatggcatccgcatagtgcgctgggttatcgctcg
ccacgggaatatctgcggcagcgggccagtaatgggttaagtgataacaggtgtctggaat
ataggggcaaatccaccaactatgacacgctgcttatgcaggaagaaaaacgctgaaatca
ga

>PS5

caaatcgggcggaatatcataggtaacggttgtaaggaaacctgtaataaataacttaa
tagtaactgccgacaatgtgtaaaatttataactataagtagccagaaaataaaagcaccac
ttccccacaaataacttatttataataatccaccaattcaactagccattaggtaaaa
acgaaatagtagcagtgcaaagtatagttatatagaagttattaataaggaattcctttcct
tg

>PS6

taagtgaggctatcgaccttgcttgcgcggtggcgcatcacaagggcagataaagcgcggcg
accgcgcgtccgctgcaagcagtagggccggcgtagcaatagagacgcggcccttaataa
cgcccgtagcggactacgccgctgcaaagagacggggcggcaaaaaacgactcgcggactc
tatgcccaagagtcgtttctttttgtgttagacagacggactggaacacgcgaacccaagg
ac

>PS7

aggatccgtatctcttactggaatacctgacaatgtgcctgagccggatataattagtcgt
gaaactgttactgataaatttaataaggagaaatataaattaccatgtacaggtctgtcaa
taaggcataaacaagtattattaagacattcagtttaataaacgtgatgaatatcaatata
atatttaaggctttatttttaaccaagggaatggcaagtgggtaaatatcgttgtaatag
gg

>PS8

ttgaattcagatattatttagggccaggcgttacctccccctgactctttcgtttatgata
aacatcccttgaccaaaaaataccatctaacttcgaaactctcgcgggcaggggagatattt

cctttccctgaagcggcaggttaataatccagccgaacacacggcggatgagggatatccgt
ccgggcatttggctcctgtctgcagaccaatgaaagtgaaaggcacaggaggtgcaatgaaga
tt

>PS9

tctgatatatcgatatgcacataaataatatttatgatgaatgagatgattgtcattaacaa
aattaaacagggcgtcctgtgaattgatttttgatagtcattacagcggagtaaacaggaa
tataatggatacctcttcaataattaccggcaccacccttaaccgttaccagctggatattc
cttcatgtgccgcatcgctggatgtggaagaattcagtggggcagaaaaactgagtgaactt
ta

>PS10

gcatttcctgacttgcttgacggtaggcaaagaaggcgtccctctgtagtatggcaacct
atttaatttagaacagggcggacgcgaggacgagcgaaccgctatttaccgacgccctctct
ccgtccacgagaccgctgactgccgctccgtatttagaggagaatgccgctcaatagcgac
cgtaattatggatatttaaatcctgttttagtatacgcatttaaaaattagcctaaaaata
aA

>PS11

tataagtcgcacgcgcgacgacttgacaaaggcgaactggctgctatatgggccgcttct
ccctatgacttgaagaggggagtgactaacgcgaactggacgcttgcggtccggaggaagcga
gtaccggcgggctgccggttgattagactggcgcgctcggtacgacgcacagaggacatctg
cttattcccttctaagtattacctgctaaattgggacgctataaaaactacgactttacgcga
tg

>PS12

ggccccgtgcccgtcctccactaggtcgtcggccaataggtagcgaccatgccccgcctac
aggcggagcaggtcgtgggcctgttgccgaggggggtccaaatagaccgcggtattaggggc
gcccgcgcttgacctagccttagcgcggcgcgaatgcgaaagaccggttattaacggacgta
atataaactcccagtaaatataaatttttagatatgggtaccgcattattgggttacttta
tA

>PS13

atcgtctaaccgtgcccagagccgacgatagagtccttgtttctgcccgggttgccggtt
ttccagcgaataatgaaacaatacgaagtaataagctaaccgagatattttaaacggggcag
ggggcgaggggaatataaacaagagtagagtcataattactatcaaggccatcgttttata
agtatattgaaaggccagatgaacataagaatatgtgcccgaacaatgaatgatctaactc
gtC

>PS14

cgccgtgtttccagttaccgtcctgcagtttttgctggggcggctcgcgtagatttcattcggcg
gattatcaggcacctcttccttagagccggagcaaccaccaaagccaggctcagtggtggcg
gctgccaccaggatatttcatgcgctcatgacgttttgacttccctcaaatgtttatgcggg
agaatctctgttccctgctcccgattaagaccagctacaatagcacactatattaacggcaa
agC

>PS15

cctgtctgctgcccgaagccgtggagtagcgggctggctcgcactctattttgaccatcgtaatt
gatgcccaaggcgcgaaggggtgccgctttacgcttcatccttctgccacgtctgtgagcgc
ggtggcttttattcggccgcttgtctggcaggcttgccggcgggtggcttcgctaccggctcg
atgtagcgcgaaggattttgtgtatataactaatgaagggcggcacaacgctatatggcggct
tgC

>PS16

ttgttgaaccacttaacagtacccatttccatacttcgacatccttcgtaaactctttattag
taagatggaatgaaccgggtggtggagtgacgagttgttcaattgctcgccaactctcgccac
tacaatttagagaaatacgaaaacgcggtcaagcgattgaggagggggaacggcgaaaaatgc
gtcaaaatttgaagcagttaacgctattgcccggggaatgtgacagccgctcgcgatagaagc
gaT

>PS17

cctgctaggccaacgcccggcccggttgacccggctcgaccaattatagccggctcggcaagtaca
agtcagggccgctttgcgacgacggccgccccttactggacgaccaccgggtggagcacagg
ccggacgacgtctgagctgcccggcccgacggcgacccaactacgttattggcctac
cgcgctggtctattttgactctgtactattaaggaacgatacttactttccaatactacgg
cgG

HS11286

>PS1

caaaatcgggcggaataccataggtaacgtttgtaaggaaaccacgtaataaataacttaa
tagtaactgacaacaagtataaaatttataactataagtaccacgaaaataaaagcaccac
ttcaccacaaaatacttatttataataatccaccaattcaactagccattaggtaaaa
acgaaatagtagcagtgcaaagtgtagttatatagaagttattaataaggaattcctttcct
tgA

>PS2

taagtgaggctatcgaccttgcttgccgggtggcgcatcaccggcagataaagcgcggcg
accgcgctccgctgcaagcgagtagggccggcgtagcaatagagacgcccggcccctaataa
cgcccgctagcggactacggccgctgcaaagagacgggcggaacacgactcgcggactc
tatgcccaagagtcgtttctttttgtgttagacagacggactggaacacgcatccccaagg
acA

>PS3

gctgtatctgcagagcctgggagaggtcaacatcaccacggcacagcgtgttctctgaatg
ccagccaggctatcttactgctggcgcaacaggaggggatgaggctggtctccgcaaaggt
ccgctgcaagtagagtcgcacggcgatattctatcgctgaccgcgctaaaagatatcaccgt
ccagtcaaccaggggcatctgcagttgacggctaaaaatggcatcacctgggttgccggcg
gcG

>PS4

ttgaattcagatattatcttagggccaggcgttacctccccctgactctttcgctttatgata
aacatcccttgactaaaaataccatctaacttcgaaactctcgcgggcaggggagatattt
cctttccctgaagcggcaggtaataaatccagccgaacacacggcggtatgagggatatccgt
ccgggcatttggtcctgtctgcagaccaatgaaagcgaaaggcacaggaggtgcaatgaaga
ttT

>PS5

tctgatatatcgatatgcacataaataatatttatgatgaatgagatgattgtcattaacaa
aattaaacagggcgtcctgtgaattgatttttgatagtcattacagcggagtaaaacaggaa
tataatggatacctcttcaataattaccggcaccacccttaaccggtaccagctggatattc
cttcatgtaccgcatcgctggatgtggaagaattcagtggggcagaaaaactgagtgactt
taT

>PS6

aacatatctcgtacggttatctttaattccgctgaaagcgctcatattccagcagaccggttaa
ccgcaagcttcactttttaataattacatcccctatagtggaccccccttcttcatacataat
tcattttctctccacacgctctataaggtctataataaaaatacaataataagaaggaaactgt
cctcagcggcctcaaaaatagacaataatttacaataacggagtgagaaatccaactttatcta
cgC

>PS7

tataagtcgcacgcgcgcagacgacttgcaacaaggcgactggctgctatatgggcccgttct
ccctatgacttggaagaggggagtgactaacgcaactggacgctgcggtccggaggaagcga
gtaccggcgggctgccggttgattagactggcgcgctcggtagcagcagagggacatctg
cttattcccttataaggtattacctgctaaattgggacgctataaaaactacgactttacgcga
tgG

>PS8

gggtcccgtgcgggcgtcctccactaggtcgtcggccaataggtagcaccatgccccgccta
caggcggagcaggtcgtgggctgtgcgccagggggtccaaatagaccgcggtattagggg
cgcgcccgttgaccttaggcctagcgcggcgcgaatgcaaagaccggtattaacggacgt
aatataaactcccagtaaatataaatttttagatatggttaccgcattattgggttacttt
atA

>PS9

tgattttcggcggcgatcgtgacgctaagcgaagagatgatttttctcttaccggcgc
gacaagaccaccggtcgtatggcaagtttcattttggctgatataacctagagaatcaagacg
atccagaaggcaaacttatttttcacatagttcgggtctttaaccttgaataattgagggat
gacctcatttaatctccagtagcaaatgtgacctgtttatgggaggagttatgcgtctggat
cgT

>PS10

ttgttgaaccacttaacagtacccatttccatacttcgacatccttcgtaaacttttattag
taagatggaatgaaccggtggtggagtgacgagttgttcaattgctcgccaactctcgccac
tacaatttagagaaatacgaacacgctcaagcgattgaggagggggaacggcgaaaaatgc
gtcaaaatttgaagcagttaacgctattgccggggaatgtgacagccgctcgcggatagaagc
gaT

>PS11

cctgtctgctgcggaagccggtggagtacgggctggtcgcactctattttgacccatcgtaatt
gatgcccaaggcgcgaaggggtgccgctttacgcttcctctccttctcgccacgtctgtgcgagcg
ggtggcttttattcggccgcttgtctggcaggcttgccggcgggtggcttcgctaccggctcg
atgtagcgcgaaggattttgtgtatataactaatgaagggcggcacaacgctatatggcggct
tgC

>PS12

gcggcatcataaccctttcattcatagcaaggaattatcatgtctcagttttatctggtcgc
gccatccgggtattgcatcaaccagcaggcggcagcccggcgctcgagcgtctgcagcagg
ccggacacgaggtggcccaccagcaggtcattccccgcggcagcagcgtttcgccgggact
gaacatgaacggctggccgatattaaccagctggcccagttgccggcggcgaaccggatcgt
ccT

NTUH-K2044

>PS1

taatataaataagttatttgggggaaagtgggtgcttttattttcgtgggtacttatagttat
aaattttacacattgtcggcagttactatttaagttattttattacgtggtttccttacaac

gttacctatgatattccgcccgatTTTTgtaattcgccacttatttcctttgggtatctctgc
tggccccgttatatttgcaaagtgtagttatatagaagttattaataaagaattcctttcct
tgA

>PS2

gtaagtgaggctatcgaccttgcttgcgcggtggcgcat AACGGCAGTATAAAGCGCGGCG
accgCGCGTCCGCTGCAAGCGAGTAGGGCCGGCGTAGCAATAGAGACGCCGGCCCTAATAA
CGCCCGCTAGCGGACTACGCCCGTCGAAAGAGACGGGCGGCAAAAAACGACTCGCGGACTC
TATGCCCAAGAGTCGTTTTCTTTTTGTGTTAGACAGACGGACTGGAACACGCGAACCCCAAGG
acA

>PS3

cttataaattggatgtattggtgggatagaccagttactgtcaactatttcgCGTATAAAA
TGGATTGAAATATGCCCTACGAGGGGAAGAGAATATCATAGTTGAAAACGGTAGGTGTGGTG
GAGAAAACCTCCGATTCCTGATTATAATTAaaaaagacaagTCTCTGTTAAATTACATGCT
GAAAAAAATGGAAGGCCACTTCAACGACAAAAGTAACTAGCACGTTGATTTCTGTTTTTT
acC

>PS4

caacctcccaaccaccagttatgcttatattgCGACGTCCATCGCCACCAGTTTTAAATAA
GAAAGAAGTTTTCTCCCGCACACGATGCGTAGCTAAATATTTTTCTACTAAAAAATTCCC
GCGCTAATTTTTACCAAGTTCGTTATTTACAAATTTTTTTTATTTCCGTAATTAATATTCA
TAATTAAGTCCACAATATTTCCAATAAAAAAATCAACGGTCAAAGTTTTAGACTACTAACTT
tcC

>PS5

aaagccctgaaagagtaggacgacgcccataatggagccgaccgctaaggcccggttgtagtg
ggactggcgCGGCTGGGCTGACGACCGACAGGGCGGCGACCCCTGCTCGGCAAGACCGAGC
CATACTTGCCGACGACCCCGACCCACTCCCGCTCTCACTCCCGCTATAAGGAGTCTGGAA
TGGTGGCTCGACCCATGGATATGCCCGGACGTCGGACGGTGGAAAGATGTTCTTTGGCTGC
acA

>PS6

ttgaattcagatattatTTAGGGCCAGGCGTTACCTCCCCCTGACTCTTTCGCTTTATGATA
AACATCCCTTGACCAAAAATATCATCTGACTTCGAAACTCTCGCGGGCAGGGGGAGATATTT
CTTTCCCTGAAGCGGCAGGTAATAAATCCAGCCGAACACACGGCGGATGAGGGATATCCGT
CCGGGCATTTGGTCTGTCTGCAGACCAATGAAAGCGAAAGGCACAGGAGGTGCAATGAAGA
ttT

>PS7

gCGCGTCTTTCCCTACCTATTTCTCTAGTAATACCTATCGCGATTCTTGACACCGTCTC
TTACGACCGCCGCTATTTGCAACTCCGAGGCCAAAGACCGACGAAAAAATAAAGAGTCA
GTCATTATATTTATAGGGACCCCAATGGAAAATCAATGCTTTGAGGACCCCATAGACT
ATATAGCTATACGTGTATTTATATAAATACTACTTACTCTACTAATAGTAATTTGTTTTAAT
ttG

>PS8

ttagtaaagacgatgactatgaaaaatatggtggtgggtccgcagtatgtatttcgTTGTGTC
gttaataaataatgtgtgatgctatatcttatttttaaaaaatcggtggcgTTCAACATATAA
GGTTTTTTCTAATGGATGGTAAAATCATGTACAGTAGTGAAATTATTTTCATCGTCATCAAC
GATGATTGTCAACGGAAAAGGTGGTGCTCATAAAAATCTGAAAGAAGAAATGTTAAGGTTT
gaA

>PS9

ctttacgctaccggagacgatattcagcgtgcgcgcgctctgctgaacgtgttccgctgaccg
acgatatacccggaagaggatactgaaccttctccctcactgattggcggttaacctgca
acgccaggcctccttcgctcatggccgcccgcggccaactaatctgaccgcgagccatgc
tgcgtgtctccctgtagacgaataatggaatattcataatggacgatttaaccctgcgatat
ttT

>PS10

ttgttgaaccacttaacagtaccatttccatacttcgacatccttcgtaaattctttattag
taagatggaatgaaccgggtggtggagtgcgagttgttcaattgctcgccaactctcgccac
tacaatttagagaaatacgaaaacgctcaagcgattgaggagggggaacggcgaaaaatgc
gtcaaaaatttgaagcagttaacgctattgccggggaatgtgacagccgctcgcgatagaagc
gaT

>PS11

cgctgtttccagttaccgtcctgcagtttttgctgggcggtcgcgtagatttcattcggcg
gattatcaggcacctcttcccttagagccggagcaaccaccaaagccaggctcagtggtggcg
gctgccaccagggtatttcatgctgctcatgacgttttgacttccctcaaaatgtttatgctggg
agaatctctgttccctgctcccgattaagaccagctacaatagcacactatattaacggcaa
agC

>PS12

cctgtctgctgcggaagccgtggagtacgggctggtcgactctattttgacccatcgtaatt
gatgcccaaggcgcaaggggtgccgctttacgcttcatccttctcgccacgtctgtgagcg
gttggcttttattcggccgcttgtctggcaggcttgcggccgggtggcttcgctaccggctcg
atgtagcgcgaaggattttgtgtatataactaatgaagggcggcacaacgctatatggcggct
tgC

>PS13

gcgaccgcatttcaaataagtcactacgtaaaataaaaaagcaacgtccctatagtccacat
atataaatagaatgtaataaatagttttattagtaaccaccactatatccagggccataac
ttataagtactacgaaatagaaaacttaacttaacaactgaaccacttatttaatttaagaa
ttggacaattcaactagccattagatgaaaacaaaatagtagtaaggcaaaataagggcgat
taT

>PS14

agcagattggcgcgctctcggctgctatctcagggacaaagacgcgggcccgcgcgcaaaa
aggctcgcttttatactttattatgcttcattatcgattggctctataaaattcggcccgtcc
cccgctcccttatatttgttttctcatctcagttataatgatagttccggtagcaaaatatt
catataactttccggtctacttgtattctttatacacggccttgttacttactagattattt
ccT

>PS15

ttgaccgggtcgaccaattatagccggctggcaagtacaagtcaggggccgctttgcgacgac
ggccgccccttactggacgaccaccgggtggagcacaggccggacgacgtctgagctgagc
gccccgacggcgacgaccaactacgttatggcctaccgcgctggtctattttgactctg
tactattaaggaacgatacttactttccaataactacggcggaccttaccgctaacagtaga
ggT

APÊNDICE F - Sequências promotoras para σ^{70} de genes do T6SS de Kp52.145.

	SP	T6SS adj (BN49_RS)	-35	Região Espaçadora	-10	SIT	Score
Kp52_RI	SP1	06040	ttcata	taaataaaacta	tcttaggat	acgagcA	7.07
	SP2	06000	ttgatg	acttattttattaaat	ggataaact	ttatggC	5.98
	SP3	05970	ttgata	attgctcgcgatatat	tcgtaaaat	aatacA	5.94
	SP4	05945	gtgata	gtcttaataactagt	tttagact	agtcatT	2.81
Kp52_RII	SP5	14020	gtgaaa	gggggtgtttatgaat	aaatataat	tattagG	5.42
	SP6	14070	ttaaat	aaacgtgatgaatatcaat	atataat	taaggcC	2.53
	SP7	14075	ttgttt	tattaatttctcagat	aaataaaat	ccctgcT	7.08
Kp52_RIII	SP8	18815	ttcctg	acggagggattaattcgct	aattaatat	tcaacaG	6.54
	SP9	18800	ttgaag	gctccggtttctggctgc	ttttttat	ttctcaG	7.75
	SP10	18750	ttaag	gacaaatcatat	gcgtaaatt	ttaaatC	5.75
	SP11	18735	gtgaag	tctctctgccgctgc	tggcaaaat	acctcgA	2.64
Kp52_Orphans	SP12	06500	ttatat	tattgagggtcattat	atttaaaat	ctatacC	5.49
	SP13	15545	tttatg	gggatggggta	ttgtattat	taatcaT	7.23
	SP14	07300	tcaca	tagttcgggtcttaacc	ttgaataat	tgagggA	4.95
	SP15	08275	tcgatg	tagcgcgaaggattttgt	gtatatact	aatgaaG	4.41
	SP16	11635	ttgacg	tgttacaaaaattcttt	tcttatgat	gttgaaC	3.86
	SP17	19005	ttcctt	gctatgaatgaaag	ggttatgat	gccgccT	4.32

Legenda - SP: Sequência promotora; adj: adjacente; SIT: Sítio de Início de Transcrição

APÊNDICE G- Sequências promotoras para σ^{70} de genes do T6SS de HS11286

	SP	T6SS adj (KPHS_)	-35	Região Espaçadora	-10	SIT	Score
HS11_RI	SP1	22970	ttaaaa	tattgatattcatgggctc	ttttatatt	cgtgggG	5.44
	SP2	23120	ctgcaa	gtagagtcgcacgg	cgatattct	atcgctG	3.02
	SP3	23140	ttgaat	accctcgcaattggaa	ttgtaaaac	ctgatgG	3.29
HS11_RII	SP4	32770	ttcctg	acggagggattaattcgct	aattaatat	tcaacaG	6.54
	SP5	32730	ctccgg	tttctggctgcttt	ttttatatt	ctcagtC	7.04
	SP6	32520	ttaata	attactgtatcgctg	atatataaa	aggaaaT	6.29
	SP7	32490	gtgaag	tctctctgccgctgc	tggcaaaat	acctcaA	2.75
HS11_Orphans	SP8	41670	ttatat	tattgagggtcatttat	atttaaaat	ctatacC	5.49
	SP9	39850	ttcaca	tagttcgggtctttaacc	ttgaataat	tgagggA	4.95
	SP10	17930	ttgacg	tgttacaaaaattcttt	tcttatgat	gttgaaC	3.86
	SP11	11410	tcgatg	tagcgcgaaggattttgt	gtatatact	aatgaaG	4.41
	SP12	33150	ttcctt	gctatgaatgaaag	ggttatgat	gccgccT	4.32

Legenda - SP: Sequência promotora; adj: adjacente; SIT: Sítio de Início de Transcrição

APÊNDICE H - Sequências promotoras para σ^{70} de genes do T6SS de NTUH-K2044

	SP	T6SS adj (KP1_RS)	-35	Região Espaçadora	-10	SIT	Score
NTUH_RI	SP1	11220	gtgaaa	gggggtgtttatgaat	aaatataat	tattagG	5.42
	SP2	11295	ttaaaa	aagacaagtctc	tgtaaatt	acatgcT	5.02
	SP3	11315	ttaatt	caggtgtataaaa	ggttatTT	ttagtT	9.85
NTUH_RII	SP4	15790	ttacct	ccccctgactcttcg	ctttatgat	aaacatC	3.99
	SP5	15775	ttgaag	gctccggtttctggctgc	ttttttat	ttctcaG	7.75
	SP6	15725	ttgttg	ttagtgatTTT	ttttattgt	atagagC	10.05
	SP7	15710	atgctg	cgtgtctccctgtag	acgaataat	ggaataT	2.32
NTUH_Orphans	SP8	08880	ttgacg	tgttacaaaaattcttt	tcttatgat	gttgaaC	3.86
	SP9	19445	ttgaat	aattgagggatgacctc	atTTaatct	ccagtaG	4.02
	SP10	06015	tcgatg	tagcgcgaaggattttgt	gtatatact	aatgaaG	4.41
	SP11	11085	ttctta	acctgttaagttgat	cggtaatct	acttttG	6.92
	SP12	12525	tttatg	gggatgggta	ttgtattat	taatcaT	7.23
	SP13	15965	ttcctt	gctatgaatgaaag	ggttatgat	gccgccT	4.32

Legenda - SP: Sequência promotora; adj: adjacente; SIT: Sítio de Início de Transcrição

APÊNDICE I- Sequências promotoras para σ^{54} de genes do T6SS de Kp52.145

	SP	T6SS adjacente (BN49_RS)	Sítio de ligação sig54	Fita	Score
Kp52_RI	SP1	06040	GACGTCCAGCACGTTAAGTCACGCTGGCACGGTCTCAGGTTTCTCAAAT	+	15.61
	SP2	06025	ATCAGGTTTACATCAGGCTAAATTGTGAATAGCGGTTGCTCGGTACCGG	-	14.45
	SP3	05970	AACATCGGTTTTACATCTGAAGACTTGCACTATATTTACTTTTAATGAA	+	16.16
	SP4	05945	CCAACTATGACACGCTGCTTATGCAGGAAGAAAAACGCTGAAATCAGA	+	15.02
Kp52_RII	SP5	14020	GTTCTTTTCTTAAGGAATAATTATTGAAGATATATTGATATGAAACGT	-	16.73
	SP6	14050	CAGGAACCCCAAGCGCACAAAGGTCAGGCAGACAGATTGTGTTTTCTTT	-	16.17
	SP7	14070	TTTTTAACCAAAGGAATGGCAAGTGGGTAATATATCGTTGTAATAGGG	+	15.56
Kp52_RIII	SP8	18815	TGTCTGCAGACCAATGAAAGTGAAAGGCACAGGAGGTGCAATGAAGATT	-	15.54
	SP9	18800	CGCTGGATGTGGAAGAATTCAGTGGGGCAGAAAACTGAGTGAACTTTA	-	17.73
	SP10	18750	AATAAAAAATCCGATTA AAAATTTACGCATATGATTTGTCCTTAAATTT	-	15.60
	SP11	18735	GTAGCGCATTTTCAGCATCAAATATCGCAGGGTTAAATCGTCCATTATG	-	16.34
Kp52_Orphans	SP12	06500	TATTTTATTGGGTTATTACGCCATTGGTATAGATTTTAAATATAAATGA	+	15.01
	SP13	15545	CAGATGAACATAAGAAATATGTGCCGGAACAATGAATGATCTAATCTGT	+	13.91
	SP14	07300	GCTCCCGATTAAGACCAGCTACAATAGCACACTATATTAACGGCAAAG	+	13.72
	SP15	08275	TTTTGTGTATATACTAATGAAGGGCGGCACAACGCTATATGGCGGCTTG	+	14.86
	SP16	11635	AGTTAACGCTATTGCCGGGAATGTGACAGCCGTCGCGGATAGAAGCGA	+	13.21
	SP17	19005	GCGGCATCATAACCCTTTCATTCATAGCAAGGAATTATCATGTCTCAGT	+	15.21

Legenda - SP: Sequência promotora;

APÊNDICE J - Sequências promotoras para σ^{54} de genes do T6SS de HS11286

	SP	T6SS adjacente (KPHS_)	Sítio de ligação sig54	Fita	Score
HS11_RI	PS1	KPHS_22970	TGCAAAGTGTAGTTATATAGAAGTTATTAATAAGGAATTCCTTTCCTTG	-	16.39
	PS2	KPHS_23030	TTTCTTTTTGTGTTAGACAGACGGACTGGAACACGCGATCCCCAAGGAC	-	15.99
	PS3	KPHS_23060	GCATCTGCAGTTGACGGCTAAAAATGGCATCACCCCTGGGTTGCGGCGGC	+	16.46
HS11_RII	PS4	KPHS_32770	TGTCTGCAGACCAATGAAAGCGAAAGGCACAGGAGGTGCAATGAAGATT	+	15.28
	PS5	KPHS_32730	CGCTGGATGTGGAAGAATTCAGTGGGGCAGAAAACTGAGTGAACITTA	+	17.52
	PS6	KPHS_32520	ATAGACAATAATTTACAATAACGGAGTGAGAAATCCAACITTTATCTACG	-	13.52
	PS7	KPHS_32490	GTATTACCTGCTAAATTGGGACGCTATAAACTACGACTTTACGCGATG	-	16.30
HS11_Orphans	PS8	KPHS_41670	AGTAAATATAAATTTTAGATATGGTTACCGCATTATTGGGTTACTTTAT	-	15.16
	PS9	KPHS_39850	CCAGTAGCAAATTTGACCTGTTTATGGGAGGAGTTATGCGTCTGGATCG	-	14.19
	PS10	KPHS_17930	AGTTAACGCTATTGCCGGGAATGTGACAGCCGTCGCGGATAGAAGCGA	+	13.34
	PS11	KPHS_11410	TTTTGTGTATATACTAATGAAGGGCGGCACAACGCTATATGGCGGCTTG	+	14.93
	PS12	KPHS_33150	GCGGCATCATAACCCTTTCATTCATAGCAAGGAATTATCATGTCTCAGT	+	15.34

Legenda - SP: Sequência promotora;

APÊNDICE K- Sequências promotoras para σ^{54} de genes do T6SS de NTUH-K2044

	T6SS adjacente		Sítio de ligação sig54	Fita	Score
	SP	(KP1_RS)			
NTUH_RI	SP1	11220	TGCAAAGTGTAGTTATATAGAAGTTATTAATAAAGAATTCCTTTCCTTG	-	16.88
	SP2	11250	TTTCTTTTTGTGTTAGACAGACGGACTGGAACACGCGAACCCCAAGGAC	-	16.18
	SP3	11295	GCCACTTCAACGACAAAAGTAACTAGCACGTTGATTTCTGTTTTTAC	+	15.79
	SP4	11315	TATTTCCAATAAAAAAATCAACGGTCAAAGTTTTAGACTACTAACTTTC	-	16.45
NTUH_RII	SP5	15790	TGTCTGCAGACCAATGAAAGCGAAAGGCACAGGAGGTGCAATGAAGATT	+	15.68
	SP6	15775	TATTTATATAAATACTACTTACTCTACTAATAGTAATTTGTTTTAATTT	-	15.99
	SP7	15725	AAAAGGTGGTGCTCATTAAAAATCTGAAAGAAGAAATGTTAAGGTTTGA	+	15.12
	SP8	15710	AGACGAATAATGGAATATTCATAATGGACGATTTAACCCTGCGATATTT	+	15.03
NTUH_Orphans	SP10	08880	AGTTAACGCTATTGCCGGGAATGTGACAGCCGTCGCGGATAGAAGCGA	+	13.39
	SP11	19445	GCTCCCGATTAAGACCAGCTACAATAGCACACTATATTAACGGCAAAG	-	13.72
	SP12	06015	TTTTGTGTATACTAATGAAGGGCGGCACAACGCTATATGGCGGCTTG	+	14.94
	SP13	11085	AGCCATTAGATGAAAACAAAATAGTAGTAAGGCCAAAATAAGGGCGATTA	-	13.81
	SP14	12525	TCTACTTGTATTCTTTATACACGGCCTTGTTACTTACTAGATTATTTCC	-	14.56
	SP15	15965	ATACTTACTTTCCAATACTACGGCGGACCTTACCGCTAACAGTAGAGG	+	13.51

Legenda - SP: Sequência promotora;

APÊNDICE L – Sítios putativos dos promotores dependentes de σ^{70} em Kp52.145

SP1

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	24	28	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	143	152	-	5.88	CCGATAAACG
IciA Escherichia coli (strain K12)	54	65	-	5.15	TGATTTGGTGGC
RcsAB Escherichia coli (strain K12)	105	112	-	5.35	TCGGAATA

SP2

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	117	128	+	7.76	CGGAAACTGGTT
GcvA Escherichia coli (strain K12)	196	200	+	10.00	CTAAT

SP3

PWM (species)	Start Position	End Position	Strand	Score	Sequence
FhIA Escherichia coli (strain K12)	233	239	+	10.53	TTTTCGT
GcvA Escherichia coli (strain K12)	78	82	-	10.00	CTAAT
H-NS Escherichia coli (strain K12)	224	233	+	5.91	GCGATATATT
Nac Escherichia coli (strain K12)	28	38	-	10.27	ATCGGCACCTC
RcsAB Escherichia coli (strain K12)	42	49	+	5.39	TAGGGAGT

SP4

PWM (species)	Start Position	End Position	Strand	Score	Sequence
H-NS Escherichia coli (strain K12)	65	74	-	5.67	GCGATACAAA

IHF Escherichia coli (strain K12)	206	221	-	6.24	TATCACTTATTTAAGT
IciA Escherichia coli (strain K12)	61	72	+	5.33	CGATTTTGTATC
Lrp (SELEX) Escherichia coli (strain K12)	36	60	-	11.16	AGATGCAGCGTTTTGATCTGCTTGT
RcsAB Escherichia coli (strain K12)	29	36	-	5.20	TAGAAAAT

SP5

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fis Escherichia coli (strain K12)	196	210	-	3.49	GAAAATAAAAGCACC
GcvA Escherichia coli (strain K12)	246	250	-	10.00	CTAAT
H-NS Escherichia coli (strain K12)	141	150	-	6.06	TCAATAAATA
Nac Escherichia coli (strain K12)	149	159	-	10.33	ATGATAAATTC
RcsAB Escherichia coli (strain K12)	212	219	+	5.05	TGGTGAAA

SP6

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	107	111	-	10.00	CTAAT
H-NS Escherichia coli (strain K12)	9	18	-	5.61	GCCATAAATT
Nac Escherichia coli (strain K12)	57	67	+	9.87	ATCCGTATCTC
RcsAB Escherichia coli (strain K12)	154	161	-	5.26	TGGGTAAT

SP7

PWM (species)	Start Position	End Position	Strand	Score	Sequence
H-NS Escherichia coli (strain K12)	231	240	+	5.95	CAGATAAATA

IciA Escherichia coli (strain K12)	232	243	+	5.12	AGATAAATAAAA
RcsAB Escherichia coli (strain K12)	7	14	-	5.09	CAGGAAAA

SP8

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArgR Escherichia coli (strain K12)	234	247	-	9.26	TGAATATTAATTAG
GcvA Escherichia coli (strain K12)	234	238	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	153	162	-	5.78	TCGATATACC
IciA Escherichia coli (strain K12)	223	234	+	5.22	GGATTAATTTCGC
Nac Escherichia coli (strain K12)	170	180	-	9.92	ATAAAAAATATC
RcsAB Escherichia coli (strain K12)	229	236	-	5.02	TAGCGAAT
OxyR Escherichia coli (strain K12)	174	184	-	3.98	AATAATAAAAA

SP9

PWM (species)	Start Position	End Position	Strand	Score	Sequence
H-NS Escherichia coli (strain K12)	202	211	+	5.58	GCGATAAACG
IciA Escherichia coli (strain K12)	155	166	+	5.26	AGATCAGTATGG
RcsAB Escherichia coli (strain K12)	132	139	+	5.13	TCGGGAAA
OxyR Escherichia coli (strain K12)	235	245	-	3.75	AATAAAAAAAG

SP10

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	237	248	+	7.80	CGTAAATTTTTA

H-NS Escherichia coli (strain K12)	115	124	+	6.24	GCGATAAATG
IHF Escherichia coli (strain K12)	62	77	+	6.45	CATCATACCGTTGGAT
IciA Escherichia coli (strain K12)	74	85	+	5.50	GGATAAATTAAA
RcsAB Escherichia coli (strain K12)	206	213	-	5.06	TAGGTATT

SP11

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fis Escherichia coli (strain K12)	199	213	+	3.61	TTTAAAAATACCAGC
GlpR Escherichia coli (strain K12)	51	70	-	6.30	TTTGCCCAATCCACTCCTG
H-NS Escherichia coli (strain K12)	1	10	-	5.53	TCGGTTAAAT
RcsAB Escherichia coli (strain K12)	236	243	+	5.03	TGGCAAAA
OmpR Escherichia coli (strain K12)	201	207	+	9.08	TAAAAAT
OxyR Escherichia coli (strain K12)	199	209	+	4.05	TTTAAAAATAC

SP12

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	234	243	-	6.48	TTTTAAATAT
CytR Escherichia coli (strain K12)	231	242	-	7.66	TTTAAATATAAA
Fis Escherichia coli (strain K12)	226	240	-	3.56	TAAATATAAATGACC
H-NS Escherichia coli (strain K12)	127	136	-	5.60	CAGATAAACC
IHF Escherichia coli (strain K12)	208	223	+	6.81	CTGCATTATATTATTG
IciA Escherichia coli (strain K12)	124	135	-	5.50	AGATAAACCTGG

OxyR Escherichia coli (strain K12)	233	243	+	4.23	TATATTTAAAA
---	-----	-----	---	------	-------------

SP13

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	80	84	+	10.00	CTAAT
IciA Escherichia coli (strain K12)	70	81	-	5.18	AGATCATTCATT
Nac Escherichia coli (strain K12)	144	154	-	9.94	ATAAGTAATTC

SP14

PWM (species)	Start Position	End Position	Strand	Score	Sequence
IHF Escherichia coli (strain K12)	3	18	-	6.31	CAGCAAAATATTTTTC
IciA Escherichia coli (strain K12)	104	115	+	5.13	GGATGATTTTTT
RcsAB Escherichia coli (strain K12)	176	183	+	5.06	TAGAGAAT

SP15

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArgR Escherichia coli (strain K12)	69	82	-	9.29	TGGGTCAAATAGA
Fis Escherichia coli (strain K12)	152	166	-	4.42	GAATAAAAGCCAACC
GcvA Escherichia coli (strain K12)	243	247	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	158	167	-	6.05	CGAATAAAAG
IHF Escherichia coli (strain K12)	144	159	-	6.31	AGCCAACCGCTCGCAC
IciA Escherichia coli (strain K12)	118	129	-	5.13	GGATGAAGCGTA
RcsAB Escherichia coli (strain K12)	83	90	+	5.04	TCGTAATT

OxyR Escherichia coli (strain K12)	229	239	-	5.16	ATACACAAAAT
---	-----	-----	---	------	-------------

SP16

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GlpR Escherichia coli (strain K12)	100	119	-	6.45	GGTGCTCGAAGTCCAGCCAA
H-NS Escherichia coli (strain K12)	143	152	+	5.76	GCGATTAAAC
Nac Escherichia coli (strain K12)	203	213	-	10.19	ATAACAAACTC

SP17

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	35	44	+	6.74	GGTTAATATC
Fis Escherichia coli (strain K12)	53	67	-	3.54	GGACTGAACATGAAC
Fnr Escherichia coli (strain K12)	219	232	+	8.69	TTGCTATGAATGAA
H-NS Escherichia coli (strain K12)	192	201	+	5.62	CAGATAAAAC
IHF Escherichia coli (strain K12)	209	224	-	7.43	TAGCAAGGAATTATCA
IciA Escherichia coli (strain K12)	209	220	+	5.79	TGATAATTCCTT
RcsAB Escherichia coli (strain K12)	213	220	-	5.19	AAGGAATT
OxyR Escherichia coli (strain K12)	191	201	+	3.83	CCAGATAAAAC

Sítios putativos dos promotores dependentes de σ^{70} em HS11286

SP1

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fis Escherichia coli (strain K12)	231	245	-	3.77	GAAAATAAAAGCACC
H-NS Escherichia coli (strain K12)	176	185	-	6.34	TCAATAAATA

IciA Escherichia coli (strain K12)	182	193	-	5.17	TGATAAATTCAA
Nac Escherichia coli (strain K12)	184	194	-	10.45	ATGATAAATTC
RcsAB Escherichia coli (strain K12)	24	31	-	5.47	AGGGAAAA

SP2

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	1	10	-	6.54	TGTTTAAGAT
Fis Escherichia coli (strain K12)	111	125	+	3.33	GAGGTCAACATCACC
H-NS Escherichia coli (strain K12)	240	249	-	5.71	GCGATAGAAT
IHF Escherichia coli (strain K12)	80	95	-	6.39	ATACAGCCCCTTACCG
IciA Escherichia coli (strain K12)	236	247	+	5.53	CGATATTCTATC

SP3

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	82	91	-	6.67	GGTTAATATC
Fis Escherichia coli (strain K12)	36	50	-	3.63	GGCTGAAAAAACAGC
IciA Escherichia coli (strain K12)	207	218	+	5.18	CGATCCCTTGAA
RcsAB Escherichia coli (strain K12)	230	237	+	5.38	TTGGAATT
OxyR Escherichia coli (strain K12)	234	244	+	4.35	AATTGTAAAC

SP4

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fis Escherichia coli (strain K12)	167	181	-	3.27	AATAAAAAATATCACC

GcvA Escherichia coli (strain K12)	234	238	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	222	231	+	5.61	GGGATTAATT
IciA Escherichia coli (strain K12)	223	234	+	5.24	GGATTAATTCGC
Nac Escherichia coli (strain K12)	170	180	-	9.91	ATAAAAAATATC
RcsAB Escherichia coli (strain K12)	229	236	-	5.04	TAGCGAAT
OxyR Escherichia coli (strain K12)	174	184	-	3.99	AATAATAAAAA

SP5

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fis Escherichia coli (strain K12)	231	245	-	3.23	GAAAATAAAAAAGC
GlpR Escherichia coli (strain K12)	61	80	+	6.32	GATGGGCGGTAGCGTGCCTG
IciA Escherichia coli (strain K12)	43	54	+	5.13	TGATAAACAGGA
RcsAB Escherichia coli (strain K12)	129	136	+	5.14	TCGGGAAA

SP6

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	211	215	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	224	233	-	5.94	GCGATACAGT
RcsAB Escherichia coli (strain K12)	39	46	+	5.15	TAGGGGAT

SP7

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fis Escherichia coli (strain K12)	198	212	+	3.61	TTTAAAAATACCAGC

NagC (Selex) Escherichia coli (strain K12)	74	96	+	15.48	GTTTGTTCGCCCGATGGAACAGG
RcsAB Escherichia coli (strain K12)	235	242	+	5.03	TGGCAAAA
OmpR Escherichia coli (strain K12)	200	206	+	9.12	TAAAAAT
OxyR Escherichia coli (strain K12)	198	208	+	4.11	TTTAAAAATAC

SP8

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	234	243	-	6.47	TTTTAAATAT
CytR Escherichia coli (strain K12)	231	242	-	7.64	TTTTAAATATAAA
Fis Escherichia coli (strain K12)	226	240	-	3.57	TAAATATAAATGACC
H-NS Escherichia coli (strain K12)	127	136	-	5.63	CAGATAAACC
IHF Escherichia coli (strain K12)	208	223	+	6.79	CTGCATTATATTATTG
IciA Escherichia coli (strain K12)	124	135	-	5.50	AGATAAACCTGG
OxyR Escherichia coli (strain K12)	233	243	+	4.21	TATATTTAAAA

SP9

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fnr Escherichia coli (strain K12)	16	29	-	7.37	TTGGTAGATATCAG
IHF Escherichia coli (strain K12)	3	18	-	6.30	CAGCAAAAATATTTTTC
IciA Escherichia coli (strain K12)	104	115	+	5.11	GGATGATTTTTT
RcsAB Escherichia coli (strain K12)	176	183	+	5.05	TAGAGAAT

SP10

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GlpR Escherichia coli (strain K12)	100	119	-	6.43	GGTGCTCGAAGTCCAGCCAA
H-NS Escherichia coli (strain K12)	143	152	+	5.77	GCGATTAAAC
Nac Escherichia coli (strain K12)	203	213	-	10.18	ATAACAAACTC

SP11

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArgR Escherichia coli (strain K12)	69	82	-	9.29	TGGGTCAAAATAGA
Fis Escherichia coli (strain K12)	152	166	-	4.43	GAATAAAAGCCAACC
GcvA Escherichia coli (strain K12)	243	247	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	158	167	-	6.05	CGAATAAAAG
IHF Escherichia coli (strain K12)	144	159	-	6.32	AGCCAACCGCTCGCAC
IciA Escherichia coli (strain K12)	118	129	-	5.12	GGATGAAGCGTA
RcsAB Escherichia coli (strain K12)	83	90	+	5.05	TCGTAAATT
OxyR Escherichia coli (strain K12)	229	239	-	5.16	ATACACAAAAT

SP12

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	35	44	+	6.76	GGTTAATATC
Fis Escherichia coli (strain K12)	53	67	-	3.56	GGACTGAACATGAAC
Fnr Escherichia coli (strain K12)	219	232	+	8.72	TTGCTATGAATGAA
H-NS Escherichia coli (strain K12)	192	201	+	5.66	CAGATAAAAC

IHF Escherichia coli (strain K12)	209	224	-	7.46	TAGCAAGGAATTATCA
IciA Escherichia coli (strain K12)	209	220	+	5.81	TGATAATTCCTT
RcsAB Escherichia coli (strain K12)	213	220	-	5.21	AAGGAATT
OxyR Escherichia coli (strain K12)	191	201	+	3.89	CCAGATAAAAC

Sítios putativos dos promotores dependentes de σ^{70} em NTUH-K2044

SP1

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fis Escherichia coli (strain K12)	196	210	-	3.52	GAAAAATAAAGCACC
GcvA Escherichia coli (strain K12)	246	250	-	10.00	CTAAT
H-NS Escherichia coli (strain K12)	3	12	-	6.14	GCAATAAATT
IciA Escherichia coli (strain K12)	147	158	-	5.03	TGATAAATTCAA
Nac Escherichia coli (strain K12)	149	159	-	10.36	ATGATAAATTC
RcsAB Escherichia coli (strain K12)	212	219	+	5.05	TGGTGAAA

SP2

PWM (species)	Start Position	End Position	Strand	Score	Sequence
H-NS Escherichia coli (strain K12)	77	86	-	5.50	CCAATACATC

SP3

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	2	6	-	10.00	CTAAT
GlpR Escherichia coli (strain K12)	31	50	-	7.04	ATTGACCACCAACCCTCCAA

H-NS Escherichia coli (strain K12)	122	131	+	5.58	TCGATTTATA
NagC (Selex) Escherichia coli (strain K12)	85	107	+	16.72	TTTTATTCTTTCTTCAAAAGAGG
RcsAB Escherichia coli (strain K12)	4	11	+	5.14	TAGGAAAAG

SP4

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	92	96	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	11	20	-	5.74	TCGATATACC
IciA Escherichia coli (strain K12)	81	92	+	5.03	GGATTAATTCGC
Nac Escherichia coli (strain K12)	28	38	-	9.87	ATAAAAAATATC

SP5

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GlpR Escherichia coli (strain K12)	64	83	+	6.26	GATGGGCGGTAGCGTGCCTG
H-NS Escherichia coli (strain K12)	202	211	+	5.66	GCGATAAACG
IciA Escherichia coli (strain K12)	155	166	+	5.27	AGATCATTATGG
RcsAB Escherichia coli (strain K12)	243	250	-	5.10	TGAGAAAT
OxyR Escherichia coli (strain K12)	235	245	-	3.88	AATAAAAAAAG

SP6

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	173	177	+	10.00	CTAAT
GntR Escherichia coli (strain K12)	102	117	-	10.25	GCGTTCGCCATAACCT

H-NS Escherichia coli (strain K12)	240	249	-	5.66	TCTATACAAT
Nac Escherichia coli (strain K12)	176	186	+	10.10	ATCAGAAAATG
RcsAB Escherichia coli (strain K12)	24	31	-	5.81	TAGGAAAT

SP7

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fis Escherichia coli (strain K12)	46	60	+	3.47	GAGACGATATTCAGC
GcvA Escherichia coli (strain K12)	196	200	+	10.00	CTAAT
IciA Escherichia coli (strain K12)	191	202	-	5.89	AGATTAGTTGGC

SP8

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GlpR Escherichia coli (strain K12)	100	119	-	6.43	GGTGCTCGAAGTCCAGCCAA
H-NS Escherichia coli (strain K12)	143	152	+	5.77	GCGATTAAAC
Nac Escherichia coli (strain K12)	203	213	-	10.18	ATAACAAACTC

SP9

PWM (species)	Start Position	End Position	Strand	Score	Sequence
IciA Escherichia coli (strain K12)	233	244	-	5.18	AGATTAAATGAG
RcsAB Escherichia coli (strain K12)	153	160	+	5.07	TAGAGAAT

SP10

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArgR Escherichia coli (strain K12)	69	82	-	9.29	TGGGTCAAAATAGA

Fis Escherichia coli (strain K12)	152	166	-	4.43	GAATAAAAGCCAACC
GcvA Escherichia coli (strain K12)	243	247	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	158	167	-	6.05	CGAATAAAAAG
IHF Escherichia coli (strain K12)	144	159	-	6.32	AGCCAACCGCTCGCAC
IciA Escherichia coli (strain K12)	118	129	-	5.12	GGATGAAGCGTA
RcsAB Escherichia coli (strain K12)	83	90	+	5.05	TCGTAATT
OxyR Escherichia coli (strain K12)	229	239	-	5.16	ATACACAAAAT

SP11

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GlpR Escherichia coli (strain K12)	132	151	-	6.59	CGGGACCTATATCACCCACC
H-NS Escherichia coli (strain K12)	7	16	+	6.40	GCGATAAATG
IciA Escherichia coli (strain K12)	167	178	-	5.08	AGATAAAGCATC
RcsAB Escherichia coli (strain K12)	235	242	+	5.20	TCGGTAAT

SP12

PWM (species)	Start Position	End Position	Strand	Score	Sequence
H-NS Escherichia coli (strain K12)	217	226	-	5.60	CCCATAAAAA
Nac Escherichia coli (strain K12)	144	154	-	10.03	ATAAGTAATTC

SP13

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	35	44	+	6.76	GGTTAATATC

Fis Escherichia coli (strain K12)	53	67	-	3.56	GGACTGAACATGAAC
Fnr Escherichia coli (strain K12)	219	232	+	8.72	TTGCTATGAATGAA
H-NS Escherichia coli (strain K12)	192	201	+	5.66	CAGATAAAAC
IHF Escherichia coli (strain K12)	209	224	-	7.46	TAGCAAGGAATTATCA
IciA Escherichia coli (strain K12)	209	220	+	5.81	TGATAATTCCTT
RcsAB Escherichia coli (strain K12)	213	220	-	5.21	AAGGAATT
OxyR Escherichia coli (strain K12)	191	201	+	3.89	CCAGATAAAAC

Sítios putativos dos promotores dependentes de σ^{54} em Kp52.145

SP1

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	26	37	+	7.75	GGGAAATTTTGA
GlpR Escherichia coli (strain K12)	36	55	-	6.41	TATGTTCTTTGCAGACATC
IciA Escherichia coli (strain K12)	88	99	+	5.16	GGATCAGGCGGC
Nac Escherichia coli (strain K12)	28	38	-	10.35	ATCAAAATTTTC

SP2

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	225	234	+	6.65	TGTTAAATCG
CytR Escherichia coli (strain K12)	97	108	-	7.61	CCTAAATTGACT
FhlA Escherichia coli (strain K12)	153	159	-	10.50	TTTTTCGT
Fnr Escherichia coli (strain K12)	24	37	-	7.48	TTGGTCTGACTAAA

IHF Escherichia coli (strain K12)	88	103	+	6.25	CCGCAATAGAGTCAAT
IciA Escherichia coli (strain K12)	218	229	+	5.22	CGATAAGTGTTA

SP3

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CysB Escherichia coli (strain K12)	176	206	-	9.13	ATGTTGATATTATTTTACGAAAAATATATCG
FhIA Escherichia coli (strain K12)	184	190	+	10.62	TTTTCGT
GcvA Escherichia coli (strain K12)	29	33	-	10.00	CTAAT
H-NS Escherichia coli (strain K12)	175	184	+	5.95	GCGATATATT
OmpR (F box) Escherichia coli (strain K12)	236	245	+	13.94	TTTACTTTTA
RcsAB Escherichia coli (strain K12)	3	10	+	5.51	TCGGGGAA

SP4

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	164	173	-	6.36	TGTTATCACT
H-NS Escherichia coli (strain K12)	81	90	-	5.60	TCGTTATAAT
IHF Escherichia coli (strain K12)	230	245	-	6.54	TTTCAGCGTTTTTCT
IciA Escherichia coli (strain K12)	127	138	-	5.34	AGATATTCCCGT

SP5

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	119	130	-	7.61	GGGAAAGTGGTG

FhA Escherichia coli (strain K12)	106	112	-	10.97	TTTTCGT
GcvA Escherichia coli (strain K12)	176	180	-	10.00	CTAAT
IciA Escherichia coli (strain K12)	148	159	-	5.11	GGATTATTAATA
RcsAB Escherichia coli (strain K12)	126	133	-	5.71	TGGGGGAA

SP6

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	212	221	+	6.50	TGTTAGACAG
GcvA Escherichia coli (strain K12)	118	122	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	199	208	+	5.62	TCGTTTCTTT
IciA Escherichia coli (strain K12)	4	15	-	5.26	CGATAGCCTCAC
Lrp (SELEX) Escherichia coli (strain K12)	41	65	-	11.53	GGTCGCCGCGCTTTATACTGCCCGT
OmpR (F box) Escherichia coli (strain K12)	201	210	+	13.47	GTTTCTTTTT
RcsAB Escherichia coli (strain K12)	97	104	+	5.26	TAGCAATA

SP7

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	54	58	-	10.00	CTAAT
H-NS Escherichia coli (strain K12)	231	240	-	5.79	ACGATATATT
Nac Escherichia coli (strain K12)	4	14	+	10.00	ATCCGTATCTC
RcsAB Escherichia coli (strain K12)	101	108	-	5.31	TGGGTAAT

SP8

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	57	66	-	6.34	TGTTTATCAT
CytR Escherichia coli (strain K12)	116	127	-	7.32	AGGAAATATCTC
H-NS Escherichia coli (strain K12)	146	155	-	5.62	TGGATTTATT
IciA Escherichia coli (strain K12)	117	128	+	5.31	AGATATTCCTT
RcsAB Escherichia coli (strain K12)	121	128	-	5.03	AAGGAAAT

SP9

PWM (species)	Start Position	End Position	Strand	Score	Sequence
H-NS Escherichia coli (strain K12)	4	13	-	6.00	TCGATATATC
IciA Escherichia coli (strain K12)	1	12	-	5.03	CGATATATCAGA

SP10

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	206	215	-	6.37	AGTTAACTAT
CytR Escherichia coli (strain K12)	173	184	-	7.81	AAGAAATTAGGA
GcvA Escherichia coli (strain K12)	175	179	+	10.00	CTAAT
IciA Escherichia coli (strain K12)	204	215	+	5.05	AGATAGTTAACT
RcsAB Escherichia coli (strain K12)	70	77	-	5.26	TAGGAAGT

SP11

PWM (species)	Start Position	End Position	Strand	Score	Sequence
---------------	----------------	--------------	--------	-------	----------

CytR Escherichia coli (strain K12)	211	222	+	7.85	GCTAAATGGGA
GcvA Escherichia coli (strain K12)	146	150	-	10.00	CTAAT
H-NS Escherichia coli (strain K12)	224	233	+	5.65	GCTATAAAAC
IciA Escherichia coli (strain K12)	144	155	+	5.04	TGATTAGACTGG

SP12

PWM (species)	Start Position	End Position	Strand	Score	Sequence
AraC Escherichia coli (strain K12)	208	236	-	13.59	AATCTATACCAATGGCGTAATAACCCAAT
GcvA Escherichia coli (strain K12)	77	81	-	10.00	CTAAT
H-NS Escherichia coli (strain K12)	91	100	-	6.00	TCGATACACT
Nac Escherichia coli (strain K12)	119	129	+	8.96	ATAATCAACGT
RcsAB Escherichia coli (strain K12)	195	202	+	5.10	TCGTAAAT

SP13

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	242	246	+	10.00	CTAAT
IciA Escherichia coli (strain K12)	232	243	-	5.38	AGATCATTCATT
RcsAB Escherichia coli (strain K12)	160	167	-	5.22	TAGTAATA

SP14

PWM (species)	Start Position	End Position	Strand	Score	Sequence
H-NS Escherichia coli (strain K12)	206	215	+	5.88	CCGATTAAGA

IHF Escherichia coli (strain K12)	167	182	+	6.51	CCTCAAAATGTTTATG
IciA Escherichia coli (strain K12)	59	70	-	5.20	TGATAATCCGCC
OxyR Escherichia coli (strain K12)	165	175	+	3.85	TTCCTCAAAAT

SP15

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArgR Escherichia coli (strain K12)	41	54	-	9.29	TGGGTCAAAATAGA
Fis Escherichia coli (strain K12)	124	138	-	4.42	GAATAAAAGCCAACC
GcvA Escherichia coli (strain K12)	215	219	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	130	139	-	6.04	CGAATAAAAG
IHF Escherichia coli (strain K12)	116	131	-	6.32	AGCCAACCGCTCGCAC
IciA Escherichia coli (strain K12)	90	101	-	5.12	GGATGAAGCGTA
RcsAB Escherichia coli (strain K12)	55	62	+	5.07	TCGTAAAT
OxyR Escherichia coli (strain K12)	201	211	-	5.15	ATACACAAAAT

SP16

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	188	199	-	7.72	TTCAAATTTTGA
FhIA Escherichia coli (strain K12)	141	147	-	10.65	TTTTTCGT
GcvA Escherichia coli (strain K12)	58	62	-	10.00	CTAAT
GlpR Escherichia coli (strain K12)	103	122	+	6.69	ATTGCTCGCCAACCTCTCGCC
IHF Escherichia coli (strain K12)	150	165	-	6.48	CCTCAATCGCTTGACG

RcsAB Escherichia coli (strain K12)	46	53	+	5.53	TCGTAAAT
--	----	----	---	------	----------

SP17

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	205	214	-	6.84	GGTTAATATC
Fis Escherichia coli (strain K12)	182	196	+	3.48	GGACTGAACATGAAC
Fnr Escherichia coli (strain K12)	17	30	-	8.85	TTGCTATGAATGAA
H-NS Escherichia coli (strain K12)	48	57	-	5.59	CAGATAAAAC
IHF Escherichia coli (strain K12)	25	40	+	7.31	TAGCAAGGAATTATCA
IciA Escherichia coli (strain K12)	29	40	-	5.95	TGATAATTCCTT
RcsAB Escherichia coli (strain K12)	29	36	+	5.16	AAGGAATT
OxyR Escherichia coli (strain K12)	233	243	+	4.13	TTTCCCAATAC

Sítios putativos dos promotores dependentes de σ^{54} em HS11286

SP1

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	36	47	+	7.70	AGGAAACCACGT
FhIA Escherichia coli (strain K12)	106	112	-	11.06	TTTTCGT
GcvA Escherichia coli (strain K12)	176	180	-	10.00	CTAAT
RcsAB Escherichia coli (strain K12)	156	163	-	5.65	TGGTGGAT

SP2

PWM (species)	Start Position	End Position	Strand	Score	Sequence
------------------	-------------------	-----------------	--------	-------	----------

ArcA Escherichia coli (strain K12)	212	221	+	6.50	TGTTAGACAG
GcvA Escherichia coli (strain K12)	118	122	+	10.00	CTAAT
IciA Escherichia coli (strain K12)	231	242	-	5.35	GGATCGCGTGTT
Lrp (SELEX) Escherichia coli (strain K12)	41	65	-	11.48	GGTCGCCGCGCTTTATACTGCCCGT
OmpR (F box) Escherichia coli (strain K12)	201	210	+	13.42	GTTTCTTTTT
RcsAB Escherichia coli (strain K12)	97	104	+	5.25	TAGCAATA

SP3

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	220	231	+	7.63	TAAAAATGGCAT
IciA Escherichia coli (strain K12)	171	182	-	5.70	TGATATCTTTTA
OmpR Escherichia coli (strain K12)	220	226	+	8.93	TAAAAAT

SP4

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	57	66	-	6.34	TGTTTATCAT
H-NS Escherichia coli (strain K12)	146	155	-	5.62	TGGATTTATT
IciA Escherichia coli (strain K12)	117	128	+	5.31	AGATATTTTCCTT
RcsAB Escherichia coli (strain K12)	121	128	-	5.03	AAGGAAAT

SP5

PWM (species)	Start Position	End Position	Strand	Score	Sequence
H-NS Escherichia coli (strain K12)	4	13	-	6.00	TCGATATATC

IciA Escherichia coli (strain K12)	1	12	-	5.05	CGATATATCAGA
---	---	----	---	------	--------------

SP6

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	207	218	-	7.95	TGTAAATTATTG
H-NS Escherichia coli (strain K12)	239	248	-	5.75	TAGATAAAGT
IciA Escherichia coli (strain K12)	236	247	-	5.70	AGATAAAGTTGG
RcsAB Escherichia coli (strain K12)	90	97	-	6.09	TAGGGGAT

SP7

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	211	222	+	7.84	GCTAAATTGGGA
GcvA Escherichia coli (strain K12)	146	150	-	10.00	CTAAT
H-NS Escherichia coli (strain K12)	224	233	+	5.63	GCTATAAAAC
IciA Escherichia coli (strain K12)	144	155	+	5.03	TGATTAGACTGG

SP8

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	172	181	-	6.91	CGTTAATAAC
CysB Escherichia coli (strain K12)	167	197	+	9.70	AGACCGTTATTAACGGACGTAATATAATAAC
CytR Escherichia coli (strain K12)	208	219	-	7.70	CTAAAATTTATA
GcvA Escherichia coli (strain K12)	117	121	-	10.00	CTAAT
IHF Escherichia coli (strain K12)	207	222	-	6.41	TATCTAAAATTTATAT

IciA Escherichia coli (strain K12)	218	229	+	5.28	AGATATGGTTAC
Nac Escherichia coli (strain K12)	211	221	-	9.92	ATCTAAAATTT

SP9

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	207	218	+	7.61	AGCAAATTTGAC
IciA Escherichia coli (strain K12)	190	201	-	5.04	AGATTAAATGAG
RcsAB Escherichia coli (strain K12)	206	213	+	5.15	TAGCAAAT

SP10

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	188	199	+	7.74	TCAAATTTGAA
FhlA Escherichia coli (strain K12)	141	147	-	10.66	TTTTCGT
GcvA Escherichia coli (strain K12)	58	62	-	10.00	CTAAT
GlpR Escherichia coli (strain K12)	103	122	+	6.69	ATTGCTCGCCAACTCTCGCC
IHF Escherichia coli (strain K12)	150	165	-	6.50	CCTCAATCGCTTGACG
RcsAB Escherichia coli (strain K12)	46	53	+	5.54	TCGTAAT

SP11

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArgR Escherichia coli (strain K12)	41	54	-	9.27	TGGGTCAAAATAGA
Fis Escherichia coli (strain K12)	124	138	-	4.41	GAATAAAAGCCAACC
GcvA Escherichia coli (strain K12)	215	219	+	10.00	CTAAT

H-NS Escherichia coli (strain K12)	130	139	-	6.03	CGAATAAAAG
IHF Escherichia coli (strain K12)	116	131	-	6.32	AGCCAACCGCTCGCAC
IciA Escherichia coli (strain K12)	90	101	-	5.11	GGATGAAGCGTA
RcsAB Escherichia coli (strain K12)	55	62	+	5.07	TCGTAATT
OxyR Escherichia coli (strain K12)	201	211	-	5.15	ATACACAAAAT

SP12

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	205	214	-	6.84	GGTTAATATC
Fis Escherichia coli (strain K12)	182	196	+	3.48	GGACTGAACATGAAC
Fnr Escherichia coli (strain K12)	17	30	-	8.85	TTGCTATGAATGAA
H-NS Escherichia coli (strain K12)	48	57	-	5.59	CAGATAAAAC
IHF Escherichia coli (strain K12)	25	40	+	7.31	TAGCAAGGAATTATCA
IciA Escherichia coli (strain K12)	29	40	-	5.95	TGATAATTCCCTT
OxyR Escherichia coli (strain K12)	48	58	-	3.74	CCAGATAAAAC
RcsAB Escherichia coli (strain K12)	29	36	+	5.16	AAGGAATT

Sítios putativos dos promotores dependentes de σ^{54} em NTUH-K2044

SP1

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	162	173	-	7.60	AGGAAATAAGTA
FhIA Escherichia coli (strain K12)	41	47	+	10.17	TTTTTCGT

Fis Escherichia coli (strain K12)	31	45	-	3.48	GAAAATAAAAGCACC
H-NS Escherichia coli (strain K12)	213	222	-	5.58	TCTATATAAC
Nac Escherichia coli (strain K12)	146	156	-	10.29	ATTACAAAATC
RcsAB Escherichia coli (strain K12)	167	174	-	5.38	AAGGAAAT

SP2

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	212	221	+	6.50	TGTTAGACAG
GcvA Escherichia coli (strain K12)	118	122	+	10.00	CTAAT
IciA Escherichia coli (strain K12)	5	16	-	5.26	CGATAGCCTCAC
Lrp (SELEX) Escherichia coli (strain K12)	41	65	-	11.43	GGTCGCGCGCTTATACTGCCGTT
OmpR (F box) Escherichia coli (strain K12)	201	210	+	13.47	GTTTCTTTTT
RcsAB Escherichia coli (strain K12)	97	104	+	5.26	TAGCAATA

SP3

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	219	230	+	7.80	AGTAAACTAGCA
OmpR (F box) Escherichia coli (strain K12)	215	224	-	13.45	TTTACTTTTG
RcsAB Escherichia coli (strain K12)	74	81	-	5.06	TAGGGCAT

SP4

PWM (species)	Start Position	End Position	Strand	Score	Sequence
---------------	----------------	--------------	--------	-------	----------

CytR Escherichia coli (strain K12)	144	155	-	8.32	TGTAATAACGA
Fis Escherichia coli (strain K12)	125	139	-	3.51	GGTAAAAATTAGCGC
GcvA Escherichia coli (strain K12)	128	132	+	10.00	CTAAT
GlpR Escherichia coli (strain K12)	67	86	-	6.20	CGTGTGCGGGAGAAACTTC
IciA Escherichia coli (strain K12)	210	221	-	5.25	TGATTTTTTTAT
Nac Escherichia coli (strain K12)	210	220	+	10.16	ATAAAAAAATC
RcsAB Escherichia coli (strain K12)	203	210	-	5.81	TTGGAAAT

SP5

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	57	66	-	6.35	TGTTTATCAT
H-NS Escherichia coli (strain K12)	146	155	-	5.61	TGGATTTATT
IciA Escherichia coli (strain K12)	117	128	+	5.31	AGATATTTCTT
RcsAB Escherichia coli (strain K12)	121	128	-	5.06	AAGGAAAT

SP6

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	16	27	-	8.10	AGGAAATAGGTA
FhlA Escherichia coli (strain K12)	105	111	-	10.78	TTTTCGT
GcvA Escherichia coli (strain K12)	228	232	+	10.00	CTAAT
IciA Escherichia coli (strain K12)	149	160	-	5.09	TGATTTCCATT

RcsAB Escherichia coli (strain K12)	10	17	-	5.52	TAGGGGAA
--	----	----	---	------	----------

SP7

PWM (species)	Start Position	End Position	Strand	Score	Sequence
GcvA Escherichia coli (strain K12)	134	138	+	10.00	CTAAT
RcsAB Escherichia coli (strain K12)	160	167	+	5.13	TAGTGAAA

SP8

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	230	239	-	6.29	GGTTAAATCG
GcvA Escherichia coli (strain K12)	164	168	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	231	240	-	5.59	GGGTAAATC
IciA Escherichia coli (strain K12)	159	170	-	5.75	AGATTAGTTGGC

SP9

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArcA Escherichia coli (strain K12)	230	239	-	6.3	GGTTAAATCG
GcvA Escherichia coli (strain K12)	164	168	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	231	240	-	5.60	GGGTAAATC
IciA Escherichia coli (strain K12)	159	170	-	5.74	AGATTAGTTGGC
OxyR Escherichia coli (strain K12)	216	226	-	3.77	ATTATGAATAT

SP10

PWM (species)	Start Position	End Position	Strand	Score	Sequence
------------------	-------------------	-----------------	--------	-------	----------

CytR Escherichia coli (strain K12)	188	199	-	7.72	TTCAAATTTTGA
FhlA Escherichia coli (strain K12)	141	147	-	10.65	TTTTTCGT
GcvA Escherichia coli (strain K12)	58	62	-	10.00	CTAAT
GlpR Escherichia coli (strain K12)	103	122	+	6.69	ATTGCTCGCCAACTCTCGCC
IHF Escherichia coli (strain K12)	150	165	-	6.48	CCTCAATCGCTTGACG
RcsAB Escherichia coli (strain K12)	46	53	+	5.53	TCGTAAT

SP11

PWM (species)	Start Position	End Position	Strand	Score	Sequence
H-NS Escherichia coli (strain K12)	206	215	+	5.88	CCGATTAAGA
IHF Escherichia coli (strain K12)	167	182	+	6.51	CCTCAAAATGTTTATG
IciA Escherichia coli (strain K12)	59	70	-	5.20	TGATAATCCGCC
OxyR Escherichia coli (strain K12)	165	175	+	3.85	TTCTCAAAAT

SP12

PWM (species)	Start Position	End Position	Strand	Score	Sequence
ArgR Escherichia coli (strain K12)	41	54	-	9.29	TGGGTCAAAATAGA
Fis Escherichia coli (strain K12)	124	138	-	4.42	GAATAAAAGCCAACC
GcvA Escherichia coli (strain K12)	215	219	+	10.00	CTAAT
H-NS Escherichia coli (strain K12)	130	139	-	6.04	CGAATAAAAG
IHF Escherichia coli (strain K12)	116	131	-	6.32	AGCCAACCGCTCGCAC

IciA Escherichia coli (strain K12)	90	101	-	5.12	GGATGAAGCGTA
RcsAB Escherichia coli (strain K12)	55	62	+	5.07	TCGTAATT
OxyR Escherichia coli (strain K12)	201	211	-	5.15	ATACACAAAAT

SP13

PWM (species)	Start Position	End Position	Strand	Score	Sequence
CytR Escherichia coli (strain K12)	4	15	-	7.88	TTGAAATGCCGT
FhlA Escherichia coli (strain K12)	136	142	-	10.48	ATTTTCGT
GcvA Escherichia coli (strain K12)	206	210	-	10.00	CTAAT
IciA Escherichia coli (strain K12)	102	113	-	5.54	GGATATAGTGGG
RcsAB Escherichia coli (strain K12)	19	26	-	5.34	TAGTGACT

SP14

PWM (species)	Start Position	End Position	Strand	Score	Sequence
Fis Escherichia coli (strain K12)	200	214	-	3.60	GAATACAAGTAGACC
H-NS Escherichia coli (strain K12)	104	113	+	6.22	TCTATAAAAAT
IHF Escherichia coli (strain K12)	175	190	+	6.27	TAGCAAAAATATTCATA
IciA Escherichia coli (strain K12)	86	97	-	5.41	CGATAATGAAGC
NagC (Selex) Escherichia coli (strain K12)	164	186	+	18.88	GATAGTTCGGGTAGCAAAAATATT
RcsAB Escherichia coli (strain K12)	175	182	+	5.08	TAGCAAAA
OxyR Escherichia coli (strain K12)	142	152	-	4.03	GATGAGAAAAC

RESEARCH ARTICLE

Open Access

Insights into *Klebsiella pneumoniae* type VI secretion system transcriptional regulation



Victor Augusto Araújo Barbosa and Leticia Miranda Santos Lery*

Abstract

Background: *Klebsiella pneumoniae* (KP) is an opportunistic pathogen that mainly causes respiratory and urinary tract infections. The frequent occurrence of simultaneously virulent and multiple drug-resistant isolates led WHO to include this species in the list of top priorities for research and development of therapeutic alternatives. The comprehensive knowledge of the molecular mechanisms underlying KP virulence may lead to the proposal of more efficient and specific drugs. One of its virulence factors is the Type VI Secretion System (T6SS), which contributes to bacterial competition, cell invasion and in vivo colonisation. Despite the few studies showing the involvement of T6SS in KP pathogenesis, little is known concerning the regulation of its expression. The understanding of regulatory mechanisms may give more clues about the function of the system and the possibilities of future interference in this process. This work aimed to standardise the annotation of T6SS genes in KP strains and identify mechanisms of their transcriptional regulation through computational predictions.

Results: We analyzed the genomes of Kp52.145, HS11286 and NTUH-K2044 strains to perform a broad prediction and re-annotation of T6SS genes through similarity searches, comparative and linear discriminant analysis. 38 genes were found in Kp52.145, while 29 in HS11286 and 30 in NTUH-K2044. Genes coding for iron uptake systems are encoded in adjacencies of T6SS, suggesting that KP T6SS might also play a role in iron import. Some of the T6SS genes are comprised in syntenic regions. 17 sigma 70-dependent promoter regions were identified in Kp52.145, 12 in HS11286 and 12 in NTUH-K2044. Using VirtualFootprint algorithm, binding sites for 13 transcriptional regulators were found in Kp52.145 and 9 in HS11286 and 17 in NTUH-K2044. Six of them are common to the 3 strains: OxyR, H-NS, RcsAB, GcvA, Fis, and OmpR.

Conclusions: The data presented herein are derived from computational analysis. Although future experimental studies are required to confirm those predictions, they suggest that KP T6SS might be regulated in response to environmental signals that are indeed sensed by the bacteria inside the human host: temperature (H-NS), nutrition-limitation (GcvA and Fis), oxidative stress (OxyR) and osmolarity (RcsAB and OmpR).

Keywords: Type VI secretion system, T6SS, *Klebsiella pneumoniae*, Transcriptional regulator; bacterial genome analysis

Background

The Type Six Secretion System (T6SS) was initially described when it was demonstrated that secretion of the Hcp (Hemolysin-Coregulated Protein) and VgrG (Valine-Glycine Repeats) proteins were independent of the other known secretion mechanisms [1, 2]. In pathogenic bacteria, secretion systems can be used at various stages of the bacterial infection pathway, such

as toxin export, cell adhesion and direct translocation of effectors into the host cell or to delivery toxins against competitor bacteria [3]. A great diversity of these effectors and toxins have been identified and associated with the T6SS, which makes it a versatile weapon [4–6].

A functional T6SS apparatus is composed of the products of at least 13 conserved genes (*tssA-M*) [7]. The T6SS components assemble in a contractile needle-shaped apparatus which translocates effectors to neighboring cells [8]. It comprises a transmembrane complex, attached to a baseplate complex and a long-tailed

* Correspondence: leticia.lery@ioc.fiocruz.br; leryleticia@gmail.com
 Cellular Microbiology Laboratory, Oswaldo Cruz Foundation - Oswaldo Cruz Institute, Av. Brasil, 4365 - Manguinhos, Rio de Janeiro, RJ CEP: 21040-900, Brazil



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

cytoplasmic tubular structure surrounded by a contractile sheath. The contraction of the sheath propels the inner tube through the membranes towards the target cell. Effector proteins are carried by the T6SS tip and/or tubular components and are released after the tube disassembly inside the target cell.

In silico analyses have demonstrated the presence of T6SS genes in several Gram-negative bacteria, with the numbers of orthologs varying in each bacterium [7, 9–11]. Besides, the genomes of some bacteria can encode the set of T6SS genes in more than one *locus* [7, 12]. Commonly, these *loci* are within pathogenicity islands - for instance, the *Pseudomonas aeruginosa* HSI (Hcp-secretion island) and the *Salmonella typhimurium* SCI (*Salmonella* centrisome island), as well as in the genomes of enteroaggregative *Escherichia coli*, and *Vibrio cholerae* [4, 12–14]. In addition to the genes encoding the 13 conserved components of T6SS, these genomic *loci* may encode toxins, antitoxins, adapters, and auxiliary proteins, as well as additional effectors [4, 7].

In another hand, some functional T6SS genes are found outside the referred genomic islands: they are referred to as T6SS “orphan” genes [12]. Therefore, identifying these genes in bacterial genomes is also crucial for understanding the functionality of T6SS.

Besides, bacteria may encode several functional ‘copies’ of each T6SS gene. The expression of the different isoforms of a component may vary according to the target cell and/or by the environmental condition in which the bacterium is sensing [5, 13]. The expression of these secretion complexes can be precisely regulated by transcriptional, translational and post-translational mechanisms [15, 16].

Klebsiella pneumoniae (KP) is a ubiquitous species in nature, a gut commensal, and a human opportunistic pathogen. It can cause a wide range of infections, including pneumonia, urinary tract infections, bacteremia and liver abscesses [17, 18]. Due to the frequent occurrence of multiple antibiotic-resistant isolates, *K. pneumoniae* is considered a global public health concern [19–21]. Capsule, fimbriae, lipopolysaccharide (LPS) and siderophores are important and well-characterized virulence factors from KP [17]. Recently, additional factors have been described, such as type II (T2SS) and type VI secretion systems (T6SS) [22–25]. However, there is significant heterogeneity among KP strains and those virulence factors may play different roles in different strains [17, 26].

So far, little is known about T6SS in *K. pneumoniae*. From a genomic perspective, genes coding for putative T6SS components are present in KP genomes. In some strains, T6SS genes are grouped mainly in 2 *loci* (for instance: NTUH-K2044, Kpn2146 and HS11286), while in others they are found in 3 *loci* (Kp52.145, MGH 78578, 342) [22, 24, 27]. As KP genomes were annotated using different approaches - some using automatic pipelines -

there is no uniform annotation of T6SS genes. In a functional perspective, Lawlor and colleagues (2005) screened a transposon library and found that 2 mutants in hypothetical protein-coding genes displayed decreased ability to infect mouse spleen [28]. Those proteins were later annotated as putative T6SS components. Recently was described that HS11286 strain T6SS secretes a phospholipase effector with antibacterial activity [24]. Moreover, it has been shown that T6SS mutants in NTUH-K2044 strain significantly reduced bacterial killing, the expression of type-1 fimbriae and adherence and invasion of epithelial cells [25]. Concerning T6SS regulation, it has been observed that strain Kp52.145 genes are expressed when the bacterium colonized mice lungs, but not when the bacterium grows in TCS culture medium [22]. In Kpn2146, an RNA-seq analysis revealed that most of the T6SS genes increased expression 24 h post-infection of macrophages [29]. In HS11286 was suggested that sub-inhibitory concentrations of antibiotics might regulate T6SS secretion [24]. In NTUH-K2044 strain, was demonstrated that histone-like nucleoid structuring protein (H-NS) binds regulatory region and inhibits *tssD* (Hcp/tube component) expression [25].

Based on growing evidence for T6SS relevance for KP pathogenesis and divergent genomic features of T6SS loci between KP strains, this work aimed to identify every T6SS gene in 3 KP strains and standardize T6SS gene annotation in KP. Moreover, we hypothesized that translational mechanisms could be involved in T6SS expression in KP. Therefore, we predicted transcriptional regulator’s binding sites upstream transcriptional start sites for T6SS genes and obtained insights into T6SS role and regulation.

Results

Identification of T6SS genes in KP genomes

Previous studies have identified T6SS genes in KP genomes. Sarris and colleagues (2011) annotated T6SS genes in three fully sequenced KP strains (342, NTUH-K2044 and MGH78578) and one partially sequenced strain (KP subsp. rhinoscleromatis ATCC13884) [27]. However, at that time there were few KP complete genomes available for comparison. In this study we present a similar effort. However, we focused on the human pathogenic strains: Kp52.145 (a K2 virulent strain), HS11286 (a multiple-drug-resistant strain) and NTUH-K2044 (a K1 virulent strain). Besides, we propose herein the use of the TssA-M nomenclature for the T6SS core components in KP. This nomenclature overcome the problem that T6SSs in different organisms had historically acquired different, system-specific names for equivalent components [7]. Thus, the genomes of the 3 KP strains mentioned above were reanalyzed to identify and re-annotate all putative T6SS protein-coding genes

(*tssA-M*, *paar*, *tagA* and *tagL*). The starting point for such annotation process was the analysis provided by SecreT6 database. Then, VRProfile, Blast, COG, PFAM and CDD searches were also performed (Additional file 1). The same search and filtering criteria were applied to the 3 genomes.

In overall, we detected 38, 29 and 30 genes possibly related to T6SS in the Kp52.145, HS11286, and NTUH-K2044 genomes, respectively (Table 1).

Kp52.145

Its genome annotation has been updated recently (FO834906.1 from 07-MAR-2015/NZ_FO834906.1 from 21-FEB-2017). We identified that six T6SS-related genes identified by SecreT6 in the initial genome version were re-annotated as pseudogenes in the

most recent version (BN49_RS18740, BN49_RS14045, BN49_RS14055, and BN49_RS05960) - therefore, they do not seem to code for T6SS components and indicate that these genes may be under genome reduction process.

In agreement with previous study by Lery et al. (2014), we found that most of the T6SS genes in Kp52.145 genome (33 genes, 87%) are clustered in 3 genomic loci and only 5 (12%) are orphan genes: *tssH* (BN49_RS07300, BN49_RS08275, BN49_RS11635) and *tssD* (BN49_RS06500 e BN49_RS15545). Apart from the genes identified in this previous study, we found 7 additional putative T6SS: 1 gene coding for a PAAR domain-containing protein, 1 gene coding for an OmpA-family protein, 3 genes encoding putative TssH (Clp ATPases) components and 2 genes coding for TssD components. All the

Table 1 T6SS components in *K. pneumoniae* strains Kp52.145, HS11286 and NTUH-K2044. Locus tag for each gene is shown

	Kp52.145	HS11286	NTUH-K2044
<i>tssA</i>	BN49_RS05995; BN49_RS14080; BN49_RS18715	KPHS_23150; KPHS_32450;	KP1_RS11300; KP1_RS15690;
<i>tssB</i>	BN49_RS14020;	KPHS_22970;	KP1_RS11220; KP1_RS11085;
<i>tssC</i>	BN49_RS14025;	KPHS_22980;	KP1_RS11225;
<i>tssD</i>	BN49_RS06500; BN49_RS15545;	KPHS_23020; KPHS_41670;	KP1_RS11245; KP1_RS12525;
<i>tssE</i>	BN49_RS05940; BN49_RS18720;	KPHS_32460;	KP1_RS15695;
<i>tssF</i>	BN49_RS05970; BN49_RS14085; BN49_RS18735;	KPHS_23170; KPHS_32490;	KP1_RS11315; KP1_RS15710;
<i>tssG</i>	BN49_RS05965; BN49_RS14090; BN49_RS18730;	KPHS_23180; KPHS_32480;	KP1_RS11320; KP1_RS15705;
<i>tssH</i>	BN49_RS14050; BN49_RS07300; BN49_RS08275; BN49_RS11635;	KPHS_23030; KPHS_39850; KPHS_17930; KPHS_11410;	KP1_RS11250; KP1_RS19445; KP1_RS06015; KP1_RS08880;
<i>tssI</i>	BN49_RS06025; BN49_RS18800;	KPHS_23040; KPHS_32730;	KP1_RS11255; KP1_RS15775;
<i>tssJ</i>	BN49_RS05945; BN49_RS14095; BN49_RS18725;	KPHS_32470;	KP1_RS11325; KP1_RS15700;
<i>tssK</i>	BN49_RS06040; BN49_RS14030; BN49_RS18815;	KPHS_22990; KPHS_32770;	KP1_RS11230; KP1_RS15790;
<i>tssL</i>	BN49_RS06035; BN49_RS14035; BN49_RS18810;	KPHS_23000; KPHS_32760;	KP1_RS11235; KP1_RS15785;
<i>tssM</i>	BN49_RS06000; BN49_RS14075;	KPHS_23140; KPHS_32500;	KP1_RS11295; KP1_RS15715;
<i>Paar</i>	BN49_RS05990; BN49_RS14070; BN49_RS18750;	KPHS_23120; KPHS_32520;	KP1_RS15725; KP1_RS11305;
<i>tagL</i>	BN49_RS06030; BN49_RS14040; BN49_RS18805;	KPHS_23010; KPHS_32750;	KP1_RS11240; KP1_RS15780;

38 putative T6SS-related genes in Kp52.145 genome are described in Additional file 2.

HS11286

A previous study revealed a 23-gene T6SS cluster (KPHS_22970 to 23,190) on the chromosome of KP HS11286, containing 12 core T6SS components [24]. In our analysis we found 25 genes annotated as T6SS-related by SecreT6. Four additional genes were identified by further analysis: 3 genes encoding putative TssH components (KPHS_39850, KPHS_17930, KPHS_11410) and one gene encodes a TssD component (KPHS_41670). 25 out of the 29 (86%) T6SS-genes are clustered in two main *loci* and only 4 (14%) are orphan genes: 3 *tssH* and 1 *tssD*. All the 29 putative T6SS-related genes in HS11286 genome are described in Additional file 2.

NTUH-K2044

25 genes were annotated as T6SS-related by SecreT6. Besides those genes, we found 3 genes encoding putative TssH components, 1 TssB and 1 TssD that might be part of T6SS in this strain. Accordingly to a previous study, most of the genes (25 genes, 83%) are clustered in 2 loci, and only 5 genes (17%) are orphan: *tssB* (KP1_RS11085), *tssD* (KP1_RS12525), *tssH* (KP1_RS19445, KP1_RS06015, KP1_RS08880). All the 30 putative T6SS-related genes in NTUH-K2044 genome are described in Additional file 2.

Annotation of Clp ATPases: putative TssH components

ClpV ATPases have been initially described as the TssH member of T6SS complex, required for a functional tube formation and recycling of sheath components [30–32]. ClpV proteins in KP genomes are encoded by genes BN49_RS14050 in Kp52.145, KPHS_23030 in HS11286 and KP1_RS11250 in NTUH-K2044. Kp ClpV proteins are 99% identical.

Recently has been shown that *Francisella tularensis* has a non-canonical functional T6SS that uses ClpB ATPase instead of ClpV [33, 34]. Interestingly, the *clpB* gene KP1_RS19445 (former KP1_4170) from KP NTUH-K2044 has been associated to T6SS [27]. Therefore, we expanded the search for Clp ATPases in KP genomes - as they could be putative TssH. In addition to *clpV* genes, SecreT6 and VRprofile predicted 3 additional Clp-family ATPases, belonging to COG0542, encoded in each KP genome analyzed: ClpB, ClpA e ClpX.

According to the deposited genome annotation, Kp52.145, HS11286 and NTUH-K2044 ClpB proteins would be 857, 823 and 857 aminoacids long. According to our analysis, the start codon for HS11286 *clpB* has been previously misannotated. We suggest it is 102 bp longer, coding a protein of 857 aminoacids. Considering such re-annotation, protein sequences from the 3 KP

strains are 100% identical. KP ClpB protein is quite conserved to *F. tularensis* ClpB: 64% sequence identity and 79% similarity (Fig. 1). In addition, KP ClpB presents the same conserved domains as its *F. tularensis* orthologue: Clp_N, ClpB_D2small, AAA and AAA2 domains, as well as Walker A, Walker B, ATP binding site and arginine finger motifs - suggesting that they might be able to perform similar roles. Further studies are required to confirm it.

ClpA proteins from the 3 KP strains share 99% identity. In comparison to ClpV sequence they are ~53% similar and 35% identical in an alignment of 87% coverage. Despite such sequence dissimilarities, ClpA, ClpB and ClpV contain conserved domains (Additional file 3). The KP ClpX sequences are identical among the 3 strains, however ClpX is 424 aminoacids long, while ClpV is 884. Whether ClpA and ClpX could play the role of TssH member at T6SS complex remains to be elucidated.

Genomic context of T6SS genes

Bacterial genes required for the same functional process are often clustered in the same genomic region and frequently subjected to the same regulatory network. As mentioned above, most of the KP T6SS genes are clustered in 2 or 3 regions per genome, presenting characteristics of genomic islands (GC content different from the average genome, inserted in tRNA loci, containing transposase or other mobile elements). None of those islands encode any transcriptional regulator. We analysed the function of the gene products neighbouring T6SS genes to get insights into processes that could be co-regulated and identify putative regulatory proteins. Curiously, we found genes encoding conserved LysR-type transcriptional regulators from superfamily PBP2 ~ 3 to 4 kb from several T6SS *loci* (Table 2).

The 3 genomes also presented non-clustered T6SS-related 'orphan' genes. We found outer membrane ion transporters encoded in genes neighboring orphan *tssD* genes (Fig. 2). In Kp52.145, the *tssD* encoded by BN49_RS06500 is located in a region encoding hypothetical proteins, iron/heme ABC-family transporters (BN49_RS06505 to BN49_RS06520) and a TonB-dependent receptor (BN49_RS06525). In HS11286 the *tssD* orphan gene (KPHS_41670) neighbors genes that encode an hemin ATP transport system (KPHS_41660 to KPHS_41640), an S-adenosylmethionine-dependent methyltransferase (KPHS_41630) and an outer membrane receptor for ferric enterobactin and colicins B and D (KPHS_41620). Although these iron-related genes in Kp52.145 and HS11286 have a different names, they are orthologs and present the same sequence. Thus, both orphan *tssD* are encoded in the same genomic region. Interestingly, recently has been

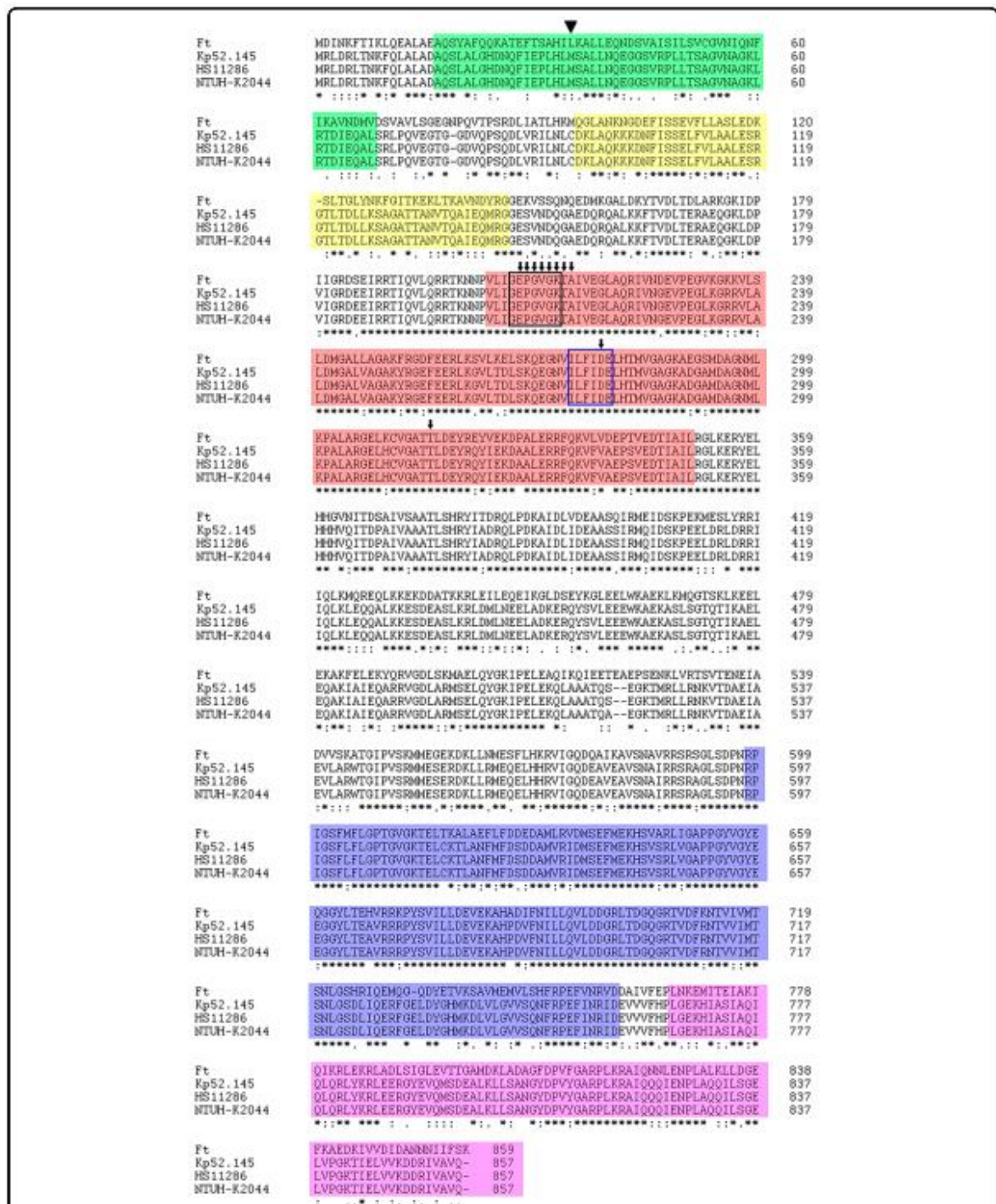


Fig. 1 KP ClpB proteins are quite similar to *F. tularensis* ClpB: Multiple sequence alignment of ClpB proteins from *F. tularensis* SCHU (Ft), *K. pneumoniae* Kp52.145, HS11286 and NTUH-K2044. Green and yellow regions indicate the Clp_N domain; Red, blue and purple are the AAA, AAA2 and Clp_D2-small domains, respectively. Black and blue boxes indicate the Walker a and Walker b domains; Arrows point to ATP binding site residues. The arrow head indicate the position 1 of ClpB from HS11286 according to its previous annotation. "*"Identity of all aminoacids in the indicated column, "*"Similarity of aminoacids aligned, "*" Low similarity of aminoacids aligned, according to ClustalW default parameters

Table 2 Transcriptional regulators predicted to be involved in KP T6SS regulation

Transcriptional regulator	Locus-Tag			T6SS relationship*
	Kp52.145	HS11286	NTUH-K2044	
Fis	BN49_RS03455	KPHS_48020	KP1_RS23255	Binding sites predicted at 17, 12, 12 promoters
OxyR	BN49_RS25290	KPHS_01030	KP1_RS00535	Binding sites predicted at 15, 10, 12 promoters
H-NS	BN49_RS18480	KPHS_31980	KP1_RS15455	Binding sites predicted at 13, 8, 10 promoters
OmpR	BN49_RS02915	KPHS_27610	KP1_RS23790	Binding sites predicted at 9, 6, 7 promoters
GcvA	BN49_RS13685	KPHS_42400	KP1_RS10795	Binding sites predicted at 8, 3, 6 promoters
RcsAB	BN49_RS21260	KPHS_37040	KP1_RS18010	Binding sites predicted at 3, 4, 5 promoters
LysR_PBP2	BN49_RS18845	KPHS_32830	KP1_RS15820	Encoded 3058 bp, 3059 bp, 3059 bp from T6SS cluster
	BN49_RS13995	-	KP1_RS11195	Encoded 4025 bp, -, 4024 bp from T6SS cluster

* the numbers indicated correspond to strains Kp52.145, HS11286 and NTUH-K2044, respectively

shown that in response to oxidative stress, T6SS from *Yersinia pseudotuberculosis* and *Burkholderia thailandensis* secrete effectors involved in ion uptake. The ion import is further mediated by ABC and/or TonB-family proteins [35–37].

In the neighborhood of the putative *tssH/clpB* (BN49_RS07300 in Kp52.145 and KP1_RS19445 in NTUH-K2044) there is *pgeF* - coding for a peptidoglycan editing factor. PgeF contributes to the maintenance of the peptidoglycan peptide chain composition in *E. coli*, thus contributing to the integrity of the bacterial peptidoglycan layer [38]. Moreover, in the neighborhood of *clpX*, another putative *tssH* (BN49_RS08275, KPHS_11410 and KP1_RS06015) we found a *clpP* ATPase and the LonA endopeptidase; and besides the *clpA tssH* (BN49_RS11635, KPHS_17930 and KP1_RS0880), a *clpS* ATPase and macrolide transport proteins. Further experiments will be performed to check whether these gene products could be related to or secreted by T6SS, used for bacterial competition and/or tissue invasion.

Syntenicity among T6SS loci in the 3 strains

We analyzed the synteny and sequence conservation among the genomic regions coding for T6SS genes in Kp52.145 (Kp52_R1; Kp52_R2; Kp52_R3), HS11286 (HS_R1; HS_R2) and NTUH-K2044 (NT_R1; NT_R2). Pairwise alignments of the entire regions coding for T6SS genes were visualized in ACT software. Regions of similarity and differences among the strains are pointed in Fig. 3. Two distinct regions > 24 kb were aligned among the 3 genomes (Fig. 3, a and b). In both regions, genes annotated as T6SS-related are encoded in the beginning and end of those sequences, while the middle regions contain genes coding for other functions. Curiously, the T6SS-related genes are syntenic among the 3 strains. Syntenic regions with > 99% sequence identity code for *tssK-tssL-tagL-tssI* and *paar-tssM-tssF-tssG-tssJ-tssE-tssA* (Fig. 3a) / *tssB-tssC-tssK-tssL-tagL-tssD-tssH-*

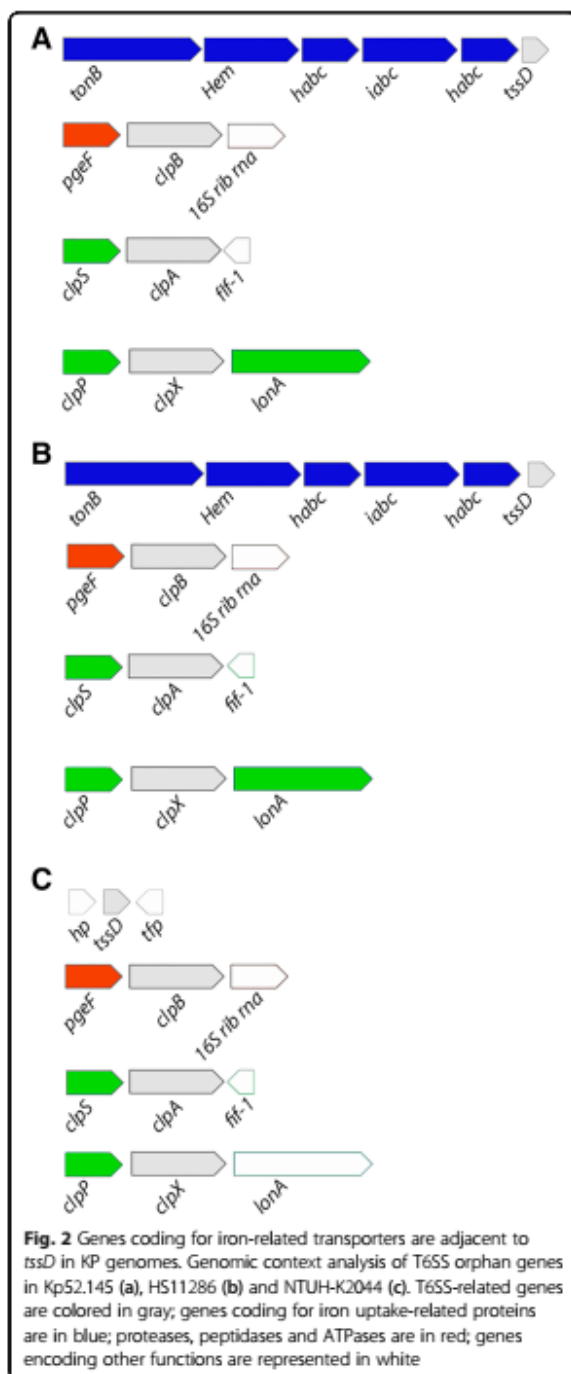
tssI and *tssM-tssA-tssF-tssG-tssJ* (Fig. 3b). It is worthy to note that *tssD* and *tssI* at Kp52_R2 and *tssM* in KP52_R3 are currently annotated as pseudogenes.

In another hand, the insertions containing T6SS-unrelated genes are less conserved or strain-specific, encoding mainly proteins of unknown function or transposases. Regions HS_R2 and NT_R2 share short regions of similarity varying from 87 to 96% identity. In another hand, the insertion observed in Kp52_R3 contains genes coding for phospholipases, Sel-1 lipoproteins and a PAAR protein. Such region has been previously characterized as implied in KP virulence [22].

Transcriptional regulator's binding sites

To get insights into T6SS transcriptional regulation in KP, we identified 70 promoter consensus sequences – 10 and – 35 upstream the CDSs coding for the T6SS-related genes (Additional file 4). In overall, 17 putative 70-dependent transcriptional start sites were found in Kp52.145 (Fig. 4a), 12 in HS11286 (Fig. 4b) and 12 in NTUH-K2044 (Fig. 4c). 250 bp upstream each of those transcriptional start sites were analyzed using position-weight matrices to identify putative binding sites (Additional file 5). In Kp52.145, 106 binding sites for 13 transcriptional regulators were predicted: Crp, CytR, FhlA, Fis, Fnr, GcvA, H-NS, MalT, MetR, Mlc, OmpR, OxyR and RcsAB (Additional file 6). In HS11286, there were 72 binding sites for 09 transcriptional regulators: Fis, GcvA, H-NS, Lrp, NarL, OmpR, OxyR, PdhR and RcsAB (Additional file 7). In NTUH-K2044 there were 114 binding sites for 17 transcriptional regulators: Crp, FadR, FhlA, Fis, Fnr, GcvA, GlpR, H-NS, Lrp, MalT, MetJ, MetR, Mlc, OmpR, OxyR, PdhR and RcsAB (Additional file 8).

Six of these regulators had binding sites upstream T6SS genes predicted in the three strains: Fis, OxyR, H-NS, OmpR, GcvA and RcsAB (Table 2) - indicating that to some extent there might be conservancy of regulatory mechanisms. Curiously, ~ 85% of the



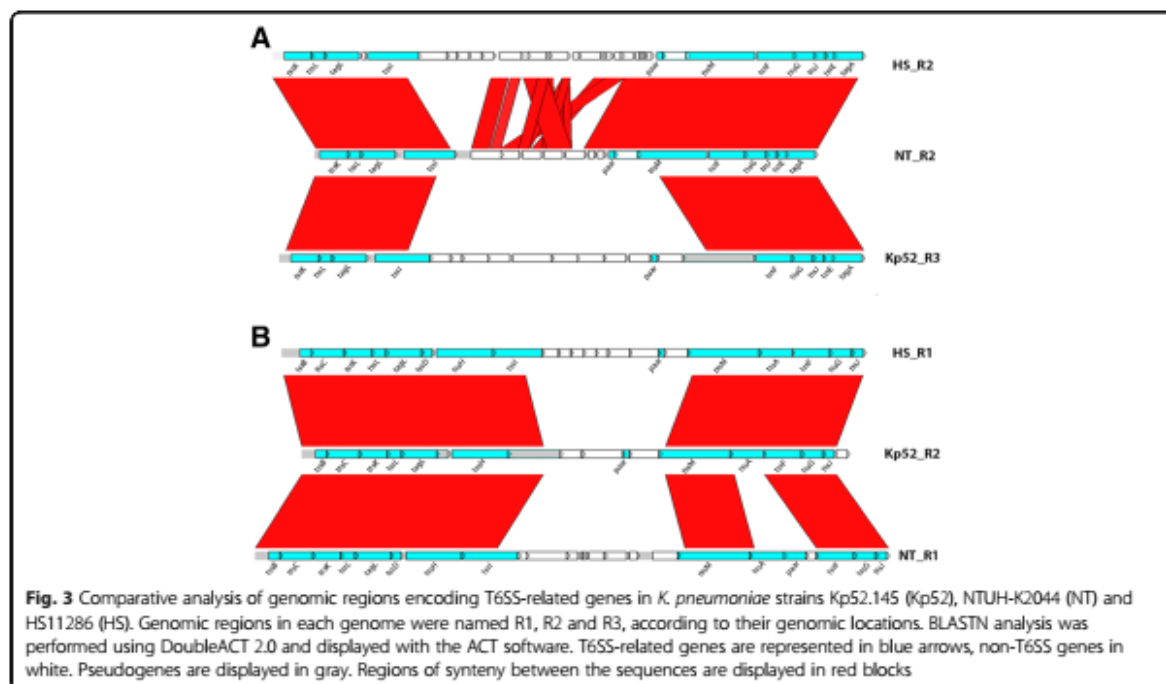
promoter sequences analyzed contain conserved Fis and OxyR binding sites, and ~75% for H-NS. All those putative binding sites described above were computationally predicted. Further studies are required to experimentally validate them.

Discussion

KP virulence factors content and expression vary among different isolates. T6SS genes are largely distributed in KP strains. Previous studies identified T6SS orthologs in KP strains 342, NTUH-K2044, MGH78578, HS11286 and BAA2146 [24, 25, 27]. In this study we performed a robust computational analysis of T6SS genes and genomic context, as well as putative protein sequences to get more insights into T6SS role in KP. Comparing the genomes of 3 human pathogenic KP strains (NTUH-K2044, HS11286 and Kp52.145), the analysis presented herein allowed us to propose a standardization of T6SS genes nomenclature in KP. The results presented herein show that KP T6SS core genes encoded in the genome of 3 KP strains are somewhat conserved in terms of sequence similarity, gene content and operon structure. In addition, they are similar to the T6SS-encoding regions in *E. coli* E042 [4] and *Pantoea ananatis* [39]. Using the approach described above, we identified putative T6SS orthologs not described previously - specially putative TssH family proteins. TssH is a ClpV ATPase shown to be involved in T6SS disassembly and subunits recycling [30, 31]. However, recently has been shown that other Clp/Hsp100 family proteins such as ClpB from *Francisella tularensis* might play a similar role [33, 34]. The approach performed in this study identified ClpB, ClpA, ClpX and ClpV proteins as putative TssH. Due to the high similarity between KP and *F. tularensis* ClpB, we propose that KP ClpB proteins might function as a TssH. However, whether those proteins are indeed required and/or assembled as part of T6SS in KP remains to be elucidated.

We observed that the majority of the T6SS genes in KP are clustered in 2 or 3 genomic islands - in agreement with Sarris et al. [27]. Interestingly, we identified genes coding for PAAR proteins in every T6SS island of the 3 strains. We observed that T6SS core genes are presented in syntenic blocks, whilst insertions of variable content are strain-specific. Those insertions mainly code for proteins of unknown function, but also for putative T6SS accessory or effector proteins. A functional relationship between T6SS and such genes inside insertions is provided by the transcriptomic analysis of Bent et al. showing the co-expression of some of those genes [29].

So far, KP T6SS had been associated to antibacterial activity, cell invasion and in vitro colonization [24, 25]. The genomic context analysis described herein pointed that several iron-related transporters are encoded around T6SS genes. This observation raises the hypothesis that KP T6SS could play a role in ion uptake. For instance, *Y. pseudotuberculosis* T6SS transports Zn^{2+} to Combat Multiple Stresses and Host Immunity [35]. *B. thailandensis* uses T6SS to uptake Mn^{2+} ions, to resist oxidative stress and compete with other bacteria [36].

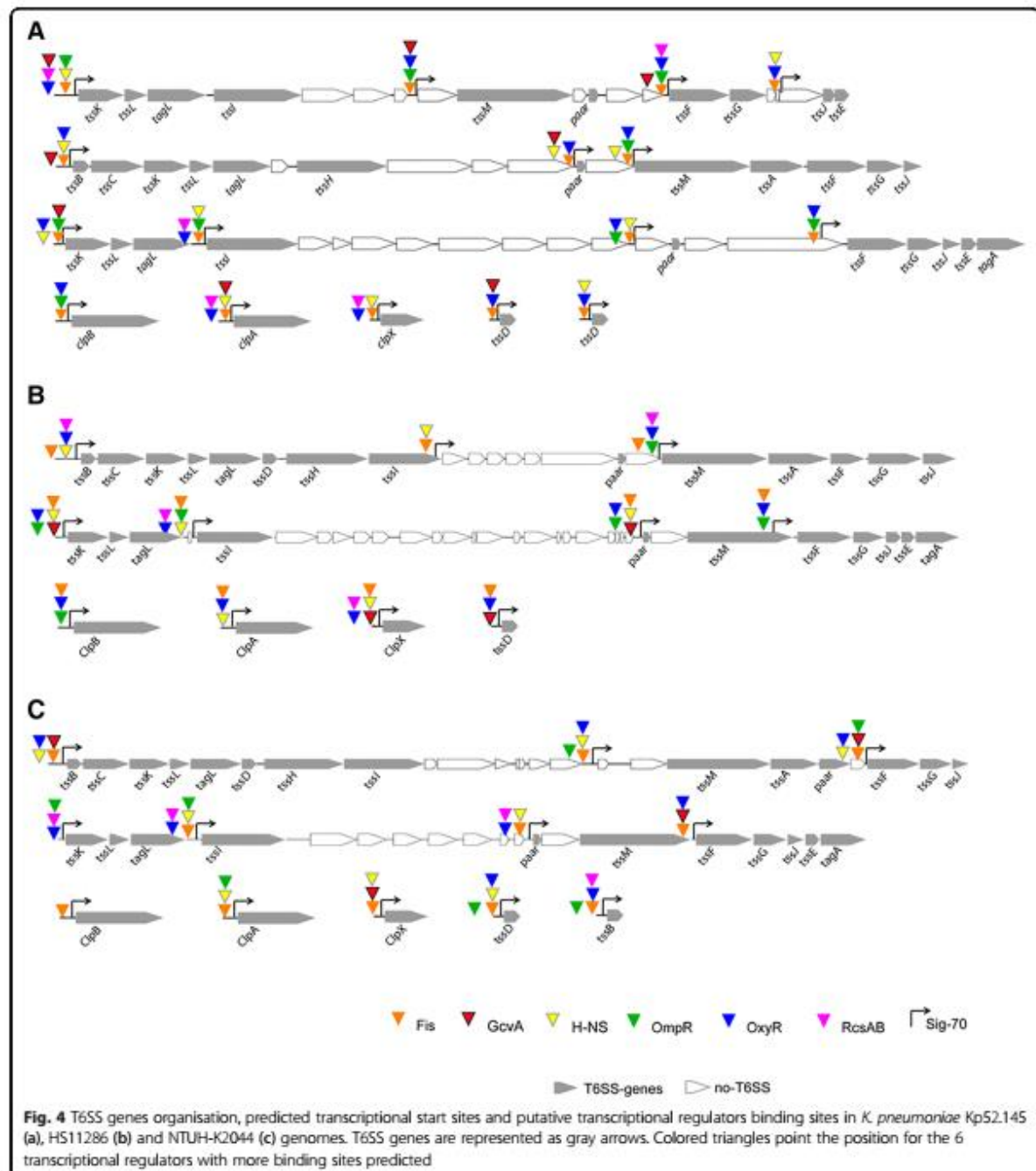


Such *B. thailandensis* T6SS is regulated by OxyR, a conserved oxidative stress response transcriptional regulator. Reactive oxygen species (ROS) can damage bacterial cells, thus, bacteria detoxify ROS by producing ROS-detoxifying enzymes, DNA repair, and sequestration of metal ions. One of those mechanisms involves catalases. Curiously, OxyR regulates the expression of KatN, a catalase secreted in a T6SS-dependent manner by enterohemorrhagic *E. coli* [40]. OxyR is one of the regulators that induce the bacterial oxidative stress response. In addition, it has been demonstrated that KP OxyR regulates biofilm formation, fimbrial genes, antibiotic resistance and adhesion to epithelial cells [41, 42]. Interestingly, we predicted conserved OxyR binding sites in most of the KP T6SS promoter regions, in all 3 strains analyzed. Altogether, these data suggest that KP T6SS might be activated under oxidative stress conditions and might help the bacteria to deal with ROS detoxification.

Another interesting finding was that putative H-NS binding sites are widely present and conserved in KP T6SS promoters. It has been previously shown that H-NS silencing of a T6SS locus limits *Salmonella enterica* interbacterial killing [43]. H-NS is a nucleoid structuring protein with global effects on silencing gene expression [44]. Its activity depends on temperature and osmolarity. It has been shown that KP H-NS represses the expression of important virulence factors, such as type-3 pili and capsule [45, 46]. Recently, Hsieh et al. (2018) showed that H-NS binds to *tssD* promoter and silences

tssD (KP1_RS11245) expression in NTUH-K2044. Thus, our prediction was confirmed by this recent experimental data.

Putative binding sites for Additional 4 regulators were predicted in promoter sequences of T6SS genes of all 3 strains: RcsAB, GcvA, Fis and OmpR. RcsAB is an unusual regulatory system that binds an *rscAB* box and modulates KP *galF* gene, thus affecting capsule expression and virulence [47, 48]. GcvA is the transcriptional regulator of the glycine cleavage system, involved in aminoacids metabolism. GcvA has not yet been studied in KP, however, it is required for *F. tularensis* fitness and full virulence [49]. Fis (factor for inversion stimulation) is a transcriptional regulator that respond to changes in the nutritional environment in enterobacteria [50]. OmpR is the response regulator of a two-component system with the sensor kinase EnvZ. OmpR binds to the promoter region of a *Yersinia pseudotuberculosis* T6SS involved in the bacterial survival in high osmolarity conditions, resistance to deoxycholate and pH homeostasis [51, 52]. It has been shown that KP OmpR regulates c-di-GMP signaling pathway, type 3 fimbriae expression, and biofilm formation in response to osmotic stresses [53]. Altogether, these results comprise quite interesting data suggesting that KP T6SS may be regulated in response to environmental signals that are indeed sensed by the bacteria inside the human host: temperature (H-NS), nutrition-limitation (Fis), oxidative stress (OxyR) and osmolarity (RscAB and OmpR). Strikingly, H-NS,



OxyR and OmpR regulators are expressed by KP BAA2146 during in vitro macrophages infection [29].

Besides, we found conserved genes encoding LysR transcriptional regulators (LTTR) containing PBP2-like substrate binding domains, in the adjacencies of almost every T6SS gene cluster in the three KP strains studied in this work. Their consensus binding sequences have

not yet been determined, thus we did not manage to further predict if they may indeed regulate KP T6SS.

It is important to highlight that, so far, we analyzed only sigma 70-dependent promoter sequences. Probably, other sigma promoters are involved in T6SS transcriptional regulation. In fact, Hsieh et al. detected that, in NTUH-K2044, *tssB-tssC-tssK-tssL-tagL-tssD* form a

single transcriptional unit and that *tssH* and *tssI* genes are independently transcribed [25]. In our analysis we detected the sigma 70-dependent promoter upstream *tssB*, but not *tssH* and *tssI*. Thus, *tssH* and *tssI* might have alternative promoters - reinforcing that it will be useful in the future to expand this analysis to other promoters.

Recently, Ho et al. (2014) proposed a model for T6SS assembly and activity [54]. In such model, the baseplate components (TssAEFGJKLM) that anchor the system through the bacterial membranes are the first to be assembled [47]. Then, the coating proteins (TssI and Paar), contributing to the overall stability of the apparatus, are recruited. In a third step, tube and sheath proteins (TssBCD) are assembled. At last, TssH ATPase is required. Interestingly, we have identified binding sites for RcsAB and OmpR - transcriptional regulators related to osmotic stress - in the promoters of all baseplate and coating coding genes, but in none of the other components. In another hand, we found that every gene coding for tube, sheath or ATPase component might be regulated by OxyR. Although those predictions are still to be confirmed and additional promoters and regulators are expected to be involved, we hypothesize that at least two signals are required for the expression of KP T6SS.

Conclusions

We presented a genomic analysis of 3 KP strains and provided new insights into T6SS role and regulation: 1) T6SS components annotation was standardised among the strains, 2) considering that KP ClpB protein sequences are quite conserved to *F. tularensis* ClpB, we suggest that they may act as TssH, 3) it was hypothesized that T6SS in KP might play a role in iron uptake, and 4) OxyR, H-NS, RcsAB, GcvA, OmpR, Fis, and LysR/PBP2 family proteins were predicted as putative regulators.

Methods

K. pneumoniae genomic sequences

The nucleotide sequences and annotations of *K. pneumoniae* Kp52.145 (NZ_FO834906, NZ_FO834904 and NZ_FO834905), HS11286 (NC_016845.1, NC_016838.1 and NC_016846.1) and NTUH-K2044 (NC_012731.1 and NC_006625.1) chromosomes and plasmids were downloaded from the NCBI RefSeq database.

In silico identification of T6SS-encoding genes

T6SS gene clusters were identified by SecReT6 [10] and VRprofile [9] algorithms, considering e-value $\leq 10^{-10}$ and identity $\geq 70\%$.

Conserved domains were identified in CDD (Conserved Domain Database), PFAM (Protein Families

Database), and COG (Clusters of Orthologous Groups) databases [55–57]. Sequence alignments were performed with Clustal [58].

Syntenic analysis

Nucleotide sequences were aligned using the Double-ACT v2.0 online resource and visualized in the Artemis Comparison Tool [59]. Syntenic regions (% identity > 87) were represented as solid red blocks.

Computational prediction of promoter sequences and operons

Genomic regions containing 250 bp upstream each of the previously identified T6SS genes were analyzed using Bprom algorithm [60]. The Bprom algorithm identifies putative binding sites for the sigma-70 factor.

Prediction of protein binding sites in promoter sequences

The 250 bp upstream each transcription start site identified above was analyzed in the Virtual Footprint prediction tool [61]. Virtual Footprint compares query sequences to a library of Position-Weight Matrixes (PWM) from Prodigic database. Those matrixes represent known transcriptional regulators binding sites.

Additional files

Additional file 1: Summary of SecReT6 and VRprofile results contributing to the annotation of T6SS genes in KP. Table describing T6SS genes in KP genomes: old and current locus-tag, genomic coordinates, gene orientation, gene size, gene product, SecReT6 and VRprofile predictions. (XLS 56 kb)

Additional file 2: Identification and location of T6SS components in the genome of KP Kp52.145, HS11286 and NTUH-K2044. Table describing locus tag, previous annotation and indicated if T6SS components to belong in loci or are orphan. (XLS 284 kb)

Additional file 3: Multiple sequence alignment of ClpV, ClpB and ClpA proteins from *Klebsiella pneumoniae* Kp52.145. Green and yellow regions indicate the Clp_N domain; Red, blue and purple are the AAA, AAA2 and ClpB_D2-small domains, respectively. Black and blue boxes indicate the Walker A and Walker B domains; Arrows point to ATP binding site residues. ***Identity of all aminoacid residues in the indicated column, **Similarity of pairs aligned, *Low similarity of pairs aligned. (PDF 58 kb)

Additional file 4: Predicted 70 promoter sequences of T6SS genes. Binding sites - 10, spacer region, - 35 and Transcriptional Start Site upstream the CDSs coding for the T6SS-related genes in genome of *Klebsiella pneumoniae* Kp52.145, HS11286 and NTUH-K2044. (XLS 98 kb)

Additional file 5: Putative promoter sequence of T6SS genes. 250 bp upstream those transcriptional start sites of T6SS genes in genome of *Klebsiella pneumoniae* Kp52.145, HS11286 and NTUH-K2044. (TXT 11 kb)

Additional file 6: Hypothetical binding sites for transcriptional regulators predicted in putative promoter sequence of T6SS genes in Kp52.145. Position Weight Matrix (PWM), Start and End position, Strand, Score and Sequence binding for transcriptional regulators. (XLS 55 kb)

Additional file 7: Hypothetical binding sites for transcriptional regulators predicted in putative promoter sequence of T6SS genes in HS11286. Position Weight Matrix (PWM), Start and End position, Strand, Score and Sequence binding for transcriptional regulators. (XLS 42 kb)

Additional file 8: Hypothetical binding sites for transcriptional regulators predicted in putative promoter sequence of T6SS genes in NTUH-K2044. Position Weight Matrix (PWM), Start and End position, Strand, Score and Sequence binding for transcriptional regulators. (XLS 47 kb)

Acknowledgements

The authors acknowledge the members of Laboratório de Microbiologia Celular and Pós-Graduação em Biologia Computacional e Sistemas at IOC/Fiocruz for their valuable suggestions.

Authors' contributions

VAAB performed all the experiments herein described. LMSL conceived the study and analyzed the data. Both authors wrote and approved the final manuscript.

Funding

This study was financed by: 1) Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 2) Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), 3) Fundação de Amparo à Pesquisa no Estado do Rio de Janeiro (FAPERJ) 4) Instituto Oswaldo Cruz/Fundação Oswaldo Cruz and 5) Programa Print-Fiocruz-CAPES.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 13 January 2019 Accepted: 5 June 2019

Published online: 18 June 2019

References

- Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, Mekalanos JJ. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dityostelium* host model system. *Proc Natl Acad Sci U S A*. 2006;103(5):1528–33.
- Mougous JD, Cuff ME, Raunser S, Shen A, Zhou M, Gifford CA, Goodman AL, Joachimiak G, Ordoñez CL, Lory S, et al. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science*. 2006;312(5779):1526–30.
- Leiman PG, Basler M, Ramagopal UA, Bonanno JB, Sauder JM, Pukatzki S, Burley SK, Almo SC, Mekalanos JJ. Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc Natl Acad Sci U S A*. 2009;106(11):4154–9.
- Joumet L, Cascales E. The type VI secretion system in *Escherichia coli* and related species. *EcoSal Plus*. 2016;7(1).
- Sana TG, Berni B, Blevess S. The T6SSs of *Pseudomonas aeruginosa* strain PAO1 and their effectors: beyond bacterial-cell targeting. *Front Cell Infect Microbiol*. 2016;6:61.
- Russell AB, Peterson SB, Mougous JD. Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol*. 2014;12(2):137–48.
- Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I. Dissecting the bacterial type VI secretion system by a genome wide *in silico* analysis: what can be learned from available microbial genomic resources? *BMC Genomics*. 2009;10:104.
- Nguyen VS, Douzi B, Durand E, Roussel A, Cascales E, Cambillau C. Towards a complete structural deciphering of type VI secretion system. *Curr Opin Struct Biol*. 2018;49:77–84.
- Li J, Tai C, Deng Z, Zhong W, He Y, Ou HY. VRprofile: gene-cluster-detection-based profiling of virulence and antibiotic resistance traits encoded within genome sequences of pathogenic bacteria. *Brief Bioinform*. 2018;19(4):566–74.
- Li J, Yao Y, Xu HH, Hao L, Deng Z, Rajakumar K, Ou HY. SecReT6: a web-based resource for type VI secretion systems found in bacteria. *Environ Microbiol*. 2015;17(7):2196–202.
- Shrivastava S, Mande SS. Identification and functional characterization of gene components of type VI secretion system in bacterial genomes. *PLoS One*. 2008;3(8):e2955.
- Blondel CJ, Jiménez JC, Contreas I, Santiviago CA. Comparative genomic analysis uncovers 3 novel loci encoding type six secretion systems differentially distributed in *Salmonella* serotypes. *BMC Genomics*. 2009;10:354.
- Chen L, Zou Y, She P, Wu Y. Composition, function, and regulation of T6SS in *Pseudomonas aeruginosa*. *Microbiol Res*. 2015;172:19–25.
- Labbate M, Orata FD, Petty NK, Jayatilake ND, King WL, Kirchberger PC, Allen C, Mann G, Mutreja A, Thomson NR, et al. A genomic island in *Vibrio cholerae* with VP-1 site-specific recombination characteristics contains CRISPR-Cas and type VI secretion modules. *Sci Rep*. 2016;6:36891.
- Silverman JM, Brunet YR, Cascales E, Mougous JD. Structure and regulation of the type VI secretion system. *Annu Rev Microbiol*. 2012;66:453–72.
- Miyata ST, Bachmann V, Pukatzki S. Type VI secretion system regulation as a consequence of evolutionary pressure. *J Med Microbiol*. 2013;62(Pt 5):663–76.
- Paczosa MK, Mecsas J. *Klebsiella pneumoniae*: going on the offense with a strong defense. *Microbiol Mol Biol Rev*. 2016;80(3):629–61.
- Bengochea JA, Pessoa JS. *Klebsiella pneumoniae* infection biology: living to counteract host defences. *FEMS Microbiol Rev*. 2018.
- Chung PY. The emerging problems of *Klebsiella pneumoniae* infections: carbapenem resistance and biofilm formation. *FEMS Microbiol Lett*. 2016;363(20).
- Lee CR, Lee JH, Park KS, Jeon JH, Kim YB, Cha CJ, Jeong BC, Lee SH. Antimicrobial resistance of Hypervirulent *Klebsiella pneumoniae*: epidemiology, Hypervirulence-associated determinants, and resistance mechanisms. *Front Cell Infect Microbiol*. 2017;7:483.
- Lee CR, Lee JH, Park KS, Kim YB, Jeong BC, Lee SH. Global dissemination of Carbapenemase-producing *Klebsiella pneumoniae*: epidemiology, genetic context, treatment options, and detection methods. *Front Microbiol*. 2016;7:895.
- Lery LM, Frangeul L, Tomas A, Passet V, Almeida AS, Bialek-Davenet S, Barbe V, Bengochea JA, Sansonetti P, Brisse S, et al. Comparative analysis of *Klebsiella pneumoniae* genomes identifies a phospholipase D family protein as a novel virulence factor. *BMC Biol*. 2014;12:41.
- Tomás A, Lery L, Regueiro V, Pérez-Gutiérrez C, Martínez V, Moranta D, Lobet E, González-Nicolau M, Insua JL, Tomas JM, et al. Functional genomic screen identifies *Klebsiella pneumoniae* factors implicated in blocking nuclear factor κ B (NF- κ B) signaling. *J Biol Chem*. 2015;290(27):16678–97.
- Liu L, Ye M, Li X, Li J, Deng Z, Yao YF, Ou HY. Identification and characterization of an antibacterial type VI secretion system in the Carbapenem-resistant strain. *Front Cell Infect Microbiol*. 2017;7:442.
- Hsieh PF, Lu YR, Lin TL, Lai LY, Wang JT. *Klebsiella pneumoniae* type VI secretion system contributes to bacterial competition, cell invasion, Type-1 fimbriae expression, and *in vivo* colonization. *J Infect Dis*. 2019; 219(4):637–47.
- Broberg CA, Palacios M, Miller VL. *Klebsiella*: a long way to go towards understanding this enigmatic jet-setter. *F1000Prime Rep*. 2014;6:64.
- Sarris PF, Zoumadakis C, Panopoulos NJ, Scoulica EV. Distribution of the putative type VI secretion system core genes in *Klebsiella* spp. *Infect Genet Evol*. 2011;11(1):157–66.
- Lawlor MS, Hsu J, Rick PD, Miller VL. Identification of *Klebsiella pneumoniae* virulence determinants using an intranasal infection model. *Mol Microbiol*. 2005;58(4):1054–73.
- Bent ZW, Poorey K, LaBauve AE, Hamblin R, Williams KP, Meagher RJ. A rapid spin column-based method to enrich pathogen transcripts from eukaryotic host cells prior to sequencing. *PLoS One*. 2016;11(12):e0168788.
- Bönemann G, Pietrosiuk A, Diemand A, Zentgraf H, Mogk A. Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. *EMBO J*. 2009;28(4):315–25.
- Kapitein N, Bönemann G, Pietrosiuk A, Seyffer F, Hausser I, Locker JK, Mogk A. ClpV recycles VipA/VipB tubules and prevents non-productive tubule formation to ensure efficient type VI protein secretion. *Mol Microbiol*. 2013; 87(5):1013–28.
- Förster A, Planamente S, Manoli E, Lossi NS, Freemont PS, Filloux A. Coevolution of the ATPase ClpV, the sheath proteins TssB and TssC, and the

- accessory protein TagJ/HslE1 distinguishes type VI secretion classes. *J Biol Chem.* 2014;289(47):33032–43.
33. Alam A, Golovliov I, Javed E, Sjöstedt A. ClpB mutants of *Francisella tularensis* subspecies holarctica and tularensis are defective for type VI secretion and intracellular replication. *Sci Rep.* 2018;8(1):11324.
 34. Brodmann M, Dreier RF, Broz P, Basler M. *Francisella* requires dynamic type VI secretion system and ClpB to deliver effectors for phagosomal escape. *Nat Commun.* 2017;8:15853.
 35. Wang T, Si M, Song Y, Zhu W, Gao F, Wang Y, Zhang L, Zhang W, Wei G, Luo ZQ, et al. Type VI secretion system transports Zn²⁺ to combat multiple stresses and host immunity. *PLoS Pathog.* 2015;11(7):e1005020.
 36. Si M, Zhao C, Burkinshaw B, Zhang B, Wei D, Wang Y, Dong TG, Shen X. Manganese scavenging and oxidative stress response mediated by type VI secretion system in. *Proc Natl Acad Sci U S A.* 2017;114(11):E2233–42.
 37. Si M, Wang Y, Zhang B, Zhao C, Kang Y, Bai H, Wei D, Zhu L, Zhang L, Dong TG, et al. The type VI secretion system engages a redox-regulated dual-functional Heme transporter for zinc acquisition. *Cell Rep.* 2017;20(4):949–59.
 38. Parveen S, Reddy M. Identification of YfiH (Pgef) as a factor contributing to the maintenance of bacterial peptidoglycan composition. *Mol Microbiol.* 2017;105(5):705–20.
 39. Shyntum DY, Venier SN, Moleleki LN, Toth I, Coutinho TA. Comparative genomics of type VI secretion systems in strains of *Pantoea ananatis* from different environments. *BMC Genomics.* 2014;15:163.
 40. Wan B, Zhang Q, Ni J, Li S, Wen D, Li J, Xiao H, He P, Ou HY, Tao J, et al. Type VI secretion system contributes to Enterohemorrhagic *Escherichia coli* virulence by secreting catalase against host reactive oxygen species (ROS). *PLoS Pathog.* 2017;13(3):e1005246.
 41. Hennequin C, Forestier C. oxyR, a LysR-type regulator involved in *Klebsiella pneumoniae* mucosal and abiotic colonization. *Infect Immun.* 2009;77(12):5449–57.
 42. Srinivasan VB, Mondal A, Venkataramaiah M, Chauhan NK, Rajamohan G. Role of oxyRKP, a novel LysR-family transcriptional regulator, in antimicrobial resistance and virulence in *Klebsiella pneumoniae*. *Microbiology.* 2013;159(Pt 7):1301–14.
 43. Brunet YR, Khodr A, Logger L, Aussel L, Mignot T, Rimsky S, Cascales E. H-NS silencing of the *Salmonella* Pathogenicity Island 6-encoded type VI secretion system limits *Salmonella enterica* Serovar typhimurium interbacterial killing. *Infect Immun.* 2015;83(7):2738–50.
 44. Fang FC, Rimsky S. New insights into transcriptional regulation by H-NS. *Curr Opin Microbiol.* 2008;11(2):113–20.
 45. Ares MA, Fernández-Vázquez JL, Rosales-Reyes R, Jarillo-Quijada MD, von Bargen K, Torres J, González-Y-Merchand JA, Alcántar-Curiel MD, De la Cruz MA. H-NS nucleoid protein controls virulence features of *Klebsiella pneumoniae* by regulating the expression of type 3 pill and the capsule polysaccharide. *Front Cell Infect Microbiol.* 2016;6:13.
 46. Ares MA, Fernández-Vázquez JL, Pacheco S, Martínez-Santos VI, Jarillo-Quijada MD, Torres J, Alcántar-Curiel MD, González-Y-Merchand JA, De la Cruz MA. Additional regulatory activities of MrkH for the transcriptional expression of the *Klebsiella pneumoniae* mrk genes: antagonist of H-NS and repressor. *PLoS One.* 2017;12(3):e0173285.
 47. Peng D, Li X, Liu P, Zhou X, Luo M, Su K, Chen S, Zhang Z, He Q, Qiu J, et al. Transcriptional regulation of galF by RcsAB affects capsular polysaccharide formation in *Klebsiella pneumoniae* NTUH-K2044. *Microbiol Res.* 2018;216:70–8.
 48. Su K, Zhou X, Luo M, Xu X, Liu P, Li X, Xue J, Chen S, Xu W, Li Y, et al. Genome-wide identification of genes regulated by RcsA, RcsB, and RcsAB phosphorelay regulators in *Klebsiella pneumoniae* NTUH-K2044. *Microb Pathog.* 2018;123:36–41.
 49. Brown MJ, Russo BC, O'Dee DM, Schmitt DM, Nau GJ. The contribution of the glycine cleavage system to the pathogenesis of *Francisella tularensis*. *Microbes Infect.* 2014;16(4):300–9.
 50. Mallik P, Pratt TS, Beach MB, Bradley MD, Undamatla J, Osuna R. Growth phase-dependent regulation and stringent control of *fts* are conserved processes in enteric bacteria and involve a single promoter (*fts P*) in *Escherichia coli*. *J Bacteriol.* 2004;186(1):122–35.
 51. Gueguen E, Durand E, Zhang XY, d'Amalric Q, Jourmet L, Cascales E. Expression of a *Yersinia pseudotuberculosis* type VI secretion system is responsive to envelope stresses through the OmpR transcriptional activator. *PLoS One.* 2013;8(6):e66615.
 52. Zhang W, Wang Y, Song Y, Wang T, Xu S, Peng Z, Lin X, Zhang L, Shen X. A type VI secretion system regulated by OmpR in *Yersinia pseudotuberculosis* functions to maintain intracellular pH homeostasis. *Environ Microbiol.* 2013;15(2):557–69.
 53. Lin TH, Chen Y, Kuo JT, Lai YC, Wu CC, Huang CF, Lin CT. Phosphorylated OmpR is required for type 3 fimbriae expression in *Klebsiella pneumoniae* under hypertonic conditions. *Front Microbiol.* 2018;9:2405.
 54. Ho BT, Dong TG, Mekalanos JJ. A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe.* 2014;15(1):9–21.
 55. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DJ, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 2015;43(Database issue):D222–6.
 56. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* 2019;47(D1):D427–32.
 57. Galperin MY, Makarova KS, Wolf YI, Koonin EV. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* 2015;43(Database issue):D261–9.
 58. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N, Lopez R. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* 2015;43(W1):W580–4.
 59. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. *Bioinformatics.* 2005;21(16):3422–3.
 60. Solovyev V, Salamov A. Automatic annotation of microbial genomes and metagenomic sequences. In: *Metagenomics and Its Applications in Agriculture, Biomedicine and Environmental Studies* Edited by Li R, vol. 1: Nova science publishers; 2011: 61–78.
 61. Münch R, Hiller K, Grote A, Scheer M, Klein J, Schobert M, Jahn D. Virtual footprint and PRODIGIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics.* 2005;21(22):4187–9.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

