

ALEXIA SUELLEN FERNANDES

**ANÁLISE DE SEQUÊNCIAS DE INSERÇÃO E TRANSPOSONS NOS GENOMAS
DE BACTÉRIAS FITOPATOGÊNICAS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Microbiologia Agrícola, para obtenção do título de *Magister Scientiae*.

Orientador: Mateus Ferreira Santana

Coorientadoras: Marisa Vieira de Queiroz
Denise Mara Soares Bazzolli

**VIÇOSA - MINAS GERAIS
2021**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

F363a
2021
Fernandes, Alexia Suellen, 1997-
Análise de sequências de inserção e transposons nos
genomas de bactérias fitopatogênicas / Alexia Suellen
Fernandes. – Viçosa, MG, 2021.
1 dissertação eletrônica (62 f.): il. (algumas color.).

Orientador: Mateus Ferreira Santana.
Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Microbiologia, 2021.
Inclui bibliografia.
DOI: <https://doi.org/10.47328/ufvbbt.2022.066>
Modo de acesso: World Wide Web.

1. Bactérias fitopatogênicas - Genética. 2. Elementos de
DNA transponíveis. 3. Adaptação (Fisiologia). I. Santana,
Mateus Ferreira, 1983-. II. Universidade Federal de Viçosa.
Departamento de Microbiologia. Programa de Pós-graduação em
Microbiologia Agrícola. III. Título.

CDD 22. ed. 632.32

ALEXIA SUELLEN FERNANDES

**ANÁLISE DE SEQUÊNCIAS DE INSERÇÃO E TRANSPOSONS NOS GENOMAS
DE BACTÉRIAS FITOPATOGÊNICAS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Microbiologia Agrícola, para obtenção do título de *Magister Scientiae*.

APROVADA: 06 de dezembro de 2021.

Assentimento:



Alexia Suellen Fernandes
Autora



Mateus Ferreira Santana
Orientador

**A Deus e aos meus eternos avós,
Maria da Conceição dos Santos e
Pedro Fernandes Gonçalves.**

AGRADECIMENTOS

A Deus que sempre esteve comigo durante o mestrado, me guiando e me dando forças em todos os momentos, principalmente naqueles em que achei que não conseguiria concluir essa jornada.

A minha família, principalmente meus pais e meus irmãos que sempre estiveram ao meu lado, nas lutas, nas conquistas e nos momentos em que parecia que tudo ia dar errado. Obrigada por me compreenderem e desculpe as vezes em que coloquei todo mundo doido, o alicerce de vocês e o apoio foi o que me fez chegar até aqui. Também dedico esse trabalho aos meus avós que apesar de não estarem mais nesse mundo, deixaram para nós um exemplo de força e amor a vida, sempre na hora das complicações a lembrança de vocês vinha fortemente e me ajudava a seguir em frente.

Agradeço também aos meus amigos, principalmente a Mariana Yoshinaga, que mesmo longe e também concluindo seu mestrado, sempre mandava uma mensagem, me dava apoio nas longas conversas pelo Google meet. Além de ser uma inspiração de dedicação! Também sou grata pelos amigos e pessoas que conheci em Viçosa, a Jéssica também conhecida como Jeje que sempre segurou meus surtos, me ajudou durante o projeto, nas disciplinas e esteve presente em praticamente todos os perrengues que passei por aqui! E a Mísia dos protocolos, companheira de estudos e de cafés, cachorros-quentes e de conversas até altas horas, sou muito grata por ter conhecido você! E claro que as Biogirls também não poderiam ficar de fora dessa lista, com vocês pude compartilhar minhas conquistas, frustrações e também rir muito, que saudade!

Ao Prof. Mateus Santana, pela oportunidade, orientação, paciência, boa vontade e disposição em sempre me ensinar, e por todas as oportunidades de realização de atividades acadêmicas. Também agradeço aos professores do departamento e também da Biologia da UFV/CAF pelos ensinamentos, postura ética e profissional, pelo apoio, incentivo e preocupação com a minha formação.

Por fim, deixo meus agradecimentos à Universidade Federal de Viçosa, a coordenação do Programa de Pós-graduação em Microbiologia Agrícola e às agências financiadoras, em especial a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por permitirem a realização desse trabalho.

RESUMO

FERNANDES, Alexia Suellen, M.Sc., Universidade Federal de Viçosa, dezembro de 2021. **Análise de sequências de inserção e transposons nos genomas de bactérias fitopatogênicas.** Orientador: Mateus Ferreira Santana. Coorientadoras: Marisa Vieira de Queiroz e Denise Mara Soares Bazzoli.

O desenvolvimento de resistência contra os principais métodos de controle a doenças causadas por fitopatógenos são um dos principais problemas enfrentados na agricultura. Em 2012, Mansfield et al., elaboraram um ranking classificando os dez principais fitopatógenos bacterianos de interesse agrônomo e científico, sendo esses, *Pseudomonas syringae* patovares, *Ralstonia solanacearum*, *Agrobacterium tumefaciens*, *Xanthomonas oryzae* pv. *oryzae*, *Xanthomonas campestris* patovares, *Xanthomonas axonopodis* pv. *manihotis*, *Erwinia amylovora*, *Xylella fastidiosa*, *Dickeya* (*dadantii* e *solani*), *Pectobacterium* (*carotovorum* e *astrosepticum*). Estes fitopatógenos ainda são os maiores responsáveis por causar perdas e danos significativos em lavouras, são de difícil controle, apresentam significativa variabilidade genética dentro da mesma espécie e muitos são capazes de adquirir genes de forma horizontal que influenciam na patogênese. No entanto, pouco se sabe sobre a influência de transposons (Tns) e sequências de inserção (ISs) na virulência e adaptação de bactérias fitopatogênicas. Dessa forma, nosso estudo teve o objetivo de identificar e tentar entender o papel de sequências de inserção e transposons nos genomas completos das dez principais bactérias fitopatogênicas de acordo com a classificação de Mansfield et al., 2012. Neste trabalho, foi realizada a busca em 270 genomas completos desses fitopatógenos para a identificação e caracterização de ISs e Tns coletivamente denominados elementos transponíveis (ETs). Foram identificados um total de 35.692 sequências de inserção, sendo as famílias IS5, IS3, IS4 e IS110 as mais abundantes. A mediana de ISs encontradas em cada espécie, foi: 66,5 elementos por genoma em *P. syringae*, 60 em *R. solanacearum*, 10,5 em *A. tumefaciens*, 383,5 em *X. oryzae*, 65 em *X. campestris*, 60 em *X. axonopodis*, 2 em *E. amylovora*, 5,5 em *X. fastidiosa*, 21 em *D. dadantii* e 6 em *D. solani*, 6 em *P. carotovorum* e 8 em *P. astrosepticum*. Além disso, 71 transposons foram identificados, onde em quatro deles observamos a presença de genes acessórios que codificam resistência a estreptomicina e genes de patogenicidade, como as proteínas efetoras do tipo III de secreção HopX e Popp2 e proteínas efetoras que atuam como ativadores de transcrição em plantas. Verificamos também a presença de recombinações ectópicas mediadas por elementos transponíveis nas espécies do gênero *Xanthomonas*. Os perfis de inserção de ISs revelaram a

proximidade desses elementos com genes de virulência e adaptação. Em *X. oryzae*, identificamos elementos ISs interrompendo genes de avirulência. As análises de pressão de seleção mostraram que ISs próximas a genes da família ATPases estão sendo selecionados positivamente em *P. syringae*, enquanto nessa mesma espécie, elementos próximos a genes que codificam bombas de efluxo sofrem um processo de eliminação dos genomas. A análise de dados de RNAseq das espécies *P. syringae*, *X. oryzae* e *R. solanacearum* mostraram mudanças significativas na expressão dos genes das transposases dos elementos transponíveis avaliados em diferentes condições de temperatura, meios de cultivo e estresse ao óxido nítrico e oxidativo. Por fim, detectamos dois sRNA putativos, sendo um cis-RNA codificado por uma IS que realiza possivelmente a regulação da transposição do próprio elemento e um trans-RNA codificado por um transposon e possivelmente que regula um gene que codifica uma proteína efetora da família TALE. Dessa maneira, nossos resultados sugerem que ETs interferem nos mecanismos de patogenicidade e adaptação desses fitopatógenos. Portanto, a compreensão abrangente do contexto desempenhado por ETs nos genomas de bactérias deve ser explorada com o intuito de identificar alvos críticos que no futuro podem auxiliar no gerenciamento de doenças e no controle de patógenos.

Palavras-chave: Elementos genéticos móveis. Fitopatógenos. Virulência. Adaptação

ABSTRACT

FERNANDES, Alexia Suellen, M.Sc., Universidade Federal de Viçosa, December, 2021. **Analysis of insertion sequence and transposons in phytopathogenic bacteria genomes.** Adviser: Mateus Ferreira Santana. Co-advisers: Marisa Vieira de Queiroz and Denise Mara Soares Bazzoli.

A few of the critical problems confronting agriculture is the development of resistance to the main strategies of controlling diseases caused by phytopathogens. Mansfield et al. (2012) developed a ranking classifying the 12 most important plant pathogens of agricultural and scientific interest, with these microorganisms being referred to as *Pseudomonas syringae* pathovars, *Ralstonia solanacearum*, *Agrobacterium tumefaciens*, *Xanthomonas oryzae* pv. *oryzae*, *Xanthomonas campestris* pathovars, *Xanthomonas axonopodis* pv. *manihotis*, *Erwinia amylovora*, *Xylella fastidiosa*, *Dickeya* (*dadantii* e *solani*), *Pectobacterium* (*carotovorum* e *astrosepticum*). These phytopathogens are also the most responsible for causing significant losses and damage in lavouras, are difficult to control, have significant genetic variability within the same species, and many are capable of horizontally acquiring genes that influence pathogenesis. However, little is known about the role of transposons (Tns) and insertion sequences (ISs) in the virulence and adaptation of phytopathogenic bacteria. In this way, our study aimed to identify and comprehend the role of insertion and transposon sequences in the whole genomes of the ten main plant pathogens, as classified by Mansfield et al., 2012. In this context, a search was conducted in 270 complete genomes of these plant pathogens for the identification and characterization of ISs and Tns, known as transposable elements (ETs). A total of 35.692 insertion sequences were found, with the IS5, IS3, IS4, and IS110 families being the most prevalent. The median number of ISs found in each species was: 66,5 elements per genome in *P.s syringae*, 60 in *R. solanacearum*, 10,5 in *A. tumefaciens*, 383,5 in *X. oryzae*, 65 in *X. campestris*, 60 in *X. axonopodis*, 2 in *E. amylovora*, 5,5 in *X. fastidiosa*, 21 in *D. dadantii* e 6 in *D. solani*, 6 in *P. carotovorum* e 8 in *P. astrosepticum*. In addition, 71 transposons were identified, with four of them including accessory genes that code for resistance to estreptomycin as well as pathogenic genes such as type III secretion system effector, HopX and Popp2, and effector proteins that act as transcription activators in plants. We also identified by the presence of ectopic recombination mediated by elements found in *Xanthomonas* species. Furthermore, ISs insertion profiles revealed the proximity of these elements to genes of virulence and adaptation. Furthermore, in *X. oryzae*, we discovered ISs elements that disrupt avirulence genes.

Selection pressure analyses revealed that ISs near to genes of the ATPase family are being selected positively in *P. syringae*, whereas elements near to ABC transporter efflux pumps are being selected negatively in the same species. Moreover, the analysis of RNAseq data from *P. syringae*, *X. oryzae*, and *R. solanacearum* revealed significant changes in the expression of genes encoding transposable elements tested under different temperature cultivation and stress conditions. Finally, we found two putative sRNAs, one cis-RNA codified by an IS and possibly regulating the frequency of transposition of the element itself, and one trans-RNA codified by a transposon and possibly regulating a TALE type effector protein. In this way, our results indicate that transposable elements (ETs) can impact the pathogenicity and adaptation mechanisms of these plant pathogens. Therefore, a comprehensive understanding of the context played by ETs in bacterial genomes should be explored with the objective of identifying critical targets that can help in disease management and pathogen control in the future.

Keywords: Self-transmissible mobile elements. Phytopathogens. Virulence. Adaptation

SUMÁRIO

INTRODUÇÃO GERAL	10
BIBLIOGRAFIA	17
CAPÍTULO 1	22
RESUMO	23
ABSTRACT	23
INTRODUÇÃO	24
MÉTODOS	26
Genomas e Transcriptomas	26
Identificação e caracterização das sequências de inserção e transposons	26
Inferência filogenética	27
Rearranjos cromossomais mediados por ISs e Tns	27
Pressão de seleção em ISs e Tns	28
Análise de Transcriptomas	29
Predição de sRNAs reguladores	29
Análise de alvos	30
RESULTADOS	31
Perfil de distribuição de sequências de inserção nas 10 principais bactérias fitopatogênicas	31
Perfil de distribuição de transposons nas 10 principais bactérias fitopatogênicas ...	36
Rearranjos cromossomais mediados por ISs e Tns	41
Análise de pressão de seleção em ISs e Tns	41
Análises dos elementos diferencialmente expressos	43
Pequenos RNAs reguladores codificados por ETs	45
DISCUSSÃO	47
CONSIDERAÇÕES FINAIS	55
REFERÊNCIAS	56

INTRODUÇÃO GERAL

Bactérias fitopatogênicas atualmente são consideradas uma das grandes ameaças à produção global de alimentos, pois causam diversas doenças nas principais culturas agrônômicas levando a grandes prejuízos econômicos. Em 2012, foi realizada uma revisão bibliográfica, que teve a participação de 458 pesquisadores, onde foi elaborado uma classificação das 10 principais bactérias fitopatogênicas de interesse científico e econômico (Mansfield et al., 2012).

De acordo com esse comitê, a bactéria que ocupou o primeiro lugar dessa classificação é *Pseudomonas syringae*. Essa espécie se caracteriza por ser uma bactéria Gram-negativa, heterotrófica e aeróbica, capaz de colonizar as folhas e outras partes aéreas de culturas importantes como feijão, soja, tomate, trigo e café. Os principais sintomas causados por *P. syringae* em plantas são manchas, galhas e morte rápida dos tecidos (Bender et al., 1999). Existem aproximadamente mais de 50 patovares identificados, sendo esses responsáveis por infectar espécies diferentes de plantas não relacionadas (Mansfield et al., 2012; Xin et al., 2018). *P. syringae* possui diversos mecanismos envolvidos na patogenicidade que garantem maior eficiência no processo de infecção como, por exemplo, a capacidade de superar a resistência do hospedeiro por meio de proteínas efetoras, produção de toxinas que degradam a parede celular e são secretadas pelo sistema de secreção do tipo III (T3SS) (Ichinose et al., 2013; Xin et al., 2018). Outra estratégia utilizada por esse organismo é a nucleação biológica de gelo. *P. syringae* possui um único gene (*inaZ*), que codifica uma proteína de membrana externa que pode catalisar a formação de cristais de gelo no tecido da planta e consequentemente causar ferimentos que permite a entrada da bactéria (Gurian-Sherman & Lindow, 1993).

Outra protagonista na lista de fitopatógenos em destaque, ocupando a segunda posição na classificação é a *Ralstonia solanacearum*, que apresenta uma ampla distribuição geográfica e a capacidade de infectar uma ampla variedade de hospedeiros, aproximadamente 200 espécies de plantas em mais de 50 famílias. Considerada um dos principais patógenos que causa a murcha de plantas, é uma bactéria Gram-negativa, altamente heterogênea, habitante natural dos solos, que infecta seus hospedeiros por entre as pontas das raízes ou por meio de rachaduras laterais, invadindo o tecido vascular, xilema e se dispersando para as partes aéreas (Morel et al., 2018). Visto a variabilidade genotípica e fenotípica da espécie, *R. solanacearum* tem sido considerada como um complexo de espécies (Fegan & Prior, 2005). A patogenicidade de *R. solanacearum* está relacionada ao sistema de secreção do tipo III (T3SS), ao gene *hrp*, a produção exopolissacarídeo (EPS), as enzimas de degradação da parede celular secretadas pelo

sistema secreção tipo II (T2SS), aos apêndices de motilidade e ao sistema de secreção do tipo IV (Meng, 2013; Peeters et al., 2013).

Agrobacterium tumerfacies é uma bactéria Gram-negativa, encontrada em solos e causa a doença galha da coroa, principalmente em dicotiledôneas, o que resulta na formação de tumores entre o caule e a raiz (Smith & Townsend, 1907). Essa espécie possui apenas uma estratégia de virulência que depende da presença do plasmídeo Ti (Tumor-inducing), responsável pela transferência natural de uma região pertencente ao plasmídeo denominada de T-DNA, que induz um desequilíbrio na biossíntese de hormônios da própria planta com consequente hiper proliferação celular (Chilton et al., 1977). Esse plasmídeo possui também outra região importante denominada região de virulência (*vir*) que não é transferida para a célula hospedeira, mas codifica sete proteínas que auxiliam nos processos de transporte e integração do T-DNA no genoma do hospedeiro (Zambryski, 1992). Devido a sua capacidade única de virulência por meio da transformação genética, *A. tumerfaciens* não é apenas mais um fitopatógeno, mas também uma ferramenta para transformação de plantas (Krenek et al., 2015).

Xanthomonas oryzae pv. *oryzae* (*Xoo*) é uma bactéria Gram-negativa que causa principalmente a ferrugem do crestamento bacteriano da folha em culturas do arroz, denominado crestamento bacteriano do arroz. Em condições favoráveis para a doença, os prejuízos podem chegar a perdas de até 70% da produção. *X. oryzae* pv. *oryzae* geralmente infecta o hospedeiro mediante a feridas ou estruturas de secreção como hidatódio. A bactéria se instala no xilema e causa obstrução do tecido vascular vegetal levando assim a murcha da planta (Niño-Liu et al., 2006). Ademais, possui uma grande variedade de fatores de virulência, tal como sistema de secreção do tipo II (T2SS), sistema de secreção tipo III (T3SS) e a produção de exopolissacarídeos (EPS) (Jhan et al., 2007; Xue et al., 2018). Além disso, *Xoo* apresenta uma rápida capacidade de evolução devido a um alto número de ISs e de genes efetores *TAL* (Ryan et al., 2011).

Xanthomonas campestris é responsável por causar a podridão negra, essa doença desencadeia a produção insuficiente de clorofila e o aparecimento de manchas negras sob os tecidos principalmente de plantas do gênero *Brassica* (Alvarez, 2000). *X. campestris* apresenta múltiplos patovares que possuem um alto grau de especificidade com a planta hospedeira como, *X. campestris* pv. *campestris*, agente causal da podridão negra em repolho, couve e couve-flor e *X. campestris* pv. *raphani* causador de manchas foliares em rabanetes e nabo (Mansfield et al., 2012). Esse fitopatógeno utiliza diversos mecanismos para infectar seus hospedeiros como, por exemplo, sistemas de secreção, proteínas efetoras do tipo *TAL* e produção de exopolissacarídeos (EPS). Os estudos dessas bactérias geraram grande impacto científico

devido a dois fatores, entendimento de um mecanismo de sinalização celular ativado para modular a síntese de fatores de virulência quando patógenos vegetais se encontram em condições de estresse e pela produção de goma xantana, um polissacarídeo utilizado nas indústrias alimentícia, farmacêutica e petrolífera (Ryan et al., 2011).

Outra bactéria do gênero *Xantomonas* foi escolhida para estar no sexto lugar desse rank, *X. axonopodis*. Essa bactéria também possui uma grande variedade de patovares, mas o agente causador da bacteriose ou murcha bacteriana da mandioca (CBB, do inglês: *Cassava bacterial blight*), *X. axonopodis* pv. *manihotis* (*Xam*), se destaca devido a importância do seu hospedeiro, já que a mandioca é a terceira maior fonte de carboidratos e um dos principais alimentos básicos no mundo (FAOSTAT, 2012). Apesar de existirem poucos estudos sobre esse patovar, sabe-se que *Xam* produz diversos sintomas como lesões angulares nas folhas, ferrugem, murcha e manchas foliares além de colonizar o tecido mesofílico e vascular. Essa bactéria conta com o sistema de secreção tipo III como fator de patogenicidade e possui uma alta variabilidade genética em diferentes regiões geográficas onde estão presentes (Restrepo & Verdier, 1997; Ryan et al., 2011).

O fogo bacteriano, (*fireblight*) é uma bacteriose causada por *E. amylovora* e acomete árvores frutíferas como pomares de maçãs e peras, sendo altamente destrutiva na fase jovem da planta (Vanneste, 2000). Qualificada como bactéria fitopatogênica *premier*, foi o primeiro organismo demonstrado a causar doenças em plantas. *E. amylovora* invade seu hospedeiro através de aberturas nos tecidos vegetativos, estômatos e estigmas florais, causando murcha das flores, folhas, brotos e lesões nos frutos, deixando um aspecto de queima (Peil et al., 2009). Diferente da maioria dos demais fitopatógenos, essa bactéria não produz enzimas de degradação de parede celular e toxinas de baixo peso molecular. Sua patogênese provém da ilha de patogenicidade *hrp* com cerca de 60 genes. Essa ilha apresenta sinais de ter sido adquirida por transferência horizontal (Oh e Beer, 2005). Alguns isolados são resistentes à estreptomicina, um antibiótico tradicionalmente usado em seu controle. A resistência se deve ao par de genes *strA/strB* localizado no transposon Tn5393 e no plasmídeo não conjugativo pEA29 (Tancos et al., 2016).

X. fastidiosa foi o primeiro fitopatógeno a ter o genoma completamente sequenciado, é causadora da clorose variegada em citros (CVC) em laranjeiras e da doença de Pierce em videiras (PD). Os sintomas mais comuns causados pela doença são a clorose e necrose das folhas, e o tamanho reduzido dos frutos (Chatterjee et al., 2008; Kruse et al., 2019). Essa bactéria Gram-negativa não possui genes de virulência e nem dispõem de um sistema de secreção do tipo III, apesar de possuir os sistemas de secreção do tipo I e II, que podem estar

associados a bomba de efluxo e a secreção de enzimas hidrolíticas. Dessa forma, *X. fastidiosa* é dependente de insetos vetores responsáveis por sua transmissão e introdução no tecido vascular da planta. A bactéria coloniza os vasos do xilema dificultando a disponibilidade de água e nutrientes para o hospedeiro (Simpson et al., 2000). A produção de biofilme tanto no vetor quanto por toda extensão do vegetal auxilia na sobrevivência do patógeno em ambientes com alta turbulência, pressão diferencial e baixa disponibilidade de nutrientes (Van et al., 2002).

O gênero *Dickeya* também está presente na classificação realizada por Mansfield et al. (2012) em nono lugar e é representado por duas espécies, *D. dadantii* e *D. solani*. Apesar desse gênero ter mais de seis espécies, essas duas foram selecionadas pelo fato de *D. dadantii* ter importante participação em estudos moleculares que buscam compreender o comportamento das bacterioses (Li et al., 2009; Yamazaki et al., 2011) e *D. solani* devido a eficiência de propagação, causando doenças até mesmo em baixa densidade populacional. Pouco se sabe sobre a biologia da *D. solani*, por isso a especificidade de hospedeiros não é ainda compreendida. *D. solani* possui uma extensa lista de alvos, desde espécies de hortaliças importantes como batata, cenoura, cebola a espécies ornamentais como orquídeas, cravos e tulipas (Toth et al., 2011). As espécies do gênero causam podridão mole e estão comumente presentes nos solos, sendo transmitidas pela água e vetores como os insetos. As bactérias colonizam o xilema das plantas e demonstram um aumento de patogenicidade em temperaturas elevadas, o que tem se tornado um fato preocupante devido as mudanças climáticas que podem elevar a temperatura da Terra e conseqüentemente permitir à expansão desse fitopatógeno no futuro (Sławiak et al., 2009).

Por fim, na última posição se encontra duas espécies do gênero *Pectobacterium*, *P. carotovorum* e *P. astrosepticum*. Essas espécies são causadoras de podridão mole devido a produção de enzimas pectolíticas que rompem as pectinas das células vegetais promovendo o apodrecimento dos hospedeiros (Toth & Birch, 2005). Devido à proximidade com *E. coli*, foi possível desenvolver ferramentas genéticas que contribuíram para o estudo de enzimas degradadoras de parede de plantas e a descoberta dos sistemas de secreção do tipo I e tipo II (Diolez & Coleno, 1985; Salmond, 1994) e o primeiro relato da produção do ácido 1-carbapen-2-em-3-carboxílico, membro da classe carbapenem de antibióticos β -lactâmicos (Whitehead et al., 2002; Coulthurst et al., 2005; Barnard et al., 2007). Em *P. carotovorum* foram identificados genes de virulência e de adaptação ecológica adquiridos por transferência horizontal de *P. astrosepticum*, onde também foi identificada a presença de ilhas genômicas igualmente presentes no gênero *Dickeya* (Ma et al., 2007; Glasner et al., 2008).

Os 10 fitopatógenos selecionados, além de possuírem uma importância tanto econômica como científica, possuem ainda em comum o fato de apresentarem alta plasticidade genômica (Mansfield et al. 2012). Sendo elementos genéticos móveis, como as ISs e Tns, uma das principais fontes de variabilidade e dessa plasticidade, por isso estudo da importância desses elementos em bactérias fitopatogênicas se faz necessário.

Os elementos genéticos móveis (MGEs) são segmentos de DNA que codificam proteínas que proporcionam o movimento intercelular ou intracelular dessas sequências, sendo capaz de resultar na transferência de informações genéticas e na integração em um novo sítio (Frost et al., 2005). Os MGEs podem consistir ISs, Tns, elementos integrativos e conjugativos (ICEs), fagos, plasmídeos, ilhas de patogenicidade e integrons. Os métodos de transferência podem ocorrer de dois modos, transferência horizontal, que acontece através dos métodos de conjugação, transdução e transformação, e a transferência vertical, na qual o indivíduo recebe o material genético do seu genitor (Malachowa & DeLeo, 2010). Os MGEs são importantes agentes na evolução de genomas de eucariotos e procariotos (Nagy & Chandler, 2004). Dentre os MGEs, os elementos transponíveis (ETs) são os maiores contribuidores da fração repetitiva dos genomas.

Os ETs foram descobertos em 1940 por Barbara MacClintock, em seus experimentos com milho. MacClintock percebeu que a coloração dos grãos de milhos era afetada por “elementos controladores” capazes de mudar suas posições no genoma. Essa descoberta foi encarada com grande ceticismo na época visto que o DNA não era mais considerado estático, sendo sujeito a alterações e rearranjos (Ravindran, 2012). Por muito tempo os ETs foram ignorados quanto à importância e ubiquidade nos genomas, principalmente devido à dificuldade de identificação e caracterização desses elementos em genomas sequenciados utilizando a tecnologia de sequenciamento de nova geração (NGS). Essas tecnologias geram *reads* muito pequenos, originadas de sequências repetidas, como os ETs, criam ambiguidades no alinhamento e na montagem dos genomas, produzindo erros de posicionamento, lacunas e consequentemente a fragmentação do genoma sequenciado. Com as novas tecnologias de sequenciamento (sequenciamento de terceira geração), os genomas completos ou quase completos têm sido depositados em bancos de dados o que tem permitido uma melhor análise dos ETs (Hoban et al., 2016). Hoje, ETs são considerados ubíquos nos genomas e importantes como mecanismo de evolução. Em bactérias os principais ETs são as ISs e Tns.

As ISs são consideradas como ETs mais simples e os mais frequentemente encontrados em bactérias, sendo estes elementos classificados em mais de 29 famílias. Elas se caracterizam por serem pequenas, menores que 3kb, possuírem transposição autônoma, uma, duas ou três

ORFs (*Open Read Frame*) que codificam proteínas relacionadas à mobilidade e sequências terminais curtas repetitivas e invertidas (TIRs) nas extremidades (Siguier et al., 2014). Os tipos de famílias das ISs são definidos de acordo com os quatro tipos de transposases descritas, e essas proteínas possuem diferentes mecanismos catalíticos para mediar várias quebras e junções de DNA necessárias para transposição para um novo alvo (Curcio & Derbyshire, 2003). As transposases DDE são as mais abundantes dentre as bactérias, essa contém três aminoácidos polares conservados Asp, Asp e Glu. As etapas para a transposição se iniciam com a quebra da ligação fosfodiéster nas extremidades 3' OH da pentose, em seguida acontece a excisão do elemento, sendo que uma das suas extremidades 3' OH é protegida pela transposase até o sítio alvo onde ocorre a reação de transesterificação (Nesmelova & Hackett, 2010). As transposases DEDD que são menos frequentes na natureza, limitada a família IS110, é composta por quatro aminoácidos Asp, Glu, Asp, Asp e não geram repetições diretas (DRs) de alvo, pois, essa transposase está relacionada funcionalmente com a enzima resolvase (Carareto et al., 2015). Os transposons que se movem via HUH (H= histidina e U= resíduo hidrofóbico) não apresentam TIRs, mas uma estrutura secundária formada pela molécula de DNA, que é utilizada como sítio de reconhecimento para o início da transposição, a transposase-HUH apresenta um par conservado de resíduos de histidina (His) e utilizam uma tirosina para quebrar a ligação fosfodiéster e liberar o terminal 5' OH, nesse caso o transposon é excisado de seu local original para formar um intermediário circular de DNA. Esse mecanismo de excisão também ocorre em transposases que codificam uma serina, porém nesse caso não há a formação de *hairpin* e as fitas são cortadas nas duas direções, ocorrendo a remoção total do elemento transponível formando intermediários 5' fosfoserina e extremidades 3' OH livres que sofrem um ataque nucleofílico resultando na circularização, que posteriormente será inserido no sítio alvo (Curcio & Derbyshire, 2003; Carareto et al., 2015). Comum a todos, após o final do processo ocorre o reparo dos gaps nas fitas de DNA resultantes da quebra gerada por meio da excisão dos elementos.

Os Tns são sequências maiores que as ISs, geralmente mais de 4 Kb, e se caracterizam por possuírem um gene que codifica a enzima transposase e suas extremidades contém sequências de repetição que podem ser diretas ou invertidas, a excisão e a integração ocorrem diretamente a partir do DNA (Skipper et al., 2013). Existem também os Tns chamados de complexos, pois além de codificar proteínas mobilizadoras, podem conter um ou mais genes que são transcritos em outros tipos de proteínas, como por exemplo, resistência a antibióticos e fatores de virulência. Por fim, existem os Tns compostos que são caracterizados pela presença de duas ISs flanqueando genes não relacionados à transposição (Babakhani & Oloomi, 2018).

Os transposons podem ser classificados de duas maneiras, por meio da autonomia ou não de transposição: são denominados autônomos aqueles que codificam proteínas de transposição, e não autônomos, os que não são capazes de codificar essas proteínas e logo precisam usar as proteínas disponíveis de Tns autônomos. Já a segunda classificação se refere ao processo de transferência que pode ser replicativa, onde uma cópia do Tn é transferida para o sítio alvo e a segunda, denominada conservativa, onde o elemento transponível deixa o sítio original e vai ocupar um novo sítio (Babakhani & Oloomi, 2018).

A presença de ISs e Tns podem ter várias consequências para o genoma hospedeiro tanto na estrutura quanto na função. Um efeito lógico é a interrupção de genes devido a inserção em locais que codificam proteínas. Outro efeito importante é a interferência na expressão de genes adjacentes pela deleção ou interrupção de sequências reguladoras, impedindo a síntese de uma proteína, ou mesmo alterando o padrão de expressão gênica (Siguiet et al., 2014; Vandecraen et al., 2017). As ISs e os Tns também causam rearranjos no DNA que influenciam a expressão gênica e a estrutura cromossômica do hospedeiro. Esses elementos podem também influenciar na adaptabilidade dos organismos, uma vez que podem carregar genes que codificam fatores de virulência, resistência à metal pesado, metabolismo de carbono, resistência à antimicrobianos, dentre outros (Siguiet et al., 2014). Por exemplo, alguns defensivos agrícolas possuem em sua formulação a estreptomicina (Rezzonico et al., 2009). Isso tem levado a uma ampla resistência à estreptomicina em bactérias fitopatogênicas como em *P. syringae* e *X. campestris*. Essa resistência ocorre devido a presença de um par de genes de resistência *strAB* e *smR* presente no transposon Tn5393, isolado pela primeira vez em *E. amylovora* (Sundin & Wang, 2018).

Alguns ETs também são capazes de codificar pequenos RNAs reguladores não codificantes (sRNA) como o sRNA TnpA, codificado pelo transposon IS200. Esse desempenha um papel de regulação independente do controle da transposição afetando a expressão de mais de 70 genes em *Salmonella typhimurium* (Ellis et al., 2017). O sRNA TnpA é considerado um RNA do tipo trans uma vez que são transcritos a partir de genes localizados em regiões afastadas dos seus alvos, possuindo complementariedade imperfeita com o mRNA alvo, e assim, necessitam de ser auxiliados por uma proteína chaperona de RNA denominada de Hfq a qual facilita a interação sRNA – mRNA (Ellis et al., 2017). A maioria dos sRNAs relatados na literatura que são codificados por ETs são do tipo cis-sRNA, ou seja, são sintetizados no mesmo locus do mRNA alvo (fita codificadora complementar ao do alvo), por isso, possuem uma região de interação de perfeita complementaridade com o seu alvo (Bobrovskyy & Vanderpool, 2013; Sauer, 2013). Nesse caso, os sRNAs anti-senso produzidos por ETs atuam na regulação

da transposição, controlando o número de cópias dos elementos no genoma (Kreth & Merritt, 2019). Existe ainda os sRNAs denominados *Dual-function RNAs* que atuam tanto em interações sRNA-mRNA como também são capazes de codificar peptídeos. Esses sRNA, apesar de já terem sido descritos em uma bactéria fitopatogênica, como em *E. amylovora* (Vanderpool & Bobrovskyy, 2013), ainda não foram relatados em ETs. Embora exista um número relevante de sRNAs identificados em procariotos, em bactérias fitopatogênicas os trabalhos são escassos.

Por fim, existem evidências de que ETs podem ser “exaptados” em nível molecular para novas funções, ou seja, podem ser domesticados. Neste contexto, há casos em que regiões codificadoras de ETs são modificadas ao longo do tempo de forma que perdem a função relacionada à transposição e passam a contribuir com funções relevantes para o hospedeiro (Sinzelle & Ivics, 2009; Siguier et al., 2014). Um exemplo é o sistema imune adaptativo CRISPR-Cas. O complexo proteico Cas codifica algumas enzimas que clivam o DNA invasor e em seguida o integra no *locus* CRISPR, o objetivo dessa ação é prevenir o organismo de futuras invasões pelo mesmo agente (Jinek et al., 2012). As proteínas Cas9 possuem similaridades significativas com as proteínas de ligação e endonucleases codificadas por um grupo de Tns denominado TnpB, além de compartilhar sequências terminais semelhantes, sugerindo que os sistemas CRISPR-Cas originaram-se da domesticação de Tns (Kapitonov et al., 2016; Shmakov et al., 2017).

Atualmente, estudar e entender os elementos genéticos móveis é de grande importância. Apesar disso, em bactérias fitopatogênicas, a maioria dos trabalhos com ISs e Tns está relacionado à caracterização de um único elemento que foi descoberto devido ao seu envolvimento com algum fenótipo específico como o aumento ou diminuição de virulência. Assim, informações acerca de ISs e Tns em bactérias fitopatogênicas permanece uma lacuna a ser preenchida. Com isso o objetivo dessa dissertação foi caracterizar *in silico* os Tns e as ISs presentes nos genomas completos das 10 principais espécies bacterianas causadoras de doenças em plantas no mundo e fornecer subsídios que capacitam ampliar o entendimento do seu papel quanto aos mecanismos de adaptação e patogenicidade desses fitopatógenos.

BIBLIOGRAFIA

BABAKHANI, Sajad; OLOOMI, Mana. Transposons: the agents of antibiotic resistance in bacteria. **Journal of basic microbiology**, v. 58, n. 11, p. 905-917, 2018.

- BARNARD, Anne ML et al. Quorum sensing, virulence and secondary metabolite production in plant soft-rotting bacteria. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 362, n. 1483, p. 1165-1183, 2007.
- BENDER, Carol L.; ALARCÓN-CHAIDEZ, Francisco; GROSS, Dennis C. Pseudomonas syringae phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. **Microbiology and molecular biology reviews**, v. 63, n. 2, p. 266-292, 1999.
- BOBROVSKYY, Maksym; VANDERPOOL, Carin K. Regulation of bacterial metabolism by small RNAs using diverse mechanisms. **Annual review of genetics**, v. 47, p. 209-232, 2013.
- CARARETO, Claudia Marcia Aparecida; MONTEIRO-VITORELLO, Claudia Barros; VAN SLUYS, Marie-Anne (Ed.). Elementos de transposição: diversidade, evolução, aplicações e impacto nos genomas dos seres vivos. **SciELO-Editora FIOCRUZ**, 2015.
- CHATTERJEE, Subhadeep; ALMEIDA, Rodrigo P. P.; LINDOW, Steven. Living in two worlds: the plant and insect lifestyles of Xylella fastidiosa. **Annu. Rev. Phytopathol.**, v. 46, p. 243-271, 2008.
- COULTHURST, Sarah J.; BARNARD, Anne ML; SALMOND, George PC. Regulation and biosynthesis of carbapenem antibiotics in bacteria. **Nature Reviews Microbiology**, v. 3, n. 4, p. 295-306, 2005.
- CHILTON, Mary-Dell et al. Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. **Cell**, v. 11, n. 2, p. 263-271, 1977.
- CURCIO, M. Joan; DERBYSHIRE, Keith M. The outs and ins of transposition: from mu to kangaroo. **Nature Reviews Molecular Cell Biology**, v. 4, n. 11, p. 865-877, 2003.
- DIOLEZ, A.; COLENO, ALAIN. Mu-lac insertion-directed mutagenesis in a pectate lyase gene of Erwinia chrysanthemi. **Journal of bacteriology**, v. 163, n. 3, p. 913-917, 1985.
- ELLIS, Michael J. et al. A transposon-derived small RNA regulates gene expression in Salmonella Typhimurium. **Nucleic acids research**, v. 45, n. 9, p. 5470-5486, 2017.
- FEGAN, Mark et al. How complex is the Ralstonia solanacearum species complex. **Bacterial wilt disease and the Ralstonia solanacearum species complex**, v. 1, p. 449-461, 2005.
- FAOSTAT (2012) Food and agriculture organization of the United Nations. Available at: <http://faostat3.fao.org/>. Accessed 29 May 2020.
- FROST, Laura S. et al. Mobile genetic elements: the agents of open source evolution. **Nature Reviews Microbiology**, v. 3, n. 9, p. 722-732, 2005.
- GLASNER, J. D. et al. Niche-specificity and the variable fraction of the Pectobacterium pan-genome. **Molecular plant-microbe interactions**, v. 21, n. 12, p. 1549-1560, 2008.
- GONÇALVES, Osiel Silva et al. Transposable elements contribute to the genome plasticity of Ralstonia solanacearum species complex. **Microbial genomics**, v. 6, n. 5, 2020.
- GURIAN-SHERMAN, Douglas; LINDOW, Steven E. Bacterial ice nucleation: significance and molecular basis. **The FASEB journal**, v. 7, n. 14, p. 1338-1343, 1993.

- HOBAN, Sean et al. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. **The American Naturalist**, v. 188, n. 4, p. 379-397, 2016.
- ICHINOSE, Yuki; TAGUCHI, Fumiko; MUKAIHARA, Takafumi. Pathogenicity and virulence factors of *Pseudomonas syringae*. **Journal of general plant pathology**, v. 79, n. 5, p. 285-296, 2013.
- JINEK, Martin et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. **science**, v. 337, n. 6096, p. 816-821, 2012.
- JHA, Gopaljee; RAJESHWARI, Ramanan; SONTI, Ramesh V. Functional interplay between two *Xanthomonas oryzae* pv. *oryzae* secretion systems in modulating virulence on rice. **Molecular Plant-Microbe Interactions**, v. 20, n. 1, p. 31-40, 2007.
- KAPITONOV, Vladimir V.; MAKAROVA, Kira S.; KOONIN, Eugene V. ISC, a novel group of bacterial and archaeal DNA transposons that encode Cas9 homologs. **Journal of bacteriology**, v. 198, n. 5, p. 797-807, 2015.
- KRETH, J.; BRANTL, S.; MERRITT, Justin. Regulatory RNAs. In: **Encyclopedia of Microbiology**. Elsevier, 2019. p. 62-84.
- KRUSE, Angela; FLEITES, Laura A.; HECK, Michelle. Lessons from one fastidious bacterium to another: What can we learn about *Liberibacter* species from *Xylella fastidiosa*. **Insects**, v. 10, n. 9, p. 300, 2019.
- MALACHOWA, Natalia; DELEO, Frank R. Mobile genetic elements of *Staphylococcus aureus*. **Cellular and molecular life sciences**, v. 67, n. 18, p. 3057-3071, 2010.
- MANSFIELD, John et al. Top 10 plant pathogenic bacteria in molecular plant pathology. **Molecular plant pathology**, v. 13, n. 6, p. 614-629, 2012.
- MARTINS, Paula MM et al. Persistence in phytopathogenic bacteria: do we know enough?. **Frontiers in microbiology**, v. 9, p. 1099, 2018.
- MOREL, Arry et al. Plant pathogenicity phenotyping of *Ralstonia solanacearum* strains. In: **Host-Pathogen Interactions**. Humana Press, New York, NY, 2018. p. 223-239.
- NAGY, Zita; CHANDLER, Michael. Regulation of transposition in bacteria. **Research in microbiology**, v. 155, n. 5, p. 387-398, 2004.
- NESMELOVA, Irina V.; HACKETT, Perry B. DDE transposases: Structural similarity and diversity. **Advanced drug delivery reviews**, v. 62, n. 12, p. 1187-1195, 2010.
- NIÑO-LIU, DAVID O.; RONALD, Pamela C.; BOGDANOVE, Adam J. *Xanthomonas oryzae* pathovars: model pathogens of a model crop. **Molecular plant pathology**, v. 7, n. 5, p. 303-324, 2006.
- OH, Chang-Sik; BEER, Steven V. Molecular genetics of *Erwinia amylovora* involved in the development of fire blight. **FEMS Microbiology Letters**, v. 253, n. 2, p. 185-192, 2005.
- OLIVEIRA, Amanda CP et al. Transposons and pathogenicity in *Xanthomonas*: acquisition of murein lytic transglycosylases by TnXax1 enhances *Xanthomonas citri* subsp. *citri* 306 virulence and fitness. **PeerJ**, v. 6, p. e6111, 2018.

- PARTRIDGE, Sally R. et al. Mobile genetic elements associated with antimicrobial resistance. **Clinical microbiology reviews**, v. 31, n. 4, p. e00088-17, 2018.
- PEETERS, Nemo et al. *Ralstonia solanacearum*, a widespread bacterial plant pathogen in the post-genomic era. **Molecular plant pathology**, v. 14, n. 7, p. 651-662, 2013.
- PEIL, Andreas et al. Improvement of fire blight resistance in apple and pear. **Int J Plant Breed**, v. 3, n. 1, p. 1-27, 2009.
- RAVINDRAN, Sandeep. Barbara McClintock and the discovery of jumping genes. **Proceedings of the National Academy of Sciences**, v. 109, n. 50, p. 20198-20199, 2012.
- RESTREPO, Silvia; VERDIER, Valérie. Geographical differentiation of the population of *Xanthomonas axonopodis* pv. *manihotis* in Colombia. **Applied and Environmental Microbiology**, v. 63, n. 11, p. 4427-4434, 1997.
- REZZONICO, Fabio; STOCKWELL, Virginia O.; DUFFY, Brion. Plant agricultural streptomycin formulations do not carry antibiotic resistance genes. **Antimicrobial agents and chemotherapy**, v. 53, n. 7, p. 3173-3177, 2009.
- RYAN, Robert P. et al. Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions. **Nature Reviews Microbiology**, v. 9, n. 5, p. 344-355, 2011.
- SALMOND, George PC. Secretion of extracellular virulence factors by plant pathogenic bacteria. **Annual review of phytopathology**, v. 32, n. 1, p. 181-200, 1994.
- SAUER, Evelyn. Structure and RNA-binding properties of the bacterial LSm protein Hfq. **RNA biology**, v. 10, n. 4, p. 610-618, 2013.
- SERRATO-CAPUCHINA, Antonio; MATUTE, Daniel R. The role of transposable elements in speciation. **Genes**, v. 9, n. 5, p. 254, 2018.
- SIGUIER, Patricia; GOURBEYRE, Edith; CHANDLER, Mick. Bacterial insertion sequences: their genomic impact and diversity. **FEMS microbiology reviews**, v. 38, n. 5, p. 865-891, 2014.
- SIGUIER, Patricia et al. Everyman's guide to bacterial insertion sequences. **Mobile DNA III**, p. 555-590, 2015.
- SIMPSON, Andrew John George et al. The genome sequence of the plant pathogen *Xylella fastidiosa*. **Nature**, v. 406, n. 6792, p. 151-157, 2000.
- SINZELLE, L.; IZSVAK, Z.; IVICS, Z. Molecular domestication of transposable elements: from detrimental parasites to useful host genes. **Cellular and molecular life sciences**, v. 66, n. 6, p. 1073-1093, 2009.
- SKIPPER, Kristian Alsbjerg et al. DNA transposon-based gene vehicles-scenes from an evolutionary drive. **Journal of biomedical science**, v. 20, n. 1, p. 1-23, 2013.
- SŁAWIAK, Monika; ŁOJKOWSKA, Ewa; VAN DER WOLF, J. M. First report of bacterial soft rot on potato caused by *Dickeya* sp.(syn. *Erwinia chrysanthemi*) in Poland. **Plant Pathology**, v. 58, n. 4, p. 794-794, 2009.

- SMITH, Erwin F.; TOWNSEND, Charles O. A plant-tumor of bacterial origin. **Science**, v. 25, n. 643, p. 671-673, 1907.
- STRANGE, Richard N.; SCOTT, Peter R. Plant disease: a threat to global food security. **Annu. Rev. Phytopathol.**, v. 43, p. 83-116, 2005.
- SUNDIN, George W.; WANG, Nian. Antibiotic resistance in plant-pathogenic bacteria. **Annual Review of phytopathology**, v. 56, p. 161-180, 2018.
- TANCOS, K. A. et al. Prevalence of streptomycin-resistant *Erwinia amylovora* in New York apple orchards. **Plant Disease**, v. 100, n. 4, p. 802-809, 2016.
- TOTH, Ian K.; BIRCH, Paul R.J. Rotting softly and stealthily. **Current opinion in plant biology**, v. 8, n. 4, p. 424-429, 2005.
- TOTH, I. K. et al. *Dickeya* species: an emerging problem for potato production in Europe. **Plant Pathology**, v. 60, n. 3, p. 385-399, 2011.
- VANNESTE, Joël L. (Ed.). Fire blight: the disease and its causative agent, *Erwinia amylovora*. **CABI**, 2000.
- VAN SLUYS, M. A. et al. Comparative genomic analysis of plant-associated bacteria. **Annual review of phytopathology**, v. 40, n. 1, p. 169-189, 2002.
- VANDECRAEN, Joachim et al. The impact of insertion sequences on bacterial genome plasticity and adaptability. **Critical reviews in microbiology**, v. 43, n. 6, p. 709-730, 2017.
- VANDERPOOL, Carin K.; GOTTESMAN, Susan. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. **Molecular microbiology**, v. 54, n. 4, p. 1076-1089, 2004.
- WHITEHEAD, Neil A. et al. The regulation of virulence in phytopathogenic *Erwinia* species: quorum sensing, antibiotics and ecological considerations. **Antonie Van Leeuwenhoek**, v. 81, n. 1, p. 223-231, 2002.
- XIN, Xiu-Fang; KVITKO, Brian; HE, Sheng Yang. *Pseudomonas syringae*: what it takes to be a pathogen. **Nature Reviews Microbiology**, v. 16, n. 5, p. 316-328, 2018.
- XUE, Dingrong et al. Phosphodiesterase EdpX1 promotes *Xanthomonas oryzae* pv. *oryzae* virulence, exopolysaccharide production, and biofilm formation. **Applied and environmental microbiology**, v. 84, n. 22, p. e01717-18, 2018.
- YAMAZAKI, Akihiro et al. Commensal effect of pectate lyases secreted from *Dickeya dadantii* on proliferation of *Escherichia coli* O157: H7 EDL933 on lettuce leaves. **Applied and environmental microbiology**, v. 77, n. 1, p. 156-162, 2011.
- ZAMBRYSKI, Patricia C. Chronicles from the *Agrobacterium*-plant cell DNA transfer story. **Annual review of plant biology**, v. 43, n. 1, p. 465-490, 1992.

CAPÍTULO 1

Desvendando o papel dos elementos transponíveis nos genomas das bactérias fitopatogênicas mais importantes do mundo

RESUMO:

A variabilidade genética em fitopatógenos é um dos principais problemas encontrados para o efetivo controle das doenças. Grande parte dessa variabilidade pode estar relacionada a presença de elementos transponíveis (ETs). Aqui é fornecida a mais abrangente identificação, caracterização e análise dos possíveis impactos das sequências de inserção (ISs) e transposons (Tns) nos genomas das 10 principais bactérias fitopatogênicas de maior interesse agrônomo e científico. Um total de 35.692 ISs e 71 transposons foram identificados. A distribuição dos elementos entre as espécies foi considerada bastante heterogênea, mas aparentemente dependente do grau de especialização patógeno-hospedeiro. Alguns transposons foram identificados carregando genes de virulência, como os genes que codificam proteínas efetoras do sistema de secreção do tipo III de secreção e genes de resistência ao antimicrobiano, estreptomicina. Evidências de recombinação ectópica mediadas por ISs foram identificadas em *Xanthomonas*. No mais, verificamos que elementos ISs tendem a se inserir em regiões próximas a genes de virulência e adaptação. Em *X. oryzae*, por exemplo, detectamos ISs interrompendo genes de avirulência. Na espécie, *P. syringae* foram detectados eventos de pressão de seleção positiva para elementos próximos à genes relacionados a família ATPases. No entanto, observamos também um processo de eliminação de ISs adjacentes a genes de bombas de efluxo ABC. Associados a estes achados, a partir de análise de transcriptomas em diferentes condições de estresse verificamos diferenças na expressão de genes que codificam transposases nos isolados das espécies *R. solanacearum*, *X. oryzae* e *P. syringae*. Constatamos também um possível papel dos transposons na regulação via pequenos RNAs não codificantes reguladores. Portanto, nossos resultados apontam que elementos transponíveis possuem um papel na variabilidade e virulência em bactérias fitopatogênicas. Desse modo, acreditamos que futuramente ETs podem ser um alvo potencial para o controle de doenças em culturas comerciais.

ABSTRACT

Genetic variability in phytopathogens is one of the main problems found for effective disease control. A significant proportion of this variability may be indicative of the presence of transmissible elements (ETs) in the genome. We provide the most comprehensive identification, characterization, and analysis of the potential impacts of insertion sequences (ISs) and transposons (Tns) on the genomes of the 10 most important phytopathogenic bacteria

of agricultural and scientific interest. A total of 35.692 ISs were discovered, and 71 transposons identified. The distribution of the elements among the species was considered quite heterogeneous, but apparently it depends on the degree of pathogen-host specialization. Some transposons have been identified as carrying virulence genes such as type III secretion system effector and antimicrobial resistance genes, streptomycin. Ectopic recombination events were mediated by ISs from the IS5 and IS3 families in *Xanthomonas* species. Furthermore, we found that ISs generally prefer to insert themselves in regions close to virulence and adaptation genes, for example, in *X. oryzae*, where we found inserts that disrupted avirulence genes. Positive pressure events were detected in *P. syringae* for elements adjacent to ATPase family members, however we also observed a process of elimination of ISs adjacent to ABC transporter efflux pumps. In addition to these findings, we discovered differences in the expression of genes that code for transposases in isolates of the *R. solanacearum*, *X. oryzae*, and *P. syringae* based on transcriptome analysis in different stress conditions. We also discovered a possible role for transposons in regulation via small non-coding RNAs. As a result, our study indicate that transposable elements contribute to the variability and virulence of phytopathogenic bacteria. In this way, we believe that in the future, ETs might be a potential target for disease control in commercial cultures.

INTRODUÇÃO

Existem aproximadamente 150 espécies de bactérias que causam doenças em plantas. Esses fitopatógenos conseguem infectar uma significativa variedade de culturas, principalmente de interesse agrônômico, e por isso, são considerados uma das ameaças à segurança alimentar em todo mundo (Strange & Scott, 2005; Kannan et al., 2015). As bactérias fitopatogênicas obtêm nutrientes de seus hospedeiros para seu próprio crescimento e nesse processo causam diversos sintomas, como hiperplasia, nanismo, murcha, mancha, apodrecimento dos tecidos, dentre outros (Strange & Scott, 2005). Esses prejuízos são causados por fitopatógenos, que já foram e ainda são protagonistas de grandes epidemias em plantas, como por exemplo, a síndrome do declínio rápido da azeitona na Europa causada pela bactéria *Xylella fastidiosa* (Guo, 2020).

A maioria das estratégias de patogenicidade utilizadas por bactérias para causar doenças em plantas já é conhecida. Os sistemas de secreção de fitotoxinas, por exemplo, podem causar danos diretos aos tecidos vegetais como a podridão dos tecidos, já os fitohormônios são capazes de mimetizar hormônios vegetais desencadeando respostas em vias metabólicas, a fim de

impedir a defesa do hospedeiro (Pfeilmeier & Malone, 2016). Ademais, a existência de sistemas complexos de secreção, como o sistema de secreção do tipo III (T3SS), garante a transferência de diversos efetores às células dos hospedeiros (Pfeilmeier & Malone, 2016; Khan et al., 2018; Martins et al., 2018). Entretanto, apesar do amplo conhecimento sobre mecanismos de virulência, muitos fitopatógenos bacterianos ainda são de difícil controle, principalmente devido à alta plasticidade genômica. Logo, ainda existem algumas lacunas que precisam ser preenchidas para ajudar a entender como os mecanismos de patogenicidade são capazes de promover uma rápida adaptação das bactérias a diferentes condições bióticas e abióticas. Nesse sentido, os elementos transponíveis (ETs) se destacam por serem importantes geradores de plasticidade genômica. Os principais ETs encontrados em bactérias são as sequências de inserção (ISs) e os transposons (Tns). Esses elementos podem se inserir em outros genes e interferir em suas funções, logo são considerados os principais agentes mutagênicos naturais dos genomas. A inserção de ISs e Tns pode ainda alterar regiões regulatórias ou criar rearranjos cromossômicos (Serrato & Matute, 2018).

Embora a maioria das mutações resultantes da transposição seja prejudicial, a presença e atividade dos elementos transponíveis podem ser determinantes para a adaptação dessas bactérias ao hospedeiro, à diferentes nichos ecológicos, na eficiência dos mecanismos de patogenicidade, na colonização de novos hospedeiros e na suplantação de resistência. Além do mais, os Tns podem conter genes passageiros que codificam proteínas que podem promover maior adaptação, como por exemplo, genes relacionados à resistência à antibióticos (Partridge et al., 2018) ou fatores de virulência, como por exemplo, o transposon *TnXax1*, que desempenha um papel importante na patogenicidade de *Xanthomonas citri* (Oliveira et al., 2018). Além disso, muitos elementos podem ser “domesticados”, passando a codificar produtos de interesse da célula hospedeira e, portanto, desempenhando novas funções (Kapitonov et al., 2016; Peters et al., 2017).

Até o momento, sabe-se que os principais fitopatógenos possuem um número significativo de ETs em seus genomas (Robinson et al., 2012; Mat Razali & Nadarajah, 2019). Entretanto, pouco se sabe sobre o impacto de ISs e Tns na adaptação e virulência em bactérias fitopatogênicas. Sendo assim, esse estudo teve como objetivo identificar, caracterizar e realizar uma investigação das ISs e Tns encontrados nos genomas das 10 principais espécies bacterianas causadoras de doenças em plantas no mundo segundo o trabalho de Mansfield et al., 2012, sendo essas, *Pseudomonas syringae* patovares, *Ralstonia solanacearum*, *Agrobacterium tumefaciens*, *Xanthomonas oryzae* pv. *oryzae*, *Xanthomonas campestris* patovares, *Xanthomonas axonopodis* pv. *manihotis*, *Erwinia amylovora*, *Xylella fastidiosa*, *Dickeya*

(*dadantii* e *solani*), *Pectobacterium* (*carotovorum* e *atrosepticum*). Ademais, foram investigados o envolvimento desses elementos em eventos de recombinação ectópica, regulação via pequenos RNAs (sRNAs), respostas a diferentes condições de estresse e a uma possível domesticação.

MÉTODOS

Genomas e Transcriptomas

Duzentos e setenta genomas completos representando as dez espécies de bactérias fitopatogênicas de maior impacto científico e econômico no mundo (Mansfield et al., 2012), foram baixados em formato fasta a partir do banco de dados do National Center for Biotechnology Information (NCBI - <https://www.ncbi.nlm.nih.gov/assembly>) em agosto de 2020 (Tabela suplementar S1).

Dados de RNAseq pertencentes às espécies, *Pseudomonas syringae* isolado MAFF212134, *Xanthomonas oryzae* KACC10331 e *Ralstonia solanacearum* GMI1000 (números de acesso PRJNA261679, PRJNA629827 e PRJNA671670 respectivamente) foram baixados em formato fastq na plataforma Sequence Read Archive (SRA) (www.ncbi.nlm.nih.gov/sra). Os dados de transcriptomas analisados foram, das espécies: *R. solanacearum*, SRR12888322, SRR12888321, SRR12888320, SRR12888319, SRR12888318, SRR12888317, SRR12888329, SRR12888330 e SRR12888331; *P. syringae*, SRR11680825, SRR11680824, SRR11680823, SRR11680822, SRR11680821, SRR11680820, SRR11680819, SRR11680818, SRR11680817, SRR11680816, SRR80815, SRR11680814, SRR8569155, SRR8569156 e SRR8569157; e *X. oryzae*, SRR1582669, SRR1582668, SRR1582667, SRR1582666, SRR1582665, SRR1582664, SRR1582663, SRR1582662, SRR1582661, SRR1582660, SRR1582659, SRR1582658, SRR1582657 e SRR1582656. Os genomas correspondentes aos dados de transcriptomas foram:

NC_006834 (*X. oryzae*), NC_003295 (*R. solanacearum*) e GCF_002310955.1 (*P. syringae*).

Identificação e caracterização das sequências de inserção e transposons

A identificação e localização das sequências de inserção foi realizada por meio dos programas, *Semi-Automatic IS Annotation* (ISSaga) (Varani et al., 2011) e *Optimized Annotation System for*

Insertion Sequences (OASIS) (Robinson et al., 2012). A classificação das famílias e subgrupos das ISs preditas foi realizada de acordo com a literatura e o guia *Everyman's Guide to Bacterial Insertion Sequences* (Siguier et al., 2015) e para a identificação das DRs e TIRs dos elementos utilizou-se o programa Geneious 11.1.5.

A caracterização dos Tns, foi feita a partir dos estudos comparativos entre as transposases e as sequências terminais dos Tns preditos em relação a sequência usadas como referência depositadas nos bancos de dados, ISfinder e *Prokaryotic Transposable Element Database and Web Portal for Transposon Analysis* (Tncentral) (Ross et al., 2021). A presença de genes cargos importantes foi verificada via BLAST das ORFs presentes nos Tns contra os bancos de dados, Genbank, Uniprot, Virulence Factors Database (Chen et al., 2005), Pathogen–Host Interactions database (Urban et al., 2017), Type III Secretion System Database (Wang et al., 2012), The Comprehensive Antibiotic Resistance Database (Alcock et al. 2020) e ResFinder 4.0 (Clausen et al. 2018).

Inferência filogenética

Para avaliar se os transposons se agrupam de acordo com as espécies de bactérias as quais pertencem, duas árvores filogenéticas foram construídas, uma a partir das sequências do gene que codifica o rRNA 16S dos genomas analisados e outra para os transposons identificados. Para mais, também utilizamos sequências dos transposons, TN6768, TN6769, TN6770, TN6771, TN6772 e TN6773, todos previamente identificados em *R. solanacearum* (Gonçalves et al., 2020). Respectivamente, dois arquivos *txt* com as sequências foram criados, a etapa de alinhamento foi realizada via ClustalW (Thompson et al., 1994). As árvores filogenéticas geradas a partir das sequências do gene que codifica o rRNA 16S das bactérias e dos transposons, foram obtidas pelo método Likelihood Maximum, juntamente com o modelo evolutivo General Time Reversible com 1000 replicatas de bootstrap. Todas as etapas foram realizadas no software MEGAX (Kumar et al., 2018) e a visualização e anotação da filogenia foi realizada na interface Interactive Tree of Life (iTOL) v4 (<https://iTOL.embl.de/>).

Rearranjos cromossomais mediados por ISs e Tns

Investigamos a possibilidade de ocorrência de possíveis rearranjos nos cromossomos que poderiam estar sendo causados por ETs. Para cada espécie escolheu-se dois genomas para a análise, sendo um com o número relativamente alto de ISs e outro com número baixo desses

elementos. A visualização dos possíveis rearranjos decorrentes da presença de ISs e Tns foi realizada a partir de uma análise de sintenia dos genomas utilizando o software *Progressive Mauve* disponível no software Mauve versão 2.3.1 (Darling et al., 2004). Nas posições adjacentes aos rearranjos, foram verificadas as presenças de ISs e Tns. As sequências de elementos pertencentes a mesma família e ao mesmo subgrupo presentes nas bordas das sequências avaliadas, foram analisadas quanto a existência de sinais de recombinação, para isso foi utilizado o software RDP3 (Recombination Detection Program) (Martin et al., 2010). Este programa caracterizou os eventos de recombinação em um alinhamento múltiplo prévio das sequências de nucleotídeos dos elementos pertencentes a uma mesma família, pelo software ClustalW, juntamente com um conjunto de sete métodos diferentes: Programa de detecção de recombinação (RDP), GENECONV, 3Seq, Chimera, Máximo Chi Square (MaxChi), SiScan e BootScan. Esses métodos detalharam os locais de quebra e recombinação, além de indicarem as identidades das sequências recombinantes parentais. Somente eventos de recombinação detectados por pelo menos 4 métodos foram considerados verdadeiros.

Pressão de seleção em ISs e Tns

Sequências de ISs próximas a genes relacionados à adaptação celular a condições adversas e sequências inseridas na mesma posição em diferentes cepas, mas pertencentes à mesma espécie, foram selecionadas para averiguar se estão sob algum tipo de pressão de seleção. Desse modo, foi aplicado o teste de neutralidade Tajima's D (Tajima, 1989) via DnaSP v6 (Rozas et al., 2017). Esse método trabalha com a hipótese de que em uma população todas as mutações são seletivamente neutras. Sendo assim, o teste correlaciona o número de polimorfismos com o número médio de sítios de segregação, o resultado também chamado de D é interpretado da seguinte forma, D=0, sem evidências de seleção, D < 0 seleção de varredura (selective sweep) ou expansão populacional após gargalo, D > 0 seleção balanceadora ou contração repentina da população. Nos casos em que teste de Tajima foi significativo (P < 0,001), a pressão seletiva sobre as ISs foi confirmada pelo teste Z, implementado no MEGAX (Kumar et al., 2018). Dessa forma, o alinhamento de sequências de elementos com número de substituições não sinônimas maior que o de substituição sinônimas ($d_N > d_S$), estão sob pressão positiva, sendo o contrário, ($d_N < d_S$), considerado seleção purificadora ou negativa.

Análise de Transcriptomas

Para a identificação de elementos transponíveis diferencialmente expressos em condições de estresse, três dados de RNAseq pertencentes aos genomas de *P. syringae*, *X. oryzae* e *R. solanacearum* foram analisados. O controle de qualidade das sequências foi realizado utilizando o programa FastQC (v0.11.5) e Trimmomatic/0.36 (Bolger et al, 2014). Na etapa de mapeamento, as *reads* foram alinhadas com as sequências referência, nesse caso um representante dos subgrupos das ISs preditas. Tanto o mapeamento quanto a comparação dos níveis de expressão foram realizadas pelo software Geneious 11.1.5 de acordo com o tutorial R DESeq2. As análises dos elementos diferencialmente expressos foram realizadas comparando a condição controle e os tratamentos. O perfil de expressão para os elementos pertencente a *P. syringae* foram comparados em meio LB e meios indutores de coronatina (HS e HSC) a temperaturas de 18 ° C e 27 ° C (Hirose et al., 2020). Enquanto ETs encontrados em *X. oryzae* foram expostos a um meio indutor de patogenicidade contendo extratos das folhas do arroz, denominado RLX (Kim et al, 2016). Por fim, em *R. solanacearum* analisamos a expressão de ETs quando submetidos ao estresse nitrosativo (NO) e oxidativo, onde as células foram tratadas com concentrações diferentes de óxido nítrico e peróxido de hidrogênio (Hendrich et al., 2021). A expressão diferencial foi calculada usando, Log2 Fold-Change (Log2FC > 0, up-regulated) e (Log2FC < 0, down-regulated), sendo p-value < 0,05. Para geração dos heatmaps os valores foram plotados no Interactive Tree of Life (iTOL).

Predição de sRNAs reguladores

A análise de predição sRNAs em ETs teve como objetivo identificar putativos sRNAs reguladores associados aos Tns preditos e as ISs preditas que tiveram seu nível de transcrição diferencialmente expresso. Um total de 78 elementos foram avaliados, sendo 28 presentes nos genomas de *P. syringae*, 17 em *X. campestris*, 25 em *X. oryzae*, 2 em *X. axonopodis* e 6 em *R. solanacearum*. Foram utilizadas quatro estratégias, cada uma delas com uma abordagem de predição de sRNAs diferente. A primeira identificação foi realizada por dois algoritmos, INFERNAL (Nawrocki & Eddy, 2013) e RNAz (Gruber et al., 2010). Esses localizadores de sRNAs são baseados em análise comparativa, sendo assim usam homologias em nível de sequências e estrutura para realizar as predições de potenciais sRNAs. O primeiro, constrói estruturas secundárias de sRNAs a partir da sequência de entrada, denominados de modelos de

covariância, e esses perfis buscam combinações consenso em bancos de dados de sequências de sRNAs. Já o segundo, prediz estruturas de sRNA, analisando a conservação da estrutura secundária e a estabilidade termodinâmica. Essas duas características são usadas para formar uma estrutura secundária consenso de sRNA e os dados são preditos a partir de alinhamentos múltiplos no genoma. Outro algoritmo utilizado foi o PresRAT (Kumar et al., 2020), esse usa as informações das sequências de entrada e das estruturas secundárias, nesse caso para determinar a existência de um sRNA. Assim é realizado um conjunto de cálculos que utilizam as probabilidades de matrizes de pontuação Log Odds (LOD) para avaliar todas as possíveis combinações, as conformações da estrutura e a energia média das estruturas secundárias (Kumar et al., 2020).

Outra estratégia foi realizar a detecção de sRNAs a partir de dados de sequenciamento de RNA, utilizando para isso o programa Oasis 2. O método consiste em alinhamentos das *reads* divididos em quatro etapas: alinhamentos entre os transcritos do organismo alvo com o banco de dados miRbase (Kozomara & Griffiths, 2019); alinhamento das *reads* não mapeadas utilizando o alinhador Bowtie 1.0, (Rahman et al., 2018); caso ainda haja *reads* não mapeadas, essas são alinhadas com genomas bacterianos, arqueais e virais e por fim, as *reads* ainda serão comparados com organismos não alvos depositados em banco de dados miRbase pelo programa STAR 2.4, que tem como objetivo detectar potenciais sRNAs ortólogos (Rahman et al., 2018). Todas as sequências de sRNAs preditas foram analisadas pelo software GLASSgo (Lott et al., 2018) para identificação de sRNAs homólogos. Este programa avaliou a estrutura secundária das sequências selecionadas e buscou no banco de dados do NCBI sequências com identidade acima de 52% (Lott et al., 2018).

Análise de alvos

Como o programa encontrou homólogos para todos os sRNAs preditos, foi utilizado o software CopraRNA (Wright et al., 2014), que realizou as predições de alvos a partir das sequências homólogas dos sRNA candidatos. Essa análise também permitiu o enriquecimento funcional dos principais genes alvos e a visualização das regiões de interação (Wright et al., 2014). Foram considerados sRNAs verdadeiros apenas aqueles que apresentaram o mesmo alvo nos diferentes algoritmos e que tiveram valores de probabilidade amostrados igual a zero (CopraRNA p-value = 0). A anotação funcional dos mRNAs alvos preditos foi realizada por duas formas, pesquisa manual por palavra-chave, utilizando o nome do gene sugerido pelo programa CopraRNA no

banco de dados KEGG Orthology (KO) e por alinhamento entre as sequências preditas contra os bancos de dados Uniprot e Pfam (protein families database).

RESULTADOS

Perfil de distribuição de sequências de inserção nas 10 principais bactérias fitopatogênicas

Foi realizada a predição de 35.692 de ISs, nos 270 genomas completos analisados (Tabela S2 e S3). A distribuição de ISs nos genomas das dez principais bactérias fitopatogênicas foi bastante heterogênea, já que algumas espécies possuem um maior número de cópias desses elementos em relação a outras (Figura 1A). Com base no número de elementos preditos em cada genoma, foi realizado a mediana para estimar o número de sequências de inserção para cada espécie. Sendo assim, os resultados mostraram 66,5 elementos por genoma em *P. syringae*, 60 em *R. solanacearum*, 10,5 em *A. tumefaciens*, 383,5 em *X. oryzae*, 65 em *X. campestris*, 60 em *X. axonopodis*, 2 em *E. amylovora*, 5,5 em *X. fastidiosa*, 21 em *D. dadantii* e 6 em *D. solani*, 6 em *P. carotovorum* e 8 em *P. astrosepticum*.

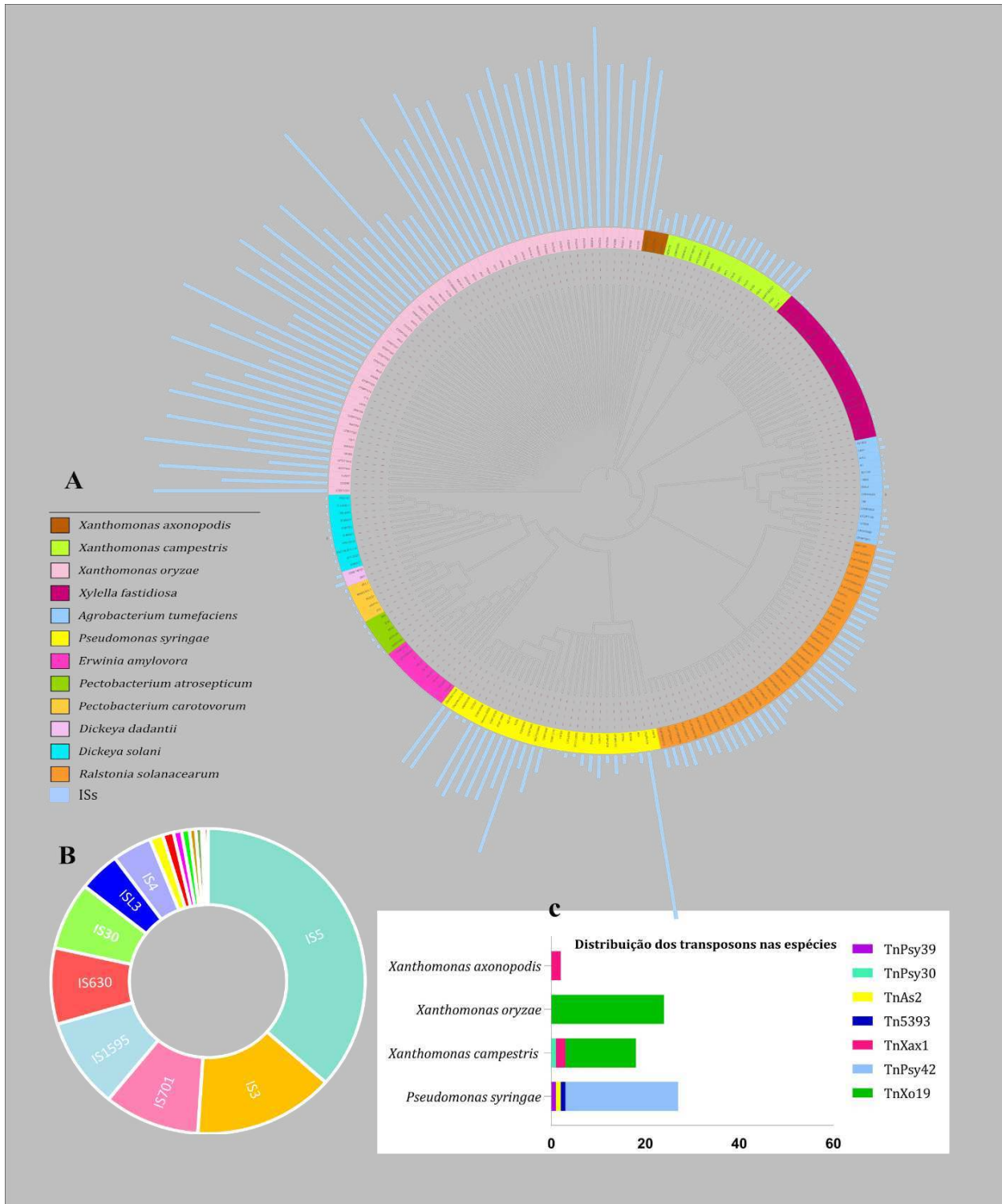


Figura 1: Perfil de distribuição de sequências de inserção e transposons encontrados nos genomas completos analisados. (A) A árvore filogenética foi inferida pelo método de neighbor-joining usando sequências do gene que codifica o rRNA 16S dos 270 genomas bacterianos investigados que estão separados por cores. O gráfico de barras ao redor da árvore indica o número de elementos preditos para cada genoma. (B) Proporção das famílias de elementos mais abundantes presentes nos genomas. (C) Distribuição e quantidade de transposons encontrados nos 270 genomas investigados.

De forma geral, as famílias de ISs que tiveram um maior número de representantes nos genomas foram, IS5, IS3, IS701, IS1595 e IS630. Embora IS5, IS701, IS630 e IS1595 não estejam presentes em todas as espécies, seu elevado número de cópias é relatado em *X. oryzae*, onde parece haver uma expansão desses elementos. No entanto, em relação a representatividade, os elementos IS3, IS110 e IS4 são os elementos presentes na maioria das espécies.

O elemento IS5, presente em sete espécies, possui 13.102 cópias e se subdivide em seis subgrupos com aproximadamente 800 a 1.500 pares de bases (pb), sendo que três subgrupos foram verificados, IS5ssgrIS5, IS5ssgrIS1031 e IS5ssgrIS427. Com 5.197 cópias, IS3 é a família que está presente em maior número de espécies, estando somente ausente em *X. fastidiosa*. Os elementos IS3 abrangem cinco subgrupos que variam de 1.000 a 1.750 pb de comprimento e codificam duas ORFs parcialmente sobrepostas. Nos genomas analisados há a presença de todos os subgrupos desse elemento. Estando ausente em apenas três espécies, IS4 e IS110 também compõem o grupo de famílias com o maior número de representantes nos dez fitopatógenos. Elementos da família IS4 possuem entre 1.400 a 1.600 pb, é uma família extremamente heterogênea, entretanto estruturas conservadas ajudam na distinção dos subgrupos. Até o momento sete membros dessa família já foram identificados, enquanto a família IS110 possui apenas dois subgrupos registrados. Elementos da família IS110 já foram detectados em mais de 130 espécies bacterianas, com tamanho entre 1.200 a 1.500 pb; é a única família a carregar uma transposase do tipo DEDD, além de parecer possuir uma região de inserção específica. Por outro lado, IS701 foi identificado em apenas duas espécies, *R. solanacearum* e *X. oryzae*,. Porém, somente em *X. oryzae* esse elemento está presente em 3.537 cópias que variam entre 1.400 a 1.550 pb e codifica apenas uma ORF. Basicamente presente apenas nos gêneros *Xanthomonas* e *Ralstonia*, IS1595 tem cerca de 1.000 pb de comprimento e codifica uma ORF. Já foi relatado anteriormente, uma expansão desse elemento em *X. oryzae* (Ross et al., 2021), e nossos dados mostram que em 3.418 cópias, 3.295 são pertencentes a essa espécie. Por fim, foram identificadas 2.893 cópias de IS630, que pode ter uma ou duas ORFs e apresenta aproximadamente 1250 bp. Em nossos dados esse elemento está presente nas quatro principais bactérias, de acordo com a classificação elaborada por Mansfield et al., 2012. Além das cinco famílias abordadas aqui, outras 15 famílias estão presentes na tabela S2 e S3 e a representação das principais famílias pode ser verificada na Figura 2.

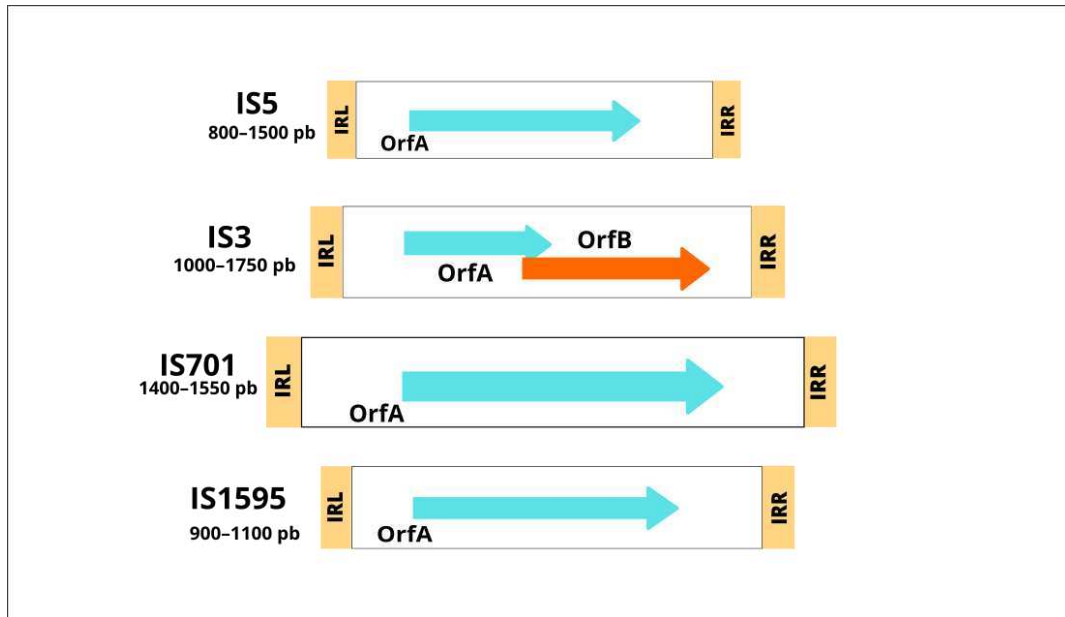


Figura 2: Representação das quatro principais famílias de ISs nos genomas das top 10 bactérias. As caixas laranjas em cada extremidade representam as repetições invertidas do terminal esquerdo (IRL) e direito (IRR) e as setas em verde e laranja os quadros abertos de leitura (ORFs).

Apesar de esperarmos que espécies do mesmo gênero apresentariam um padrão similar de famílias de ISs, nossos resultados mostraram que bactérias pertencentes aos gêneros *Xanthomonas* e bactérias pectinolíticas como *Pectobacterium* e *Dickeya* apresentaram divergências entre o tipo das famílias encontradas e ou no número de cópias delas nas diferentes linhagens (Tabela 1). Ademais, nenhuma das famílias previstas se encontram em todas as espécies e algumas somente são encontradas em determinado organismo como é o caso dos elementos IS6 e IS607 detectados apenas em *X. fastidiosa* e ISAs1 em *R. solanacearum*, sendo essa cópia única. No mais, nas espécies que investigamos apenas não foi detectada a presença das famílias IS1, ISH3, IS1634, IS6, IS982, IS1380, ISAzo13, ISH6, ISKra4 e ISLre2.

Tabela 1: Número de cópias de ISs presentes em cada espécie de bactéria investigada.

Famílias	Ps	Rs	At	Xo	Xc	Xa	Ew	Xf	Dd	Ds	Pc	Pa	Nº de ISs por família
IS1595	-	28	-	3295	7	40	-	-	-	-	-	-	3370
IS3	939	388	35	3178	453	131	22	-	19	20	8	4	5197
IS481	8	1	2	191	-	9	-	-	-	-	2	3	216
IS4	234	386	32	631	61	6	-	-	2	-	2	2	1356
IS701	-	19	-	3518	-	-	-	-	-	-	-	-	3537
IS5	322	597	13	11755	308	106	-	-	-	-	-	1	13102
IS6	-	-	-	-	-	-	-	1	-	-	-	-	1
IS1182	384	14	1	-	-	-	-	-	-	-	1	1	401
IS21	192	89	6	1	-	-	-	-	-	-	-	-	288
IS30	31	35	2	2493	-	-	-	-	-	-	-	-	2561
IS66	158	86	15	31	-	-	-	-	-	-	-	-	290
IS110	65	71	13	37	4	-	-	-	29	10	2	2	233
IS256	74	6	-	408	-	-	-	-	1	-	-	-	489
IS630	675	14	2	2192	-	-	-	-	-	-	-	10	2893
ISAs1	-	1	-	-	-	-	-	-	-	-	-	-	1
ISL3	16	67	-	1310	43	5	-	39	1	30	7	4	1522
IS607	-	-	-	-	-	-	-	35	-	-	-	-	35
IS91	14	-	-	3	3	-	8	-	-	-	-	6	34
IS200/IS605	-	-	-	-	-	6	-	2	6	-	-	-	14
ISNCY	-	20	6	3	1	-	-	-	-	8	10	3	51
Tn3	23	30	-	17	18	1	-	-	1	11	-	-	101
Nº de ISs por espécie	3135	1852	127	29063	898	304	30	77	59	79	32	36	TOTAL

Cada bactéria fitopatogênica está representada pelas iniciais do nome da espécie, sendo Ps (*Pseudomonas syringae*), Rs (*Ralstonia solanacearum*), At (*Agrobacterium tumefaciens*), Xo (*Xanthomonas oryzae*), Xc (*Xanthomonas campestris*), Xa (*Xanthomonas axonopodis*), Ew (*Erwinia amylovora*), Xf (*Xylella fastidiosa*), Dd (*Dikeya dadantii*), Ds (*Dickeya solani*), Pc (*Pectobacterium carotovorum*) e Pa (*Pectobacterium atrosepticum*).

Além das variações entre o tipo das famílias presentes em cada espécie, também detectamos variação no número de elementos encontrados para cada fitopatógeno. É possível observar que existem espécies como, *X. oryzae* e *P. syringae* que estão infladas em relação a quantidade de elementos do que outras bactérias como *E. amylovora* e *X. fastidiosa* (Figura 1A; Tabela 1). Em muitos organismos, principalmente eucariotos, o aumento do tamanho do genoma está intrinsicamente relacionado ao aumento do número ou quantidade de elementos transponíveis (Touchon & Rocha, 2007). Desse modo, calculamos o coeficiente de correlação de Pearson (r) para verificar se há essa relação também é verdadeira para as bactérias analisadas (Figura suplementar 1). Encontramos um coeficiente positivo de 0,000441 entre o comprimento de cada genoma e seu número de ISs, mostrando que para as 12 espécies estudadas não houve uma relação de dependência entre o número de elementos transponíveis encontrados nos genomas e o tamanho dos genomas.

Perfil de distribuição de transposons nas 10 principais bactérias fitopatogênicas

O número total de transposons detectados foi de 71, todos pertencentes à família Tn3 (Figura1C; Tabela2 e Tabela S4). Apenas em quatro espécies, *P. syringae*, *X. campestris*, *X. oryzae* e *X. axonopodis*, foram encontradas a presença de transposons nos cromossomos, os demais foram identificados em plasmídeos. Em *P. syringae*, foram encontrados 26 transposons, sendo 24 sequências correspondentes ao elemento TnPsy42 (Butler et al., 2013) e as demais aos elementos TnAs2 (Pfeiffer et al., 2018) e Tn5393 (Chiou & Jones, 1993). Em *X. campestris* foram encontradas 15 cópias pertencentes ao TnXo19 (Niu et al., 2015), uma de TnPsy30 (Bardaji et al., 2011) e duas do elemento TnXax1 (Ferreira et al., 2015) que também está presente em dois genomas de *X. axonopodis*. Por fim, todos os 24 transposons encontrados em *X. oryzae* se referem-se ao elemento TnXo19. Ademais, identificamos pela primeira vez vinte transposons nos genomas de três fitopatógenos que ainda não haviam sido relatados. Em *P. syringae*, nos isolados NZ-45, NZ-47 e P155 foi encontrado o transposon TnPsy42, em *X. campestris*, nos isolados NEB122, 3811, CN03, CN12, CN17 e CN18, detectamos o transposon TnPsy30 na primeira cepa e nas demais, TnXo19. Por fim, em *X. oryzae*, o transposon TnXo19 até o momento foi relatado apenas em plasmídeos, no entanto, no nosso trabalho é relatado pela primeira vez a presença desse elemento em cromossomos, nos isolados, CIAT, ICMP3125, ITCCBB0002, NX0260, PX079, PXO83, PXO86, PXO99A, PXO145, PXO211 e PXO236.

Tabela 2: Transposons encontrados nos 270 genomas completos das principais bactérias fitopatogênicas.

Espécies	Nome	Família	Número encontrado
<i>Pseudomonas syringae</i>	Tn5393	Tn3	1
	TnPsy42	Tn3	24
	TnPsy39	Tn3	1
	TnAs2	Tn3	1
<i>Xanthomonas campestris</i>	TnPsy30	Tn3	1
	TnXo19	Tn3	15
	TnXax1	Tn3	2
<i>Xanthomonas oryzae</i>	TnXo19	Tn3	24
<i>Xanthomonas axonopodis</i>	TnXax1	Tn3	2

Apesar de não terem sido identificadas inserções de transposons a menos de 100 pb de genes relacionados à virulência ou resistência a antimicrobianos, houve a identificação de genes de virulência e de resistência que podem estar sendo carregados por esses elementos preditos

(Figura 3). O elemento transponível Tn5393, encontrado em *P. syringae* (isolado B728a) apresenta um tamanho de 5.822 pb e possui duas extremidades repetitivas e invertidas (IRs). Além disso, esse elemento contém ORFs que codificam uma transposase da família Tn3 (2.886 pb) e um par de genes de resistência à estreptomicina, *strA* e *strB* que apresentam 831pb. Ainda, nesse fitopatógeno o elemento TnAs2 de 9.482 pb, foi encontrado no genoma de CFBP3846-1. Esse ET possui um gene que codifica uma transposase (Tn3) com 2.970 pb e um que codifica uma resolvase (546 pb), sendo esses dois genes responsáveis por sua transposição. Ainda foi identificado um gene que codifica um efetor, *hopX* de 1115 pb, pertencente ao sistema de secreção do Tipo III (TSS3). Do mesmo modo, o transposon TnXo19 (16.004 pb), encontrado em *X. campestris* CN03, possui um gene que codifica para uma proteína do sistema de secreção TSS3 denominada Popp2 (1698 pb) e uma transposase pertencente à família Tn3 com 2.958 pb. Ademais, nessa mesma espécie, porém nos isolados, CN17 e CN12, detectamos um transposon de 18.852 pb contendo um gene que codifica uma transposase (2.958 pb) e carregando um gene que codifica uma proteína membro da família *avrBs3* (3.192 pb), identificado como uma proteína efetora que geralmente é encontrada em transposons da família Tn3 em gêneros de *Xanthomonas* (Ferreira et al., 2015).

De acordo com a análise filogenética da família Tn3, os transposons se agruparam em grande parte, de acordo com a espécie em que estão localizados. Entretanto, foi evidenciado que o elemento Tn5393, pertencente a *P. syringae*, se encontra próximo ao grupo de elementos da espécie *R. solanacearum*. Do mesmo modo, o elemento Tn6772, presente em *R. solanacearum*, se agrupa com elementos da espécie *X. oryzae*. Ademais, o elemento TnPsy30 em *X. campestris* se encontra mais próximo de elementos da espécie *P. syringae* (Figura 4).

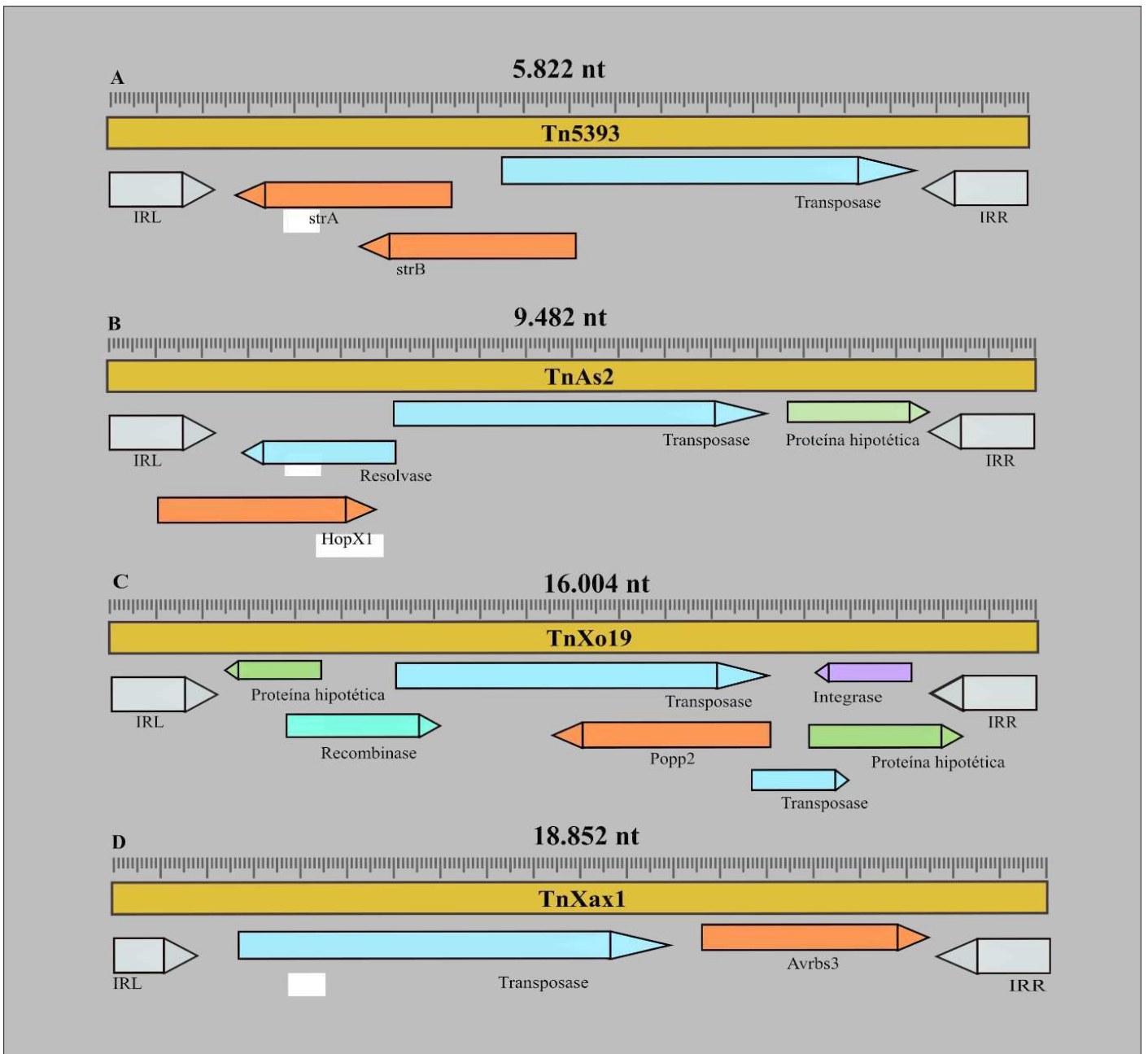
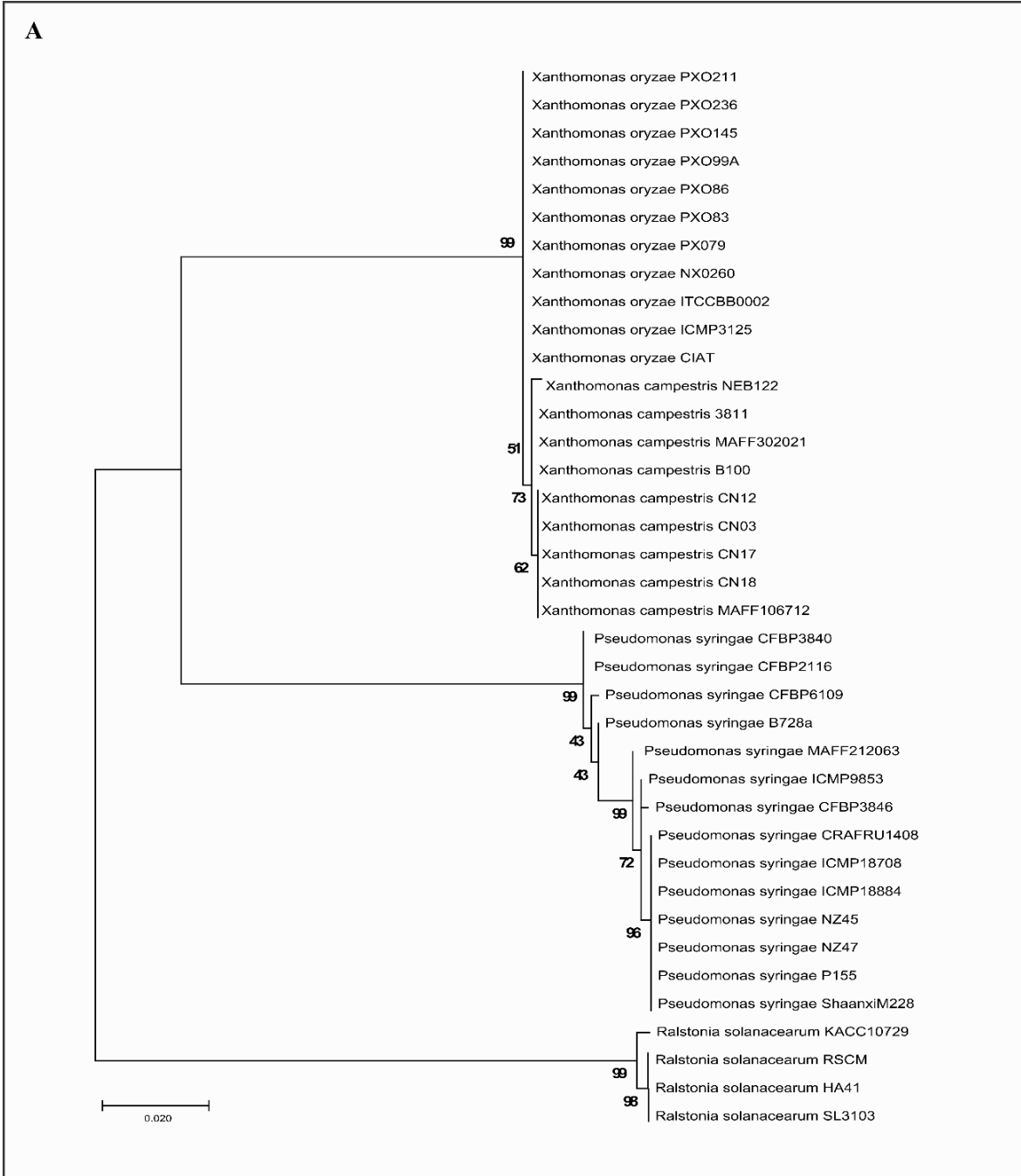


Figura 3: Representação dos transposons preditos que podem estar carregando genes de patogenicidade. (A) Elemento Tn5393 pertencente a espécie *Pseudomonas syringae*, evidenciando um par de genes de resistência a estreptomicina e TnAs2 (B) e TnXo19 (C) os quais transportam genes que codificam proteínas efetoras pertencentes ao sistema de secreção do tipo III e em (D) TnXax1 encontrado em isolados de *X. campestris* carregando um gene que codifica um fator de avirulência.



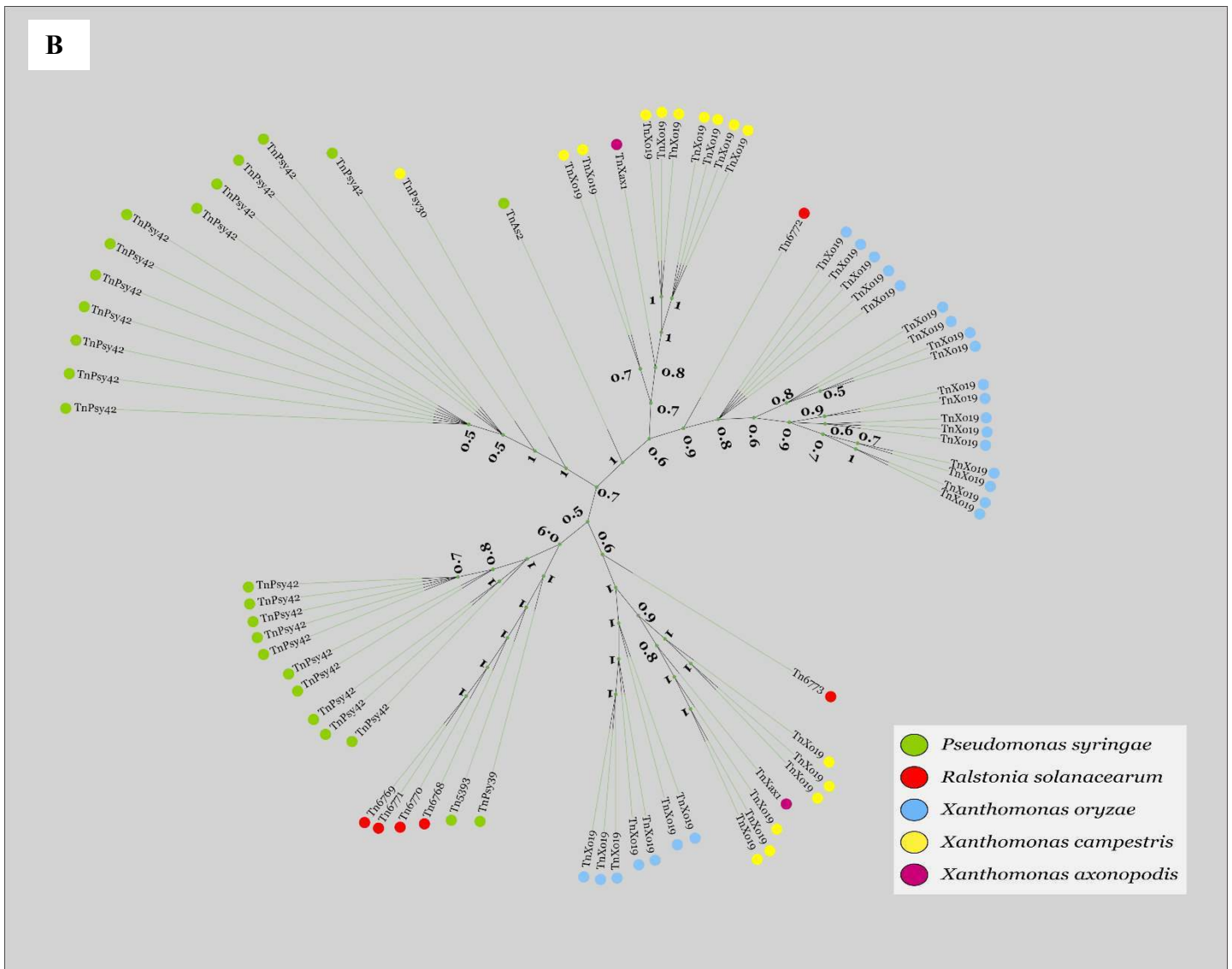


Figura 4: Interferência filogenética. (A) Distribuição dos isolados onde foram localizados os transposons, inferência filogenética realizada a partir das sequências do gene que codifica o rRNA 16S das bactérias. (B) Árvore filogenética das cópias completas dos transposons, as cores indicam as espécies as quais os elementos foram identificados. Ambas foram construídas por inferência do método Maximum Likelihood, levando em conta o modelo evolutivo General Time Reversible.

Rearranjos cromossomais mediados por ISs e Tns

A evidência de potenciais recombinações foi detectada nos genomas completos de três espécies, *X. axonopodis*, *X. campestris* e *X. oryzae*. Utilizando a regra da maioria, foram considerados eventos de recombinação verdadeiros apenas aqueles detectados por mais de quatro métodos dos sete implementados no software RDP3, sendo assim para *X. oryzae* e *X. campestris* foi identificado um evento de recombinação e quatro eventos para *X. axonapodis*. Desse modo, apenas ocorreu a identificação de recombinantes para elementos pertencentes às famílias IS3 e IS5 (Tabela 3). A visualização dos possíveis rearranjos decorrentes da presença de ISs está representada na Figura S2.

Tabela 3: Evidência de potenciais eventos de recombinação detectados.

Espécie	Família	Possíveis elementos envolvidos em eventos de recombinação	Método de detecção
<i>Xanthomonas axonopodis</i>	IS3	617140 – 618175, 3436994-3438096,	GENECOV, BootScan, MaxChi, 3Seq
<i>Xanthomonas axonopodis</i>	IS3	402772 – 401670, 4728014 – 4729116	RDP, GENECOV, Chimera, MaxChi, BootScan, SiScan, 3Seq
<i>Xanthomonas axonopodis</i>	IS3	617140-618175, 4203695 – 4202593, 836421-837523	RDP, GENECOV, Chimera, MaxChi, BootScan, SiScan, 3Seq
<i>Xanthomonas axonopodis</i>	IS3	4203695 – 4202593, 3436994 – 3438096	RDP, GENECOV, Chimera, MaxChi, BootScan, SiScan, 3Seq
<i>Xanthomonas campestris</i>	IS5	2820068 – 2820874, 2830209 – 2829229	RDP, GENECOV, Chimera, MaxChi, BootScan, SiScan, 3Seq
<i>Xanthomonas oryzae</i>	IS3	498588 - 497486,	Chimera, MaxChi, SiScan, 3Seq

Análise de pressão de seleção em ISs e Tns

Apesar do esperado impacto negativo da movimentação dos ETs, em alguns casos esses podem ser domesticados ou exaptados para novas funções. Verificamos a existência de ISs próximos a genes de virulência e adaptação em todas as espécies (Tabela S5). A maioria dos elementos estavam em regiões próximas a genes de sistemas de transporte que estão associados à captação de nutrientes e a exportação de toxinas, metais pesados e antibióticos. Ademais, também identificamos inserções próximas a regiões contendo genes de adaptação ao estresse de

solanacearum, observamos que o elemento IS3ssgrIS51 está localizado na posição 1.195.087 – 1.194.212 em 20 dos 40 genomas analisados, entretanto esse não se encontra próximo a outros genes (menos de 100 pb).

Os resultados dos testes D de Tajima confirmaram a existência de pressão de seleção em todos os elementos (Tabela 4). O cálculo de substituições sinônimas e não sinônimas entre as sequências alinhadas, realizado pelo teste Z de seleção, indicou que os elementos IS3 próximos à genes relacionados a família ATPases e IS3ssgrIS51 estão sob pressão de seleção positiva, diferente dos elementos IS3, próximo a genes de bombas de efluxo ABC, que se encontra sob pressão de seleção purificadora, indicando uma possível inserção deletéria (Tabela 4).

Tabela 4: Resultados das análises de pressão de seleção.

Gene	Tajima D	Teste Z de seleção			Seleção
		Testes	Probabilidade	Estatística	
IS3 (AAA Família ATPase)	-1,3685	Teste de neutralidade	0,79	(-0,27)	Positiva
		Seleção positiva	1	(-0,27)	
		Seleção purificadora	0,39	(-0,27)	
IS3 (ABC bombas de efluxo)	-1,63687	Teste de neutralidade	0,1	(1,64)	Purificadora
		Seleção positiva	0,5	(1,64)	
		Seleção purificadora	1	(-1,64)	
IS3ssgrIS51	-1,76746	Teste de neutralidade	0,08	(-1,76)	Positiva
		Seleção positiva	1	(-1,75)	
		Seleção purificadora	0,04	(1,79)	

Análises dos elementos diferencialmente expressos

Os níveis de expressão dos elementos transponíveis analisados em situações de estresse, diferiram entre 10 elementos pertencentes as famílias Tn3, IS66, IS21, IS256, IS91, IS3, IS701, IS1595, IS5 e IS4 (Figura 6). O grupo de elementos pertencentes a *P. syringae*, foram expostos

a condições de cultivo que utilizaram os meios indutores de coronatinas denominados HS e HSC a temperaturas de 18°C e 27°C. Nossas análises verificaram que houve um aumento da expressão dos genes que codificam transposases de alguns elementos, sendo esses Tn3, IS256, IS21 e IS91, quando expostos a temperaturas mais elevadas e uma supressão da família IS66 em todas as condições em comparação com o controle. Ademais, de acordo com nossos resultados o cultivo em meio LB comparado aos meios HS e HSC não provocou mudanças significativas na expressão dos elementos transponíveis avaliados.

Em *X. oryzae* o experimento buscou avaliar o perfil de expressão diferencial quando a bactéria foi submetida a um meio de cultivo com um indutor de patogenicidade denominado RLX (Kim et al, 2016). Observou-se que houve uma correlação significativa entre as mudanças log₂ -fold nos níveis de expressão que aumentaram nos elementos, IS4ssgrIS10, IS3ssgrIS407 e IS3ssgrIS3. Ademais, verificou-se que duas cópias do elemento IS3ssgrIS407 estão próximas a genes de adaptação e virulência, uma a 79pb de uma família de transportadores, MFS, e a outra próxima, 69 pb, de um gene que codifica uma enzima envolvida na degradação de tecidos vegetais, pectato-liase.

Quanto as análises de expressão diferencial com *R solanacearum*, essa foi submetida ao estresse oxidativo e de óxido nítrico, observou-se que houve um maior número de ISs se expressando de maneira diferencial para a segunda condição mencionada. Assim, os elementos IS3ssgrIS2, IS3ssgrIS3, IS5 e IS5ssgrIS427, tiveram a transcrição dos genes que codificam transposases aumentada, enquanto que em IS1595ssgrIS1016, IS3ssgrIS150 e IS701 esses genes foram suprimidos. Já o estresse provocado pela falta de oxigênio, levou a diminuição da transcrição dos elementos IS3ssgrIS150 e IS5ssgrIS5. Além disso, observamos que IS3ssgrIS150 está inserido próximo a dois genes, a 100 pb de proteína de membrana denominada porina e a 79 pb operon PcaQ. Esses também tiveram sua expressão diminuída sendo que a porina em escassez de óxido nítrico e PcaQ na falta de oxigênio, quando comparadas ao controle.

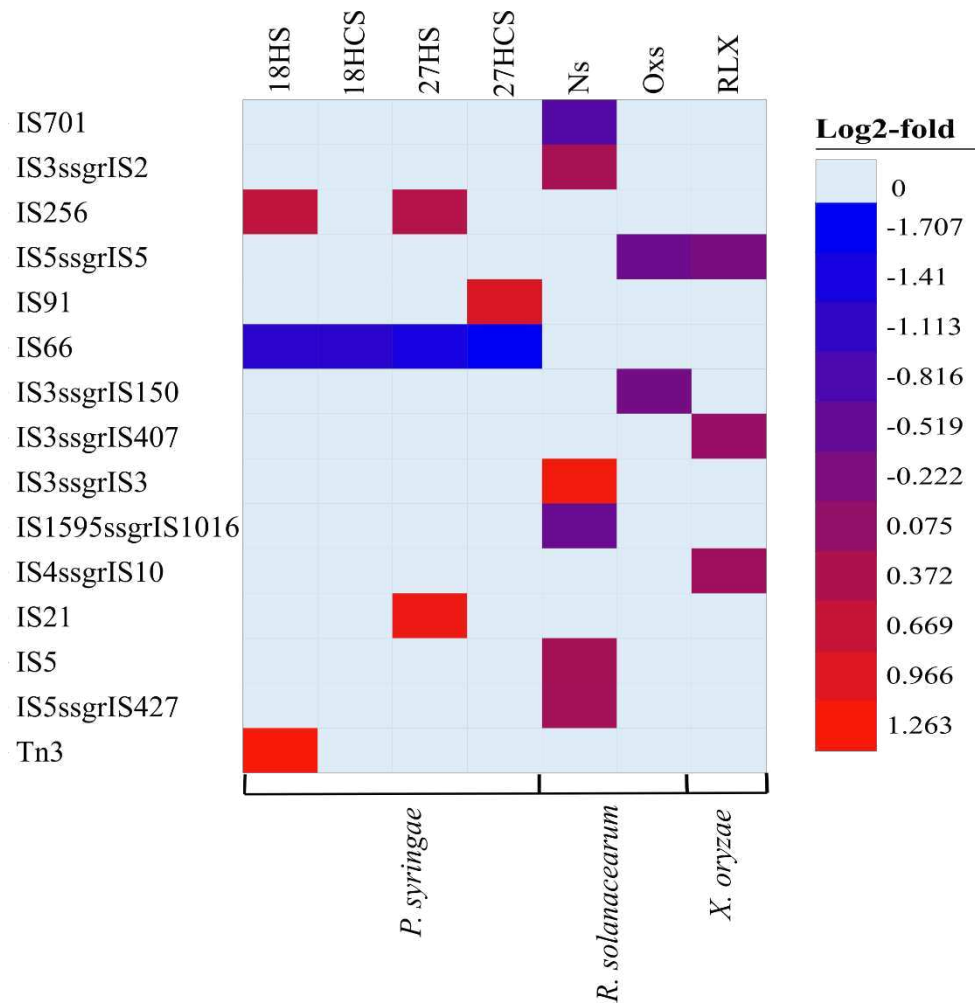


Figura 6: O nível de expressão de cada elemento é mostrado de acordo com os valores Log2-fold. Elementos que tiveram sua transcrição aumentada são mostrados em vermelho e os que foram suprimidos em azul. 18HS, 18HCS, 27HS e 27HCS: cultivo em meios indutores de coronatinas nas temperaturas de 18°C e 27°C. Ns e Oxs condição de estresse ao óxido nítrico e oxidativo, respectivamente. RLX cultivo em meio indutor de patogenicidade.

Pequenos RNAs reguladores codificados por ETs

Uma vez que observamos a diferença de expressão de algumas famílias devido a mudança das condições ambientais, realizamos uma análise para identificar possíveis sRNAs que poderiam estar envolvidos e presentes nos elementos identificados. Desse modo, observamos que o elemento IS3ssgrIS150 em *R. solanacearum* GMI1000 codifica um sRNA predito de 81 nucleotídeos, identificado por todos os algoritmos testados. Os 28 homólogos identificados para esse sRNA variaram entre 70 e 90 nucleotídeos, sendo que cinco deles foram

preditos na espécie *R. pseudosolanaceraum*. O nível de identidade entre os homólogos foi superior a 80% e a estrutura secundária predita pode ser observada na Figura 7A.

Ademais, também realizamos a mesma análise para os 71 transposons encontrados nas bactérias fitopatogênicas analisadas. Encontramos evidência de um possível sRNA codificado pelo transposon TnXo19 pertencente a *X. campestris* CN17 (Figura 7B). O sRNA predito por três dos quatro métodos utilizados (INFERNAL, RNAz e PresRAT), possui 85 nucleotídeos e foi encontrado em diferentes patovares do gênero *Xanthomonas*, sendo no total 115 homólogos identificados com identidade superior a 80%. A avaliação dos potenciais alvos de regulação dos sRNAs preditos revelou que o elemento IS3ssgrIS150 interage com alvos cognatos, sendo que no genoma de *R. solanacearum* GMI100 são encontradas três cópias desse elemento. Já o sRNA codificado pelo transposon TnXo19 apresentou como possíveis alvos genes que codificam ativadores transcricionais de células vegetais, *avrbs3*, que está sendo carregado pelo próprio elemento (Tabela 5).

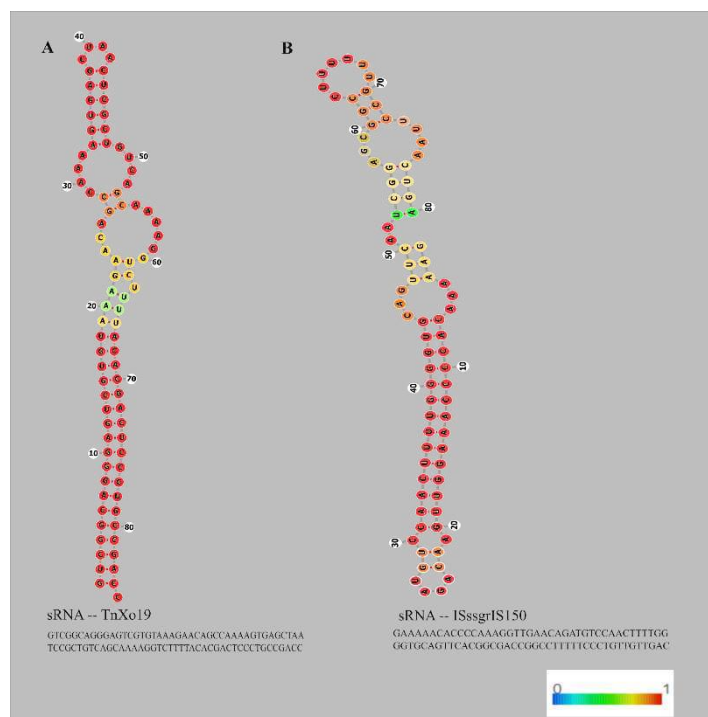


Figura 7: Estruturas secundárias dos sRNA codificados pelos elementos transponíveis (A) IS3ssgrIS150 e (B) TnXo19 obtidas por RNAfold. Abaixo da figura é possível visualizar a sequência de nucleotídeos de cada sRNA. A barra de cor indica a probabilidade do pareamento entre sRNA e o mRNA do gene alvo.

Tabela 5: Alvos preditos para os elementos transponíveis IS3 ssg IS150 e TnXo19.

Espécie	Isolado	Alvo	Locus tag	Posição-mRNA	Posição-sRNA
<i>R. solanacearum</i>	GMI1000	IS3ssgrIS150	rs_rs11380 rs_rs25525 rs_rs02895	51 -- 122	1 -- 71
<i>X. campestris</i>	CN17	<i>avrBs3</i>	xcccn17_rs1113	10 -- 45	24 -- 60

DISCUSSÃO:

Doenças em plantas causadas por bactérias fitopatogênicas estão associadas a grandes prejuízos e perdas em culturas comerciais (Strange & Scott, 2005; Nicolopoulou et al., 2016). Numerosos estudos apontam que a rápida disseminação desses patógenos, dificuldades de manejo e resistência a defensivos químicos e biológicos são os principais desafios encontrados para o controle adequado (Sundin et al., 2016; Marin et al., 2019). Os principais gêneros de bactérias que afetam culturas comerciais de grande importância codificam determinantes de resistências ou adquiriram genes de resistência a antimicrobianos via transferência horizontal de genes (HGT) (Sundin & Bender, 1995; Förster et al., 2015). Ademais, em seus genomas existem evidências que eventos de recombinação, deleções, inserções e duplicações que interferem na expressão, favorecendo a patogênese (Sundin & Bender, 1995; Mira & Andersson, 2002; Castiblanco et al., 2013). Em nosso trabalho, investigamos pela primeira vez o perfil de distribuição de ISs e Tns presentes nas principais bactérias fitopatogênicas de interesse agrônomico e científico. Além da predição desses elementos, a pesquisa também buscou entender o papel dessas sequências na evolução do genoma e no patossistema dessas bactérias.

Nossos resultados mostraram que a quantidade de elementos presentes entre as espécies analisada variou consideravelmente. De forma interessante, observamos que a abundância de cópias de ETs varia em função do estilo de vida do hospedeiro. Bactérias que possuem uma longa história evolutiva ou uma estreita dependência com seus alvos como seus vetores, tendem a ter um menor acúmulo de elementos móveis em seus genomas. Por exemplo, *E. amylovora*, *X. fastidiosa* e os isolados dos gêneros *Dickeya* e *Pectobacterium*, além de serem os organismos que apresentaram uma baixa abundância de ETs, possuem em comum associações com maior

especialização aos hospedeiros. Por exemplo, *E. amylovora* possui uma baixa diversidade genética e existem sinais claros de adaptações nos genomas em resposta ao hospedeiro (Mann et al., 2013). Essas variações foram detectadas na composição de exopolissacarídeos, na perda de genes relacionados a adaptação em ambientes livres e em genes que codificam a proteínas efetoras como Eop1, que está envolvida na especificidade do hospedeiro (Sebaihia et al., 2010; Asselin et al., 2011; Piqué et al., 2015). A bactéria *X. fastidiosa* também é outro patógeno que possui um número pequeno de ISs em seus genomas e parece ter evoluído devido a obrigatoriedade da dependência de seu vetor de transmissão, uma vez que é exclusivamente transmitida apenas por insetos da ordem Hemiptera que se alimentam da seiva do xilema (Mann et al., 2013). Do mesmo modo, bactérias que causam os sintomas da podridão mole, como as dos gêneros *Dickeya* e *Pectobacterium*, também possuem associações com insetos, algumas acidentais, porém outras relativamente íntimas, como é o caso da interação entre *P. atrosepticum* e *Drosophila melanogaster* (Brewer et al., 1981; Mariano et al., 2007). Portanto, esses resultados mostram que a coevolução entre bactéria e hospedeiro ou a presença de um vetor para infecção pode resultar em um menor número de elementos transponíveis. Nesses casos, sem a necessidade imediata de ajustes genômicos para a sobrevivência do patógeno, a seleção natural parece agir contra a plasticidade do genoma, uma vez que algumas funções não precisam mais ser executadas pelo organismo. Assim, trabalhos semelhantes concluem que nesses organismos as deleções genéticas são muito mais frequentes do que as aquisições (Moran & Plague, 2004; Moya et al., 2008; Siguier & Chandler, 2014).

No entanto, bactérias que conseguem se replicar no ambiente sem a dependência de um hospedeiro, geralmente possuem uma fase epifítica como *P. syringae*, *A. tumefaciens*, *R. solanacearum* e as espécies do gênero *Xanthomonas*. Essas bactérias apresentaram um número significativo de ETs variando de 60 a 70 por genoma, exceto em *A. tumefaciens* que possui em média 10 elementos por genoma e *Xanthomonas oryzae* entre 300 a 400 elementos. Essa abundância pode ser explicada, pois, bactérias presentes em diferentes nichos como a filosfera e a rizosfera precisam desenvolver mecanismos para superar as condições ambientais externas abióticas e bióticas. Desse modo, a geração de plasticidade a partir do ganho de genes acessórios, que confirmam funções associadas a adaptação no ambiente via transferência horizontal e a novas variações via recombinação entre genes mediados por ISs podem auxiliar na sobrevivência dessas bactérias. Um exemplo são alguns isolados de *P. syringae* e *X. campestris* que são resistentes a pesticidas a base de estreptomicina devido ao par de genes *strA/strB* localizado no transposon adquirido, Tn5393 (Tancos et al., 2016).

Xanthomonas oryzae foi a espécie que apresentou o maior número de elementos por genoma variando entre 300 a 400 elementos. Essa espécie também está presente na filosfera. No entanto, a diversidade entre as cepas pode estar de acordo com a variedade de espécies e de genes de resistência codificados pelas diferentes culturas melhoradas de cultivo de arroz. Um estudo nas Filipinas mostrou que a implantação de um gene de resistência em culturas de arroz está associada a seleção de grupos de patógenos virulentos (Quibod et al., 2020; Timilsina et al., 2020). De acordo com os pesquisadores, cada população desenvolveu diferentes mecanismos de adaptação, favorecendo seis grupos de patógenos em um curto período. Ademais, a análise da estrutura desses seis grupos mostrou a presença de ETs interrompendo genes, realizando recombinações e disseminando genes que conferem aptidão nos genomas de *X. oryzae*. Por isso, o alto número de pseudogenes nessa espécie revelou uma evolução recente, sendo um sinal claro de que a espécie está em transição.

Outra diferença às nossas observações foi *A. tumefaciens*, o patógeno apresentou dez elementos por genoma, apesar de estar presente no solo e não possuir uma especificidade em relação ao seu hospedeiro. Contudo, isso pode ser explicado devido ao seu genoma ser altamente conservado, principalmente devido ao seu principal modo de infecção bastante específico (Bourras et al., 2015). Grande parte dos microrganismos não conseguem iniciar o ciclo de infecção nas plantas principalmente devido a estratégia de reconhecimento de padrões moleculares associados a patógenos (PAMPs) (Xin et al., 2018). Nesse sentido, para superar essas estratégias que impedem a entrada da bactéria na planta, patógenos desenvolveram tipos diferentes de fatores de virulência. A patogênese de *A. tumefaciens* é totalmente dependente da presença do plasmídeo indutor de tumor, no entanto, o método de transformação mediado pela espécie é eficiente, uma vez que consegue subverter processos celulares básicos do hospedeiro, como por exemplo o reconhecimento rápido do fitopatógeno pela planta (Rivas, 2011). Por isso, nessas espécies o número de pseudogenes tende a ser menor, uma vez que ETs podem se inserir em regiões importantes provocando inativações de genes em bactérias.

De forma geral, nossos resultados mostram um perfil heterogêneo na distribuição de famílias de ISs entre os dez fitopatógenos. No total IS5 apresentou uma maior abundância de sequências nos genomas, entretanto IS3 foi a família mais amplamente distribuída, não presente em apenas uma espécie. Tanto IS5 quanto IS3 são famílias encontradas frequentemente em procariotos; no entanto IS3 é considerada a família mais abundante e amplamente distribuído em bactérias (Siguiet et al., 2015). Em contrapartida, algumas famílias como IS6 e ISAs1 foram preditas somente em determinados organismos. A família IS6 geralmente é relatada em isolados clínicos de enterobactérias e associada a genes de resistência à antibióticos (Varani et al., 2021).

Sendo uma família homogênea com elementos de aproximadamente 800 bp, essa já foi detectada no transposons Tn5393b em *X. campestris*, contudo, parece ser a primeira vez que se identifica a presença desse elemento em *X. fastidiosa*. Já a família ISAs1, presente em cópia única em *R. solanacearum*, se encontra apenas no domínio Bacteria e somente em 50 espécies bacterianas (Siguiet et al., 2015). Pouco se sabe sobre as propriedades de regulação e transposição de elementos dessa família, contudo já foram relatadas associações dessa família à genes que codificam componentes da superfície celular de seus hospedeiros (Siguiet et al., 2015).

Além das ISs, em quatro fitopatógenos foram relatados também a presença de Tns pertencentes a família Tn3, alguns carregando genes de virulência e de resistência. O transposon Tn5393, encontrado em *P. syringae* (B728a), carrega um par de genes de resistência à estreptomicina, *strA* e *strB*. Em lavouras, é comum o uso de defensivos agrícolas que possuem em sua formulação o antibiótico estreptomicina e o manejo indiscriminado desse produto químico pode selecionar fitopatógenos resistentes (Rezzonico et al., 2009). O que se torna crítico já que esses genes podem ser transferidos a outros patógenos por elementos móveis via transferência horizontal. O transposon Tn5393, por exemplo, foi relatado pela primeira vez em *E. amylovora*, contudo esse já é encontrado em isolados de bactérias fitopatogênicas como *P. syringae* e *X. campestris* (Sundin & Wang, 2018). Ademais, também identificamos dois outros transposons, TnAs2 e TnXo19, contendo sequências que codificam proteínas efetoras do sistema de secreção do tipo III, HopX e PopP2, em *P. syringae* (CFBP3846-1) e *X. campestris* (CN03), respectivamente. Proteínas efetoras do tipo T3SS estão envolvidas na patogênese, manipulando vias importantes do hospedeiro, como as de defesa. Os membros das famílias HopX, ainda não são muito bem caracterizados, entretanto, algumas pesquisas relatam a participação dessa proteína na indução da resposta de defesa por hipersensibilidade (Nimchuk et al., 2007). Já a proteína PopP2, é membro de uma importante família de efetores, a YopJ, presente em uma ampla gama de patógenos bacterianos, curiosamente em fitopatógenos o efector PopP2 havia sido relatado apenas em *R. solanacearum*, atuando como um importante inibidor da expressão de genes relacionados à defesa em plantas (Xia et al., 2021).

Outro tipo de gene efector utilizado como estratégia para a colonização em plantas foi verificado nos Tns analisados, sendo denominados como TALEs (efector do tipo ativador de transcrição), esses atuam como ativadores transcricionais em células vegetais modulando a resposta de defesa a favor do patógeno (Wu et al., 2014). Em *X. campestris* (CN17 e CN12) detectamos um transposon TnXax1 transportando um membro da família TALE, AvrBs3. Essa proteína possui como alvo o gene de resistência *Bs3* que está presente em espécies de pimenta,

desse modo regiões repetitivas de AvrBs3 se ligam em regiões promotoras do gene alvo (*Bs3*) impedindo que haja uma resposta de hipersensibilidade (Wu et al., 2014). Os TALEs já foram detectados em vários isolados de *Xanthomonas* e parece estar entre os principais mecanismos de patogênese desse gênero. Foi relatado em trabalhos anteriores que a propagação desses genes está relacionada com a transposição do TnXax1, que se encontra funcional e que no geral parece estar associado aos genes TALE como verificado em patovares de *X. citri* e *X. oryzae* (Ferreira et al., 2015).

Contudo, além dos transposons poderem ser elementos chaves para patogênese fornecendo genes determinantes para a adaptação, eles também podem ser responsáveis pela transferência de informações genéticas. Em procariotos, a transferência de Tns e ISs entre células se dá via conjugação mediada por plasmídeos, visto que esses elementos são abundantes nessa molécula extracromossômica (Malachowa & DeLeo, 2010). Nesse contexto, observamos que elementos de espécies diferentes foram agrupados em clados próximos evolutivamente, como os Tns presentes na espécie *P. syringae* se agrupando próximo ao grupo de elementos da espécie *R. solanacearum* e elementos presentes nas espécies do gênero *Xanthomonas* próximo a *R. solanacearum* e *P. syringae*. Logo, esses resultados podem representar uma possível transferência horizontal de genes entre bactérias fitopatogênicas de espécies diferentes. Esse comportamento já é relatado entre vários fitopatógenos, como já relatado anteriormente para o Tn5393 (Rezzonico et al., 2009). Ademais, esse fenômeno também foi identificado em uma análise realizada por Assis et al. (2021), em que o transposon TnXaj417, que confere resistência ao cobre, foi encontrado em diversas espécies que causam doenças em plantas como as dos gêneros *Xanthomonas*, *Xylella* e *Pseudomonas*. Curiosamente, o estudo mostrou que em cada gênero as sequências desse transposon sofreram modificações, conservando apenas os genes *copA* e *copB*, sugerindo assim que a aquisição da resistência ao cobre foi devido a uma recente HGT.

Apesar disso, a adaptação de microrganismos a novos nichos, não acontece devido apenas a aquisição de genes, ISs também possuem um importante papel na plasticidade dos genomas, por meio de inversões, translocações, deleções e inserções de DNA (Vandecraen et al., 2017). Em *R. solanacearum*, por exemplo, cepas agressivas que causam a murcha bacteriana possuem alto número de cópias de ISs, como no isolado UY031 (Erill et al., 2017). Esses elementos desempenham um papel importante na adaptação para a espécie, uma vez que a alta taxa de transposição é a principal força de diversificação desse patógeno (Erill et al., 2017). Ademais, acredita-se que a seleção de seis cepas virulentas de *X. oryzae* ocorreu a partir da adoção de variedades resistentes de arroz; entretanto, em cada população a arquitetura do

genoma é única em razão das diferentes formas de recombinação mediados também por elementos genéticos móveis (Quibod et al., 2020). Em nossos resultados, as espécies pertencentes ao gênero *Xanthomonas* foram as que apresentaram uma maior abundância em número de ETs, e como consequência nessas espécies encontramos evidências de eventos de recombinação mediados principalmente por IS3 e IS5.

No entanto, apesar de vários exemplos discutidos em que ETs podem se inserir em regiões importantes e impactar a resposta de vários processos, estudos mostram que existe uma baixa especificidade na escolha dos alvos, porém já se sabe que há alguns mecanismos sutis que medeiam a transposição (Siguier et al., 2014). Nas dez espécies analisadas, verificamos inserções próximo a genes de sistemas de transporte como bombas de efluxo ABC e proteínas de transporte relacionadas a metais pesados como cádmio, cobalto e zinco. Ademais, identificamos ISs perto de genes que codificam proteínas que atuam em sistemas de respostas a estresses, tais como de nitrogênio, oxidativo e alterações de temperatura. Além de regiões onde se localizam proteínas envolvidas no metabolismo de síntese de antimicrobianos e participação no processo de infecção e persistência. Em espécies do gênero *Pseudomonas*, a inserção do elemento ISL3 em uma via de produção de metabólitos primários e secundários, confere maior aptidão para que o organismo possa se desenvolver em diferentes ambientes (Christie-Oleza et al., 2008). Assim, uma inserção pode interferir na expressão de genes adjacentes, maximizando a expressão do gene, formando promotores híbridos, resultando no aumento da virulência (Siguier et al., 2014). No entanto, elementos ISs também podem impedir a síntese de uma proteína, nesse estudo detectamos em *X. oryzae* genes de avirulência sendo interrompidos por ISs. Os genes *avr*, são determinantes durante as interações planta-patógeno, pois podem ser reconhecidos pelas denominadas proteínas de resistência (R) da planta, que por consequência desencadeia respostas de defesas frente a infecção pelo patógeno (Leach & White, 1996; Flor, 1971). Alguns estudos apontam que a rápida adaptação patogênica em isolados de *X. oryzae* também é devido a diversificação de proteínas efetoras provocados por rearranjos e indels mediados por ISs (Salzberg et al., 2008). Desse modo, quando genes *avr* em bactérias são inativados ou modificados, o hospedeiro passa a não reconhecer essas proteínas efetoras, o que resulta em susceptibilidade da planta. Em particular, nos genomas analisados de *P. syringae* identificamos um elemento pertencente à família IS3 próximo a integrantes da superfamília de ATPases e a bombas de efluxo ABC. Para tanto, nossas análises confirmaram que no primeiro caso IS3 está sob pressão de seleção positiva, diferente da segunda situação. As proteínas AAA participam de diversas atividades celulares, como na replicação do DNA, no processo de degradação de proteínas, na expressão gênica dentre outras (Shorter & Houry,

2018). Contudo, essa família também está relacionada com o processo de regulação da transposição de alguns ETs. Um exemplo é o elemento IS21, o qual possui duas ORFs, sendo uma contendo um motivo de ligação de ATP que permite o auxílio de membros da família de proteínas AAA para a transposição em alvos definidos (Arias-Palomo & Berger, 2015). Desse modo, acreditamos que talvez essa família possa estar direcionando a transposição do elemento IS3 para alvos apropriados, por isso essa associação parece estar sendo mantida nos genomas de *P. syringae*.

Da mesma forma, os transportadores ABC exercem um papel importante para sobrevivência das bactérias. Esses transportadores estão associados ao influxo de nutrientes e o efluxo de toxinas, por consequência esse transportador também confere resistência a antimicrobianos, além de atuar como um exportador de fatores de virulência propiciando a patogênese (El-Awady et al., 2017). Sendo assim, a inativação dessa família de transportadores pode ser prejudicial ao organismo, portanto devido aos resultados, a inserção de IS3 parece inibir ou diminuir a atividade desse transportador, desse modo provavelmente a inserção de IS3 próximos a esses genes está passando por um processo de eliminação. No mais, a influência exercida por ETs não é exclusivamente dependente da distância e esses elementos podem interferir na expressão de outros genes mesmo não estando próximos (100 pb), codificando por exemplo pequenos RNAs reguladores (Vanderpool & Gottesman, 2004). Nos genomas de *R. solanacearum*, verificamos que o elemento IS3ssgrIS51 se encontra localizado na mesma posição em grande parte dos genomas analisados, não está próximo a qualquer gene, mas está sob pressão de seleção positiva. Neste contexto, ao longo do tempo, regiões codificadoras de ETs podem ser modificadas perdendo assim a função de transposição, mas passando a contribuir com outras funções relevantes para o hospedeiro, um processo que chamamos de domesticação (Kapitonov et al., 2016).

Além de desempenhar um papel importante na estrutura dos genomas, os ETs também participam da modulação da resposta do hospedeiro durante situações de estresse. Contudo, pouco se sabe sobre o efeito desses elementos em fitopatógenos, portanto utilizamos dados de RNA-seq de bactérias com o objetivo de analisar o comportamento das famílias de sequências de inserção em condições desfavoráveis. Observamos que em *P. syringae* a mudança de temperatura, 18 °C e 27°C, interferiu na expressão de alguns elementos, principalmente à 27 °C onde grande parte dos elementos tiveram sua transcrição aumentada. Estudos mostraram que em temperaturas quentes entre 28-30°C, há um efeito negativo nos isolados de *P. syringae* uma vez que suprime a transcrição de fitotoxinas e de proteínas efetoras do tipo T3SS, afetando assim a virulência do patógeno (Xin et al, 2018). Ademais, quando isolados de *X. oryzae* foram

submetidos a um meio de cultivo com um indutor de patogenicidade, elementos transponíveis próximos a transportadores da família MFS e á pectato-liase, aumentaram sua expressão.

Há evidências de que ETs costumam estar próximos de genes que codificam fatores de patogenicidade. Esses podem interferir na modulação da expressão de genes adjacentes, o que tem sido implicado na natureza adaptativa de fitopatógenos (Stapley et al, 2015). Da mesma forma, quando inseridos em contextos ao estresse oxidativo e nitrosativo, percebemos que a maioria dos elementos tiveram sua expressão aumentada quando expostos a diferentes condições de óxido nítrico. A mudança das concentrações de NO alterou um dos principais fatores de virulência em *R. solanacearum*, as proteínas efetoras T3SS (Hendrich et al., 2021). Nos isolados dessa espécie analisados por Gonçalves et al., 2020, a maioria dos elementos ISs foram encontrados inseridos dentro de genes que codificam efetores de secreção do tipo III. Desse modo, a interferência de vias metabólicas associadas ao nitrogênio parece interferir na expressão de genes de virulência em patógenos. Observamos também que ocorreu a diminuição da transcrição dos elementos adjacentes aos genes que codificam a porina e a proteína PcaQ, que também tiveram sua transcrição suprimida. A porina é um transportador presente na membrana de bactérias gram-negativas, que possui um importante papel na seleção da permeabilidade de nutrientes, toxinas e antimicrobianos (Vandecraen et al., 2017). Logo, a diminuição na transcrição desta proteína pode levar o fitopatógeno a adquirir temporariamente resistência a determinados compostos. O envolvimento de ISs com esse transportador, já foi relatado nas espécies do gênero *Pseudomonas*, onde a inativação devido a transposição de ISs aumenta a resistência a antimicrobianos em situações adversas (Vandecraen et al., 2017). Já a proteína PcaQ é um membro da família LysR que atua como regulador transcricional de diversos processos, como a resposta ao estresse oxidativo (MacLean et al., 2011; Lee & Lee, 2020). Um estudo realizado com espécies do gênero *Deinococcus* mostrou que a atividade de transposição de ISs na espécie é regulada quando ela se encontra sob estresse oxidativo. Ademais, no mesmo trabalho bactérias mutantes para os membros da família LysR tiveram modificações em determinadas famílias de ISs, onde ocorreu a diminuição da atividade de transposição (Lee & Lee, 2020). Portanto, é possível perceber que há uma modificação na expressão ISs quando o hospedeiro se encontra sob algum tipo de estresse. Observamos também que elementos ISs influenciam genes adjacentes que muitas vezes estão associados a patogenicidade ou à condição de estresse estudada. Desta forma entendemos que esses elementos parecem ter um papel importante na sobrevivência do hospedeiro e participação durante as interações planta-patógeno.

Por fim, as famílias que tiveram uma expressão diferencial nas diferentes condições ambientais anteriores, foram analisadas a fim de identificar possíveis sRNAs que poderiam estar sendo codificados. Pequenos RNAs reguladores codificados por elementos transponíveis já foram identificados tanto na forma Cis_sRNA controlando a transposição do próprio elemento, como também do tipo trans regulando outros genes (Bobrovskyy & Vanderpool, 2013; Ellis et al., 2017). Nós identificamos dois elementos, IS3ssgrIS150 pertencente a *R. solanacearum* e TnXo19 pertencente a *X. campestris* que parecem estar codificando sRNAs putativos. No caso do elemento IS3ssgrIS150, o pequeno RNA parece ser sintetizado no mesmo locus da IS se configurando como um sRNA-antisense. Na literatura, grande parte dos sRNAs relatados que são codificados por ETs são do tipo Cis_sRNA, e esse mecanismo atua na regulação da transposição, controlando assim o número de cópias dos elementos no genoma (Kreth & Merritt, 2019). Por outro lado, o sRNA codificado por TnXo19 presente no genoma de *X. campestris* parece ter o um papel biológico diferente. De acordo com nossos dados a transposase referente a família Tn3 é responsável pela transcrição do sRNA, que possui como alvo uma proteína efetora, AvrBs3, pertencente a uma família de ativadores transcricionais de plantas denominado TALEs. Assim, podemos propor que o sRNA é capaz de regular a transcrição do gene *avrBs3* durante a interação planta-patógeno, maximizando a virulência de *X. campestris*, uma vez que essas proteínas se ligam a promotores de genes de resistência da planta resultando em suscetibilidade do hospedeiro.

Neste trabalho, foi realizada a maior e mais completa predição de elementos transponíveis no genoma de bactérias fitopatogênicas. A abundância desses elementos no genoma parece estar relacionada ao tipo o mecanismo de infecção e ao nível especialização entre patógeno-hospedeiro. O envolvimento dos ETs em eventos de recombinação ectópicas, sua expressão em condições de estresse e o seu envolvimento na regulação vis sRNAs foram investigados, demonstrando que esses elementos podem desempenhar papéis na variabilidade genética e conseqüentemente na patogênese. Desse modo, acreditamos que a compreensão dos mecanismos de variabilidade de patógenos, mediados por elementos transponíveis, poderá futuramente auxiliar em novas abordagens para o controle de doenças em plantas.

CONSIDERAÇÕES FINAIS:

Os resultados deste trabalho forneceram evidências que ETs, como ISs e Tns, podem estar atuando como um dos principais mecanismos geradores de variabilidade genética nos genomas de fitopatógenos, além de atuarem como mediadores dos mecanismos de patogênese e

adaptação das principais bactérias fitopatogênicas de interesse agrônomo e científico. Diante disso, as razões que explicam e confirmam essa afirmação são:

A abundância de elementos ISs em espécies que infectam um número significativo de plantas hospedeiras ou que estão passando por um processo de seleção;

Presença de ETs, como transposons carregando genes que conferem resistência a antimicrobianos e virulência como proteínas efectoras;

Sequências de inserção sendo selecionadas positivamente, em posições adjacentes a genes referentes aos mecanismos de patogênese e adaptação;

Diferença do perfil de expressão dos elementos ISs quando submetidos a condições de estresse, o que pode ser um mecanismo de variabilidade capaz de auxiliar na adaptação dos patógenos em condições não favoráveis;

Transcrição de pequenos RNAs reguladores codificados por transposons, que parecem atuar em mecanismos que causam a suscetibilidade do hospedeiro.

REFERÊNCIAS:

ALCOCK, Brian P. et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. **Nucleic acids research**, v. 48, n. D1, p. D517-D525, 2020.

ARIAS-PALOMO, Ernesto; BERGER, James M. An atypical AAA+ ATPase assembly controls efficient transposition through DNA remodeling and transposase recruitment. **Cell**, v. 162, n. 4, p. 860-871, 2015.

ASSELIN, J. E. et al. Eop1 from a Rubus strain of Erwinia amylovora functions as a host-range limiting factor. **Phytopathology**, v. 101, n. 8, p. 935-944, 2011.

ASSIS, Renata AB et al. A comparative genomic analysis of Xanthomonas arboricola pv. juglandis strains reveal hallmarks of mobile genetic elements in the adaptation and accelerated evolution of virulence. **Genomics**, v. 113, n. 4, p. 2513-2525, 2021.

BABAKHANI, Sajad; OLOOMI, Mana. Transposons: the agents of antibiotic resistance in bacteria. **Journal of basic microbiology**, v. 58, n. 11, p. 905-917, 2018.

BARDAJI, Leire et al. Sequence and role in virulence of the three plasmid complement of the model tumor-inducing bacterium Pseudomonas savastanoi pv. savastanoi NCPPB 3335. **PLoS One**, v. 6, n. 10, p. e25705, 2011.

BOLGER, Anthony M.; LOHSE, Marc; USADEL, Bjoern. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114-2120, 2014.

- BORTOLAIA, Valeria et al. ResFinder 4.0 for predictions of phenotypes from genotypes. **Journal of Antimicrobial Chemotherapy**, v. 75, n. 12, p. 3491-3500, 2020.
- BOURRAS, Salim; ROUXEL, Thierry; MEYER, Michel. Agrobacterium tumefaciens gene transfer: how a plant pathogen hacks the nuclei of plant and nonplant organisms. **Phytopathology**, v. 105, n. 10, p. 1288-1301, 2015.
- BREWER, J. W.; HARRISON, M. D.; WINSTON, J. A. Survival of two varieties of *Erwinia carotovora* on *Drosophila melanogaster* meigen and *Drosophila busckii* Coquillett, (Diptera: Drosophilidae) vectors of potato blackleg in Colorado. **American Potato Journal**, v. 58, n. 9, p. 439-449, 1981.
- BOBROVSKYY, Maksym; VANDERPOOL, Carin K. Regulation of bacterial metabolism by small RNAs using diverse mechanisms. **Annual review of genetics**, v. 47, p. 209-232, 2013.
- BUTLER, Margi I. et al. *Pseudomonas syringae* pv. *actinidiae* from recent outbreaks of kiwifruit bacterial canker belong to different clones that originated in China. **PLoS one**, v. 8, n. 2, p. e57464, 2013.
- CASTIBLANCO, Luisa F. et al. TALE 1 from *Xanthomonas axonopodis* pv. *manihotis* acts as a transcriptional activator in plant cells and is important for pathogenicity in cassava plants. **Molecular plant pathology**, v. 14, n. 1, p. 84-95, 2013.
- CHEN, Lihong et al. VFDB: a reference database for bacterial virulence factors. **Nucleic acids research**, v. 33, n. suppl_1, p. D325-D328, 2005.
- CHIOU, C. S.; JONES, A. L. Nucleotide sequence analysis of a transposon (Tn5393) carrying streptomycin resistance genes in *Erwinia amylovora* and other gram-negative bacteria. **Journal of bacteriology**, v. 175, n. 3, p. 732-740, 1993.
- CHRISTIE-OLEZA, Joseph Alexander et al. ISPst9, an ISL3-like insertion sequence from *Pseudomonas stutzeri* AN10 involved in catabolic gene inactivation. 2008.
- CROS, Marie-Josée et al. RNAspace. org: an integrated environment for the prediction, annotation, and analysis of ncRNA. **Rna**, v. 17, n. 11, p. 1947-1956, 2011.
- DARLING, Aaron CE et al. Mauve: multiple alignment of conserved genomic sequence with rearrangements. **Genome research**, v. 14, n. 7, p. 1394-1403, 2004.
- EL-AWADY, Raafat et al. The role of eukaryotic and prokaryotic ABC transporter family in failure of chemotherapy. **Frontiers in pharmacology**, v. 7, p. 535, 2017.
- ELLIS, Michael J. et al. A transposon-derived small RNA regulates gene expression in *Salmonella Typhimurium*. **Nucleic acids research**, v. 45, n. 9, p. 5470-5486, 2017.
- ERILL, Ivan et al. Comparative analysis of *Ralstonia solanacearum* methylomes. **Frontiers in plant science**, v. 8, p. 504, 2017.
- FERREIRA, Rafael Marini et al. A TALE of transposition: Tn 3-like transposons play a major role in the spread of pathogenicity determinants of *Xanthomonas citri* and other xanthomonads. **MBio**, v. 6, n. 1, p. e02505-14, 2015.
- FÖRSTER, Helga et al. Characterization of streptomycin resistance in isolates of *Erwinia amylovora* in California. **Phytopathology**, v. 105, n. 10, p. 1302-1310, 2015.

- FLOR, Harold H. Current status of the gene-for-gene concept. **Annual review of phytopathology**, v. 9, n. 1, p. 275-296, 1971.
- GONÇALVES, Osiel Silva et al. Transposable elements contribute to the genome plasticity of *Ralstonia solanacearum* species complex. **Microbial genomics**, v. 6, n. 5, 2020.
- GUO, Yufang. The fate of European olives. **Nature Food**, v. 1, n. 5, p. 255-255, 2020.
- GRUBER, Andreas R. et al. The vienna RNA websuite. **Nucleic acids research**, v. 36, n. suppl_2, p. W70-W74, 2008.
- GRUBER, Andreas R. et al. RNAz 2.0: improved noncoding RNA detection. In: **Biocomputing 2010**. 2010. p. 69-79.
- HENDRICH, Connor G. et al. Nitric Oxide Regulates the *Ralstonia solanacearum* Type 3 Secretion System. **bioRxiv**, p. 2020.10.26.355339, 2021.
- HIROSE, Karin; ISHIGA, Yasuhiro; FUJIKAWA, Takashi. Phytotoxin synthesis genes and type III effector genes of *Pseudomonas syringae* pv. *actinidiae* biovar 6 are regulated by culture conditions. **PeerJ**, v. 8, p. e9697, 2020.
- KANNAN, Velu Rajesh; BASTAS, Kubilay Kurtulus; DEVI, Rajendran Sangeetha. 20 Scientific and Economic Impact of Plant Pathogenic Bacteria. **Sustainable approaches to controlling plant pathogenic bacteria**, p. 369, 2015.
- KAPITONOV, Vladimir V.; MAKAROVA, Kira S.; KOONIN, Eugene V. ISC, a novel group of bacterial and archaeal DNA transposons that encode Cas9 homologs. **Journal of bacteriology**, v. 198, n. 5, p. 797-807, 2015.
- KHAN, Madiha et al. Oh, the places they'll go! A survey of phytopathogen effectors and their host targets. **The Plant Journal**, v. 93, n. 4, p. 651-663, 2018.
- KIM, Seunghwan et al. Time-resolved pathogenic gene expression analysis of the plant pathogen *Xanthomonas oryzae* pv. *oryzae*. **BMC genomics**, v. 17, n. 1, p. 1-15, 2016.
- KRETH, J.; BRANTL, S.; MERRITT, Justin. Regulatory RNAs. In: **Encyclopedia of Microbiology**. Elsevier, 2019. p. 62-84.
- KUMAR, Sudhir et al. MEGA X: molecular evolutionary genetics analysis across computing platforms. **Molecular biology and evolution**, v. 35, n. 6, p. 1547, 2018.
- KUMAR, Krishna; CHAKRABORTY, Abhijit; CHAKRABARTI, Saikat. PresRAT: a server for identification of bacterial small-RNA sequences and their targets with probable binding region. **RNA biology**, v. 18, n. 8, p. 1152-1159, 2021.
- LANGMEAD, Ben; SALZBERG, Steven L. Fast gapped-read alignment with Bowtie 2. **Nature methods**, v. 9, n. 4, p. 357-359, 2012.
- LEACH, Jan E.; WHITE, Frank F. Bacterial avirulence genes. **Annual review of phytopathology**, v. 34, n. 1, p. 153-179, 1996.
- LEE, Chanjae; CHOO, Kyungsil; LEE, Sung-Jae. Active transposition of insertion sequences by oxidative stress in *Deinococcus geothermalis*. **Frontiers in microbiology**, p. 2665, 2020.

- LOTT, Steffen C. et al. GLASSgo—automated and reliable detection of sRNA homologs from a single input sequence. **Frontiers in genetics**, v. 9, p. 124, 2018.
- MANN, Rachel A. et al. Comparative genomics of 12 strains of *Erwinia amylovora* identifies a pan-genome with a large conserved core. **PloS one**, v. 8, n. 2, p. e55644, 2013.
- MALACHOWA, Natalia; DELEO, Frank R. Mobile genetic elements of *Staphylococcus aureus*. **Cellular and molecular life sciences**, v. 67, n. 18, p. 3057-3071, 2010.
- MANSFIELD, John et al. Top 10 plant pathogenic bacteria in molecular plant pathology. **Molecular plant pathology**, v. 13, n. 6, p. 614-629, 2012.
- MARIN, Vítor Rodrigues et al. Recent advances in the biocontrol of *Xanthomonas* spp. **World Journal of Microbiology and Biotechnology**, v. 35, n. 5, p. 1-11, 2019.
- MARIANO, R. L. R. et al. Transmissão de fitobacterioses por insetos. **Anais da Academia Pernambucana de Ciência Agrônômica**, p. 4,214-239, 2007.
- MARTIN, Darren P. et al. RDP3: a flexible and fast computer program for analyzing recombination. **Bioinformatics**, v. 26, n. 19, p. 2462-2463, 2010.
- MACLEAN, Allyson M. et al. The LysR-type PcaQ protein regulates expression of a protocatechuate-inducible ABC-type transport system in *Sinorhizobium meliloti*. **Microbiology**, v. 157, n. 9, p. 2522-2533, 2011.
- MARTINS, Paula MM et al. Persistence in phytopathogenic bacteria: do we know enough?. **Frontiers in microbiology**, v. 9, p. 1099, 2018.
- MAT RAZALI, Nurhani; CHEAH, Boon Huat; NADARAJAH, Kalaivani. Transposable elements adaptive role in genome plasticity, pathogenicity and evolution in fungal phytopathogens. **International journal of molecular sciences**, v. 20, n. 14, p. 3597, 2019.
- MIRA, Alex; KLASSON, Lisa; ANDERSSON, Siv GE. Microbial genome evolution: sources of variability. **Current opinion in microbiology**, v. 5, n. 5, p. 506-512, 2002.
- MORAN, Nancy A.; PLAGUE, Gordon R. Genomic changes following host restriction in bacteria. **Current opinion in genetics & development**, v. 14, n. 6, p. 627-633, 2004.
- MOYA, Andrés et al. Learning how to live together: genomic insights into prokaryote–animal symbioses. **Nature Reviews Genetics**, v. 9, n. 3, p. 218-229, 2008.
- NICOLOPOULOU-STAMATI, Polyxeni et al. Chemical pesticides and human health: the urgent need for a new concept in agriculture. **Frontiers in public health**, v. 4, p. 148, 2016.
- NIMCHUK, Zachary L. et al. The HopX (AvrPphE) family of *Pseudomonas syringae* type III effectors require a catalytic triad and a novel N-terminal domain for function. **Molecular plant-microbe interactions**, v. 20, n. 4, p. 346-357, 2007.
- NIU, Xiang-Na et al. Complete sequence and detailed analysis of the first indigenous plasmid from *Xanthomonas oryzae* pv. *oryzicola*. **BMC microbiology**, v. 15, n. 1, p. 1-15, 2015.
- NAWROCKI, Eric P.; EDDY, Sean R. Infernal 1.1: 100-fold faster RNA homology searches. **Bioinformatics**, v. 29, n. 22, p. 2933-2935, 2013.

- OLIVEIRA, Amanda CP et al. Transposons and pathogenicity in *Xanthomonas*: acquisition of murein lytic transglycosylases by TnXax1 enhances *Xanthomonas citri* subsp. *citri* 306 virulence and fitness. *PeerJ*, v. 6, p. e6111, 2018.
- PARTRIDGE, Sally R. et al. Mobile genetic elements associated with antimicrobial resistance. ***Clinical microbiology reviews***, v. 31, n. 4, p. e00088-17, 2018.
- PETERS, Joseph E. et al. Recruitment of CRISPR-Cas systems by Tn7-like transposons. ***Proceedings of the National Academy of Sciences***, v. 114, n. 35, p. E7358-E7366, 2017.
- PFEIFFER, Friedhelm et al. The complete and fully assembled genome sequence of *Aeromonas salmonicida* subsp. *pectinolytica* and its comparative analysis with other *Aeromonas* species: investigation of the mobilome in environmental and pathogenic strains. ***BMC genomics***, v. 19, n. 1, p. 1-15, 2018.
- PFEILMEIER, Sebastian; CALY, Delphine L.; MALONE, Jacob G. Bacterial pathogenesis of plants: future challenges from a microbial perspective: challenges in bacterial molecular plant pathology. ***Molecular plant pathology***, v. 17, n. 8, p. 1298-1313, 2016.
- PIQUÉ, Núria et al. Virulence factors of *Erwinia amylovora*: a review. ***International journal of molecular sciences***, v. 16, n. 6, p. 12836-12854, 2015.
- QUIBOD, Ian Lorenzo et al. The Green Revolution shaped the population structure of the rice pathogen *Xanthomonas oryzae* pv. *oryzae*. ***The ISME journal***, v. 14, n. 2, p. 492-505, 2020.
- RAHMAN, Raza-Ur et al. Oasis 2: improved online analysis of small RNA-seq data. ***BMC bioinformatics***, v. 19, n. 1, p. 1-10, 2018.
- RIVAS, Susana; GENIN, Stéphane. A plethora of virulence strategies hidden behind nuclear targeting of microbial effectors. ***Frontiers in plant science***, v. 2, p. 104, 2011.
- ROBINSON, David G.; LEE, Ming-Chun; MARX, Christopher J. OASIS: an automated program for global investigation of bacterial and archaeal insertion sequences. ***Nucleic acids research***, v. 40, n. 22, p. e174-e174, 2012.
- ROSS, Karen et al. Tncentral: a prokaryotic transposable element database and web portal for transposon analysis. ***MBio***, v. 12, n. 5, p. e02060-21, 2021.
- ROZAS, Julio et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. ***Molecular biology and evolution***, v. 34, n. 12, p. 3299-3302, 2017.
- SALZBERG, Steven L. et al. Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *oryzae* PXO99A. ***BMC genomics***, v. 9, n. 1, p. 1-16, 2008.
- SEBAIHIA, M. et al. Complete genome sequence of the plant pathogen *Erwinia amylovora* strain ATCC 49946. ***Journal of bacteriology***, v. 192, n. 7, p. 2020-2021, 2010.
- SERRATO-CAPUCHINA, Antonio; MATUTE, Daniel R. The role of transposable elements in speciation. ***Genes***, v. 9, n. 5, p. 254, 2018.
- SHORTER, James; HOURY, Walid A. the role of AAA+ proteins in protein repair and degradation. ***Frontiers in Molecular Biosciences***, p. 85, 2018.

- SIGUIER, Patricia et al. ISfinder: the reference centre for bacterial insertion sequences. **Nucleic acids research**, v. 34, n. suppl_1, p. D32-D36, 2006.
- SIGUIER, Patricia; GOURBEYRE, Edith; CHANDLER, Mick. Bacterial insertion sequences: their genomic impact and diversity. **FEMS microbiology reviews**, v. 38, n. 5, p. 865-891, 2014.
- SIGUIER, Patricia et al. Everyman's guide to bacterial insertion sequences. **Mobile DNA iii**, p. 555-590, 2015.
- STAPLEY, Jessica; SANTURE, Anna W.; DENNIS, Stuart R. Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. **Molecular ecology**, v. 24, n. 9, p. 2241-2252, 2015.
- STRANGE, Richard N.; SCOTT, Peter R. Plant disease: a threat to global food security. **Annu. Rev. Phytopathol.**, v. 43, p. 83-116, 2005.
- SUNDIN, George W.; BENDER, Carol L. Expression of the strA-strB streptomycin resistance genes in *Pseudomonas syringae* and *Xanthomonas campestris* and characterization of IS6100 in *X. campestris*. **Applied and environmental microbiology**, v. 61, n. 8, p. 2891-2897, 1995.
- SUNDIN, George W. et al. Bacterial disease management: challenges, experience, innovation and future prospects: challenges in bacterial molecular plant pathology. **Molecular plant pathology**, v. 17, n. 9, p. 1506-1518, 2016.
- TAJIMA, Fumio. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. **Genetics**, v. 123, n. 3, p. 585-595, 1989.
- THOMPSON, Julie D.; HIGGINS, Desmond G.; GIBSON, Toby J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic acids research**, v. 22, n. 22, p. 4673-4680, 1994.
- TIMILSINA, Sujana et al. *Xanthomonas* diversity, virulence and plant-pathogen interactions. **Nature Reviews Microbiology**, v. 18, n. 8, p. 415-427, 2020.
- TOUCHON, Marie; ROCHA, Eduardo PC. Causes of insertion sequences abundance in prokaryotic genomes. **Molecular biology and evolution**, v. 24, n. 4, p. 969-981, 2007.
- URBAN, Martin et al. PHI-base: a new interface and further additions for the multi-species pathogen-host interactions database. **Nucleic acids research**, v. 45, n. D1, p. D604-D610, 2017.
- VANDECRAEN, Joachim et al. The impact of insertion sequences on bacterial genome plasticity and adaptability. **Critical reviews in microbiology**, v. 43, n. 6, p. 709-730, 2017.
- VANDERPOOL, Carin K.; GOTTESMAN, Susan. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. **Molecular microbiology**, v. 54, n. 4, p. 1076-1089, 2004.
- VARANI, Alessandro M. et al. ISSaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. **Genome biology**, v. 12, n. 3, p. 1-9, 2011.

VARANI, Alessandro et al. The IS6 family, a clinically important group of insertion sequences including IS26. **Mobile Dna**, v. 12, n. 1, p. 1-18, 2021.

XIA, Yao et al. Secondary-structure switch regulates the substrate binding of a YopJ family acetyltransferase. **Nature communications**, v. 12, n. 1, p. 1-10, 2021.

XIN, Xiu-Fang; KVITKO, Brian; HE, Sheng Yang. Pseudomonas syringae: what it takes to be a pathogen. **Nature Reviews Microbiology**, v. 16, n. 5, p. 316-328, 2018.

WANG, Yejun et al. T3DB: an integrated database for bacterial type III secretion system. **BMC bioinformatics**, v. 13, n. 1, p. 1-6, 2012.

WRIGHT, Patrick R. et al. CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. **Nucleic acids research**, v. 42, n. W1, p. W119-W123, 2014.

WU, Liang et al. Go in for the kill: How plants deploy effector-triggered immunity to combat pathogens. **Virulence**, v. 5, n. 7, p. 710-721, 2014.