

FRANCISCO HENRIQUE NUNES DA SILVA ALVES

**GENÔMICA COMPARATIVA E POTENCIAIS MECANISMOS DE
PATOGENICIDADE DE *Pseudomonas* QUE INFECTAM O CAFEEIRO EM MINAS
GERAIS, BRASIL**

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

Orientador: Jorge Luis Badel Pacheco
Coorientador: Laércio Zambolim

**VIÇOSA - MINAS GERAIS
2019**

Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa

T

A474g
2019

Alves, Francisco Henrique Nunes da Silva, 1991-
Genômica comparativa e potenciais mecanismos de
patogenicidade de *Pseudomonas* que infectam o cafeeiro em
Minas Gerais, Brasil / Francisco Henrique Nunes da Silva Alves.
– Viçosa, MG, 2019.

122 f. : il. (algumas color.) ; 29 cm.

Inclui anexos.

Orientador: Jorge Luis Badel Pacheco.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. *Pseudomonas cichorii*. 2. *Pseudomonas syringae* pv.
garcae. 3. *Pseudomonas syringae* pv. *tabaci*. I. Universidade
Federal de Viçosa. Departamento de Fitopatologia. Programa de
Pós-Graduação em Fitopatologia. II. Título.

CDD 22. ed. 579.332

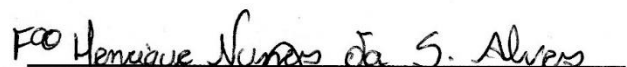
FRANCISCO HENRIQUE NUNES DA SILVA ALVES

**GENÔMICA COMPARATIVA E POTENCIAIS MECANISMOS DE
PATOGENICIDADE DE *Pseudomonas* QUE INFECTAM O CAFEIEIRO EM MINAS
GERAIS, BRASIL**

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

APROVADA: 22 de julho de 2019.

Assentimento:


Francisco Henrique Nunes da Silva Alves

Autor



Jorge Luis Badel Pacheco
Orientador

Dedico esta dissertação a Deus, à minha família, minha namorada Sara e a todos meus amigos que me apoiaram nessa jornada.

AGRADECIMENTOS

A Deus, por ser meu alicerce, minha força e por ter me dado tantas oportunidades e aprendizados.

À Universidade Federal de Viçosa, pela excelência em ensino.

Ao Departamento de Fitopatologia da UFV pela grande oportunidade de aprender com os melhores professores.

À Coordenação de Aperfeiçoamento de Pessoal do Nível Superior (CAPES) por ter me proporcionado uma bolsa de estudos de mestrado.

Ao meu orientador Jorge Luis Badel, por ter acreditado em meu potencial para conclusão da pesquisa, oferecendo uma grande oportunidade de crescimento profissional e pessoal. Pelas excelentes orientações em todos os momentos que solicitei sua ajuda. Por ser um exemplo de orientador e ter compartilhado um pouco do seu conhecimento comigo.

Ao meu coorientador Laércio Zambolim, pelo incentivo na pesquisa.

À minha mãe, por sempre ter me incentivado e ter me apoiado em todas as minhas decisões importantes.

À minha tia Eliana, que considero como mãe. Por ser meu exemplo de superação, humildade e perseverança. Que sempre me inspirou com todo seu carinho, amor e dedicação. Por ter tornado possível a minha vinda para Viçosa e sempre ter apoiado meus sonhos.

À minha avó, Bernarda, pela educação, criação, amor e por ser exemplo de força e superação.

À minha namorada Sara, por sempre permanecer me apoiando e me incentivando a seguir meus sonhos. Por toda paciência nos momentos de estresse. Por ser minha família em Viçosa e fonte de minha felicidade.

Aos meus amigos Fernanda, Ítalo, Manel, Silvio e Letícia, por todo apoio e inúmeros momentos de felicidade.

Às minhas amigas do laboratório de fitobacteriologia molecular, Adryelle Anchieta, Ana e Patrícia Matias, pelas colaborações e compartilhamento de ideias.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

RESUMO

ALVES, Francisco Henrique Nunes da Silva., Universidade Federal de Viçosa, julho de 2019. **Genômica comparativa e potenciais mecanismos de patogenicidade de *Pseudomonas* que infectam o cafeeiro em Minas Gerais, Brasil.** Orientador: Jorge Luis Badel Pacheco. Coorientador: Laércio Zambolim.

As manchas bacterianas causadas por *Pseudomonas syringae* pv. *garcae* (*Psgc*), *P. syringae* pv. *tabaci* (*Psta*) e *Pseudomonas cichorii* (*Pch*) são umas das principais doenças bacterianas que comprometem a produção de café no Brasil. Para entender a diversidade genética e os mecanismos subjacentes à patogenicidade destas espécies/patovares bacterianas, este estudo primeiro (Capítulo 1) demonstrou as suas organizações genômicas, encontrou regiões de sintonia e diferenças nas sequências, com marcada ocorrência de transposições, deleções/inserções e inversões. Comparação entre os proteomas das espécies/patovares, revelou 581 proteínas exclusivas de *Psgc*, 834 exclusivas de *Psta*, 1486 exclusivas de *Pch* e 3424 comuns entre os três patógenos bacterianos. Análise do pangenoma e árvores filogenéticas, baseadas no conteúdo gênico do pangenoma e SNP do genoma *core*, determinou que as espécies/patovares bacterianas possuem plasticidade genômica. Mediante análise filogenética baseada em identidade média de nucleotídeos (ANIb) as espécies/patovares foram separadas em distintos clados. Assim, os resultados deste estudo indicaram que as bactérias patogênicas ao cafeeiro apresentam alta diversidade genética. Na segunda parte (Capítulo 2), uma análise comparativa dos sistemas de secreção e efetores do tipo III desses fitopatógenos foi realizada, usando como referência o *cluster* de genes *hrp/hrc* da bactéria modelo *P. syringae* pv. *tomato* DC3000 (*Pst* DC3000). Encontrou-se que os genes *hrpA*, *hrpD*, *hrpF*, *hrG*, *hrcQa* e *hrpJ* de *Psgc* não possuíram identidade de sequência quando comparados com os genes do *cluster* de *Pst* DC3000, enquanto os demais genes identificados possuíram identidade de sequência que variaram de 26 a 98%. Em *Psta* os genes *hrcJ*, *hrpG*, *hrcS*, *hrpP*, *hrpQ* e *hrpJ* não obtiveram similaridade com os genes do *cluster* de *Pst* DC3000, enquanto os demais genes identificados possuíram identidades de sequência que variam de 41 a 94%. Em *Pch*, poucos genes mostraram identidade de sequência com os genes do *cluster* de *Pst* DC3000, variando de 27 a 58%. Comparações entre os pansecretomas do tipo III das diferentes espécies/patovares mostrou que de 46 famílias de efetores identificados em *Psgc*, as famílias *avrB*,

avrPto, *hopY*, *hopZ*, *hopAI*, *hopAU*, *hopBI*, *hopAJ*, *hopH*, *hopAF*, *hopBF* e 10 possíveis novos efetores foram exclusivos. Em *Psta*, das 43 famílias de efetores identificados, somente foram exclusivos *hopI*, *hopM*, *hopQ*, *hopAB*, *hopAZ* e 13 novos candidatos à efetores. Em *Pch*, *hopA*, *hopB*, *hopBN* e 10 possíveis novos efetores foram exclusivos. Foi observado que *avrE* foi a única família de efetores compartilhada entre as três espécies/patovares bacterianas. Assim, os resultados indicam que o efector *avrE* pode possuir papel fundamental nas interações dessas bactérias com o cafeeiro, e a descoberta de novos efetores pode ajudar a entender como cada espécie/patovar se adaptou ao hospedeiro. Por último (Capítulo 3) pretendendo encontrar os mecanismos de adaptação comuns entre as *Pseudomonas* patogênicas ao cafeeiro, identificou-se as proteínas secretadas pelo sistema de secreção do tipo II (T2SS) e potenciais efetores do tipo II (T2SE). Foi identificada a presença de 368 proteínas não redundantes que possuíam sinal de secreção de tipo II e ausência de domínio transmembranar em *Psgc*, 420 em *Psta* e 459 em *Pch*. Comparando os três pansecretomas do T2SS, foi observado que 240 proteínas estiveram presentes em todas as espécies/patovares analisadas. No entanto, muitas proteínas tiveram anotações relacionadas com funções associadas a membranas, pilus ou flagelo. Filtragem das proteínas associadas com essas estruturas bacterianas revelou T2SE candidatos com atividades enzimáticas potenciais, que ainda não têm sido relatadas como importantes para a patogenicidade bacteriana. Este trabalho proporciona resultados que servem como ponto de partida para pesquisas futuras, visando obter conhecimento sobre a diversidade genética e os mecanismos moleculares que conferem a capacidade de causar doença no cafeeiro aos patógenos *Psgc*, *Psta* e *Pch*. Esses resultados podem também ser úteis para desenvolver novas estratégias que visem controlar os patógenos bacterianos que causam manchas foliares em plantas de café no futuro.

Palavras-chave: *Pseudomonas cichorii*. *Pseudomonas syringae* pv. *garcae*. *Pseudomonas syringae* pv. *tabaci*.

ABSTRACT

ALVES, Francisco Henrique Nunes da Silva, Universidade Federal de Viçosa, July, 2019. **Comparative genomics and potential pathogenicity mechanisms of *Pseudomonas* infecting coffee in Minas Gerais, Brazil.** Adviser: Jorge Luis Badel Pacheco. Co-adviser: Laércio Zambolim.

Bacterial spots caused by *Pseudomonas syringae* pv. *garcae* (*Psgc*), *P. syringae* pv. *tabaci* (*Psta*) and *Pseudomonas cichorii* (*Pch*) are some of the major bacterial diseases that compromise coffee production in Brazil. In order to understand the genetic diversity and the molecular mechanisms underlying the pathogenicity of these bacterial species/pathovars, this study first (Chapter 1) showed their genomic organizations, found syntenic regions and sequence differences, with striking occurrence of transpositions, deletions/insertions and inversions. Comparisons among species/pathovar proteomes, revealed that 581 proteins are exclusive to *Psgc*, 834 exclusive to *Psta*, 1486 exclusive to *Pch* and 3424 common to all three bacterial pathogens. Pangenomic analysis and phylogenetic trees, based on gene content of the pangenome and SNP in the core genome, demonstrated that the species/pathovars exhibit genomic plasticity. Phylogenetic analysis based on average nucleotide identity (ANI_b) separated the species/pathovars in different clades. Hence, the results of this study indicated that the bacteria pathogenic to coffee show high genetic diversity. In the second part (Chapter 2), a comparative analysis of the type III secretion system and effectors of these three phytopathogens was conducted using the *hrp/hrc* gene cluster of the model bacterium *P. syringae* pv. *tomato* DC3000 (*Pst* DC3000) as a reference. It was found that genes *hrpA*, *hrpD*, *hrpF*, *hrG*, *hrcQa* and *hrpJ* from *Psgc* did not show sequence identity when compared with genes of the *Pst* DC3000 cluster, whereas the rest of the genes identified showed sequence identity varying from 26 to 98%. In *Psta*, the genes *hrcJ*, *hrpG*, *hrcS*, *hrpP*, *hrpQ* and *hrpJ* did not exhibit sequence similarity with genes of the *Pst* DC3000 cluster, whereas the other genes identified had sequence identity varying from 41 to 94%. In *Pch*, only a few genes showed sequence identity with the genes of the *Pst* DC3000 cluster, varying from 27 to 58%. Comparisons among type III pansecretomes of the different species/pathovars showed that out of 46 effector families identified in *Psgc*, the families *avrB*, *avrPto*, *hopY*, *hopZ*, *hopAl*, *hopAU*, *hopBl*, *hopAJ*, *hopH*, *hopAF*, *hopBF* and 10 potential new effectors were exclusive. In *Psta*, out of 43 effector families identified,

only *hopI*, *hopM*, *hopQ*, *hopAB*, *hopAZ* and 13 new potential effector candidates were exclusive. In *Pch*, *hopA*, *hopB*, *hopBN* and 10 possible new effectors were exclusive. It was observed that *avrE* was the only effector family shared among the three bacterial species/pathovars. Hence, the results indicate that the *avrE* effector could play a fundamental role in the interactions of these bacteria with coffee plants, and that the discovery of new effectors could aid in understanding how each species/pathovar adapted to the host. Lastly (Chapter 3), aiming to find common mechanisms of adaptation among the *Pseudomonas* pathogenic to coffee plants, proteins secreted by the type II secretion system (T2SS) and putative type II effectors (T2SE) were identified. The presence of 368 non-redundant proteins that carried a type II secretion signal and lacked transmembrane domains were identified in *Psgc*, 420 in *Psta* and 459 in *Pch*. Comparing the three T2SS pansecretomes, it was observed that 240 proteins were present in all species/pathovars analyzed. However, many proteins were annotated with functions associated with the membrane, pilus or flagellum. Filtering out proteins associated with these bacterial structures revealed T2SE candidates with putative enzymatic activities that have not yet been reported as important for bacterial pathogenicity. This work provides results that constitute a starting point for future research aimed to obtain knowledge on the genetic diversity and the molecular mechanisms that confer the capacity to cause disease on coffee plants to the pathogens *Psgc*, *Psta* and *Pch*. These results could also be useful to develop novel strategies to control the bacterial pathogens that cause foliar spots on coffee plants in the future.

Keywords: *Pseudomonas cichorii*. *Pseudomonas syringae* pv. *garcae*. *Pseudomonas syringae* pv. *tabaci*.

SUMÁRIO

INTRODUÇÃO GERAL	11
REFERÊNCIAS	14
Capítulo 1: Variabilidade e plasticidade genômica de <i>Pseudomonas</i> patogênicas ao cafeeiro em Minas Gerais, Brasil	16
INTRODUÇÃO	17
MATERIAL E MÉTODOS	19
Obtenção das sequências genômicas	19
Ordenamento e alinhamento de genomas rascunho	20
Predição da similaridade entre proteomas	20
Pangenômica e análise de divergência	20
Análise da identidade média de nucleotídeos (ANI)	21
RESULTADOS	21
Genomas de bactérias que infectam o cafeeiro apresentam transposições, deleções/inserções e inversões	21
<i>Pseudomonas cichorii</i> apresenta o proteoma mais divergente	22
Bactérias fitopatogênicas ao cafeeiro apresentam plasticidade genômica	22
Patovares de <i>P. syringae</i> que infectam plantas de café são separados em distintos clados	24
DISCUSSÃO	24
REFERÊNCIAS	27
FIGURAS	32
MATERIAL SUPLEMENTAR	40
Capítulo 2: Divergência e conservação gênica dos sistemas de secreção e efetores de tipo III em <i>Pseudomonas</i> patogênicas ao cafeeiro de Minas Gerais	44
RESUMO	44
INTRODUÇÃO	45
MATERIAL E MÉTODOS	47
Obtenção das sequências genômicas	47
Análise comparativa do <i>cluster</i> de genes <i>hrp/hrc</i>	47
Predição e genômica comparativa de complementos de T3SE	48
RESULTADOS	49
<i>Pseudomonas cichorii</i> possui um <i>hrp/hrc cluster</i> divergente dos patovares de <i>Pseudomonas syringae</i>	49
Diversidade de efetores novos e conhecidos entre <i>Pseudomonas</i> que infectam o cafeeiro	50

Um efector altamente conservado e importante para a patogenicidade é comum entre bactérias patogênicas ao cafeeiro.....	51
DISCUSSÃO.....	52
REFERÊNCIAS.....	57
FIGURAS.....	64
TABELAS.....	70
MATERIAL SUPLEMENTAR.....	73
Capítulo 3: Efectores de tipo II comuns e específicos de <i>Pseudomonas</i> patogênicas ao cafeeiro em Minas Gerais.....	96
RESUMO.....	96
INTRODUÇÃO.....	97
MATERIAL E MÉTODOS.....	98
Obtenção das sequências das proteínas.....	99
Predição e proteômica comparativa das proteínas secretadas pelo T2SS.....	99
RESULTADOS.....	100
Pansecretomas do tipo II.....	100
<i>P. cichorii</i> possui o maior número de proteínas específicas, associadas ao T2SS....	100
Identificação das proteínas dos secretomas do T2SS.....	101
DISCUSSÃO.....	102
REFERÊNCIAS.....	106
FIGURAS.....	109
TABELAS.....	117

INTRODUÇÃO GERAL

O café (*Coffea* spp.) é um dos produtos agrícolas mais consumidos no mundo, e o Brasil possui papel de destaque como maior produtor mundial, seguido pelo Vietnã e pela Colômbia (ICO, 2018). Em 2019, em decorrência da incidência de altas temperaturas, atrelado à escassez de chuvas no período importante do ciclo (principalmente em dezembro de 2018 e janeiro de 2019), observou-se a queda da expectativa de sacas beneficiadas no Brasil, estimada em 50,9 milhões, uma produção menor que a obtida em 2018 (CONAB 2019).

Cultivado nas mais distintas condições de clima, solo, altitude e topografia, o café brasileiro está sujeito aos mais diversos fatores bióticos, que incluem doenças causadas por bactérias fitopatogênicas (Rodrigues *et al.*, 2013). Dentre os patógenos bacterianos, *Pseudomonas syringae* pv. *garcae* (*Psgc*) (Amaral *et al.*, 1956), *Pseudomonas syringae* pv. *tabaci* (*Psta*) (Rodrigues Neto *et al.*, 2006) e *Pseudomonas cichorii* (*Pch*) (Swingle 1925) se destacam dentre os principais agentes causais de doenças da cultura, sendo fatores limitantes para o cultivo do café em regiões mais frias e com fortes ventos, sobretudo em regiões acima de 1000 m de altitude. Esses patógenos causam perdas estimadas de 70% da produção, tanto em condições de viveiro quanto de campo, principalmente nos estados do Paraná, São Paulo e no sul de Minas Gerais (Rodrigues *et al.*, 2013; Kimura *et al.*, 1973)

Além da dificuldade de distinguir morfologicamente as manchas foliares causadas por esses três patógenos no cafeeiro, a determinação do agente etiológico também é dificultada por causa da proximidade filogenética dessas espécies bacterianas, tornando desafiadoras as suas diagnoses. Os métodos atuais de detecção têm se baseado principalmente no isolamento e identificação dos patógenos, usando testes morfológicos, fisiológicos e bioquímicos (Belan *et al.*, 2016; Hinkosa *et al.*, 2016; Rodrigues *et al.*, 2017b; Zoccoli *et al.*, 2011). Apesar destas dificuldades, a distinção entre *Psgc* e outros patovares de *Pseudomonas syringae* é possível pelos testes de LOPAT (Rodrigues *et al.*, 2017a). Métodos sorológicos e baseados na reação em cadeia da polimerase (*Polymerase Chain Reaction*, PCR) têm sido desenvolvidos somente para detecção de estirpes de *Pch* que infectam outras espécies de plantas hospedeiras (Sánchez *et al.*, 2003). Um método baseado na

amplificação da região *hrpRST* se mostrou eficiente na detecção de *Pch* em amostras de cafeeiro infectadas pelo patógeno (Raimundi, 2017).

O conhecimento atual da diversidade genética e patogênica das espécies bacterianas que causam manchas foliar no cafeeiro é escasso. Estudos iniciais com *Psgc* demonstraram diferenças na produção de pigmentos fluorescentes em meio King's B e de bacteriocinas (Kairu *et al.*, 1997; Mwangi *et al.*, 2018). Mediante análise de sequências de 16S rRNA, foi possível agrupar as estirpes de *Psgc* com base na agressividade, embora não tenha sido possível ter uma correlação com base na origem geográfica, enquanto, utilizando Rep-PCR, o contrário foi observado (Maciel *et al.*, 2018; Mwangi *et al.*, 2018). No entanto, trabalhos de variabilidade genética precisam ser conduzidos com populações ainda maiores e com estirpes advindas de diferentes origens geográficas, com o objetivo de garantir conclusões mais claras. Quanto ao Brasil não existem relatos sobre a diversidade genética de *Psgc*.

Apesar de *Psta* e *Pch* terem ampla gama de hospedeiros, sendo capazes de infectar outras espécies de plantas de interesse econômico e que pertencem a diferentes famílias botânicas (Beriam *et al.*, 2006; Malavolta Júnior *et al.*, 2008), não existem relatos sobre a sua diversidade genética. Resultados de um único estudo com estirpes que infectam outros hospedeiros sugerem que *Psta* tem uma alta variabilidade (Peng *et al.*, 2003).

O advento de tecnologias de sequenciamento de alto desempenho possibilitou o desenvolvimento de novas estratégias para a elucidação da variabilidade em espécies bacterianas. A genômica comparativa pode ser feita em níveis de genes, transcritos, proteínas ou mesmo vias bioquímicas. A partir da sequência genômica, é possível entender os processos de evolução de uma determinada espécie mediante o alinhamento de genes ou do genoma completo com aqueles de outras espécies. Um dos objetivos da genômica comparativa é obter informações que possam responder qual a motivação de uma espécie possuir determinada característica e a presença ou ausência dessa mesma característica em outros organismos, por exemplo, a capacidade de causar doença em um determinado hospedeiro (Costa, 2015). Desse modo a genômica comparativa permite uma melhor compreensão da biologia das espécies e seus processos evolutivos, conhecimento este que aumenta a acurácia da classificação taxonômica e da identificação de patógenos (Pauling *et al.*, 2012).

Neste estudo, estirpes de *Psgc*, *Psta* e *Pch* oriundas do estado de Minas Gerais, um dos maiores produtores de café do Brasil, foram analisadas por meio de comparações genômicas para compreender sua diversidade genética. Biologia computacional foi utilizada para determinar o pangenoma e o genoma *core* de cada um desses patógenos e entender a sua variabilidade tanto nas sequências nucleotídicas quanto no ganho/perda de genes. Adicionalmente, a predição de proteínas potencialmente envolvidas na patogenicidade (efetores candidatos) foi predita, com o intuito de entender a adaptação desses patógenos para causar doença em plantas de café. Em particular, foram preditos efetores candidatos secretados pelos sistemas de secreção do tipo II (T2SS) e tipo III (T3SS).

Os resultados deste trabalho são apresentados em três capítulos, sendo o primeiro dedicado à diversidade genética, o segundo ao sistema de secreção de tipo III e o terceiro ao sistema de secreção de tipo II. Esses resultados são importantes para o avanço na procura de entender como *Psgc*, *Psta* e *Pch* causam doença no cafeeiro e desenvolver novas estratégias de controle das manchas foliares causadas por eles.

REFERÊNCIAS

Amaral JF, Do Teixeira C, Pinheiro ED, 1956. A bactéria causadora da mancha aureolada do cafeeiro. *Arquivos do Instituto Biológico* **23**, 151-155.

Belan LL, Pozza EA, Freitas ML, Raimundi MK, Souza RM, Machado JC, 2016. Occurrence of *Pseudomonas syringae* pv. *garcae* in coffee seeds. *Australasian Journal of Crop Science* **10**, 1015-1021.

Beriam, LOS, Almeida IMG, Suzete AL, Grabert E, Balani DM, Ferreira M, Rodrigues Neto J, 2006. *Pseudomonas syringae* pv. *tabaci* in papaya seedlings. *Summa Phytopathologica* **32**, 21-26.

Costa DR, 2015. Genômica comparativa de linhagens de *Saccharomyces* e *Kluyveromyces* de interesse biotecnológico. Tese de Doutorado. Universidade Federal de Viçosa.

Hinkosa GG, Lencho A, Selvaraj T, Chavhan A, Sadessa K, 2016. Screening, cultural and biochemical characterization of coffee bacterial blight (*Pseudomonas syringae* pv. *garcae* van Hall) isolates from Sidama and Gedeo zones, SNNP Regional State, Ethiopia. *International Journal of Life Sciences* **4**, 491–508.

ICO, 2018 Total production by all exporting countries. London.

Kairu GM, 1997. Biochemical and pathogenic differences between Kenyan and Brazilian isolates of *Pseudomonas syringae* pv. *garcae*. *Plant Pathology* **46**, 239–246.

Kimura O, Robbs CF, Ribeiro RLD, 1973. Estudos sobre o agente da “mancha aureolada do cafeeiro” (*Pseudomonas garcae* Amaral et al.). *Arquivos da Universidade Federal Rural do Rio de Janeiro* **3**, 15–18.

Maciel KW, Destefano SAL, Beriam LOS, Almeida IMG, Patrício FRA, Rodrigues LMR, Guerreiro Filho O, 2018. Bacterial halo blight of coffee crop: aggressiveness and genetic diversity of strains. *Bragantia* **77**, 96–106.

Malavolta Júnior VA, Almeida IMG, Rodrigues Neto J, Beriam LOS, 2008. Bactérias fitopatogênicas assinaladas no Brasil: uma atualização. *Summa Phytopathologica* **34**, 988.

Mwangi IJ, Gichuru EK, Masanga J, Alakonya AE, 2018. A comparative analysis associated to virulence of *Pseudomonas syringae* pv. *garcae*, the causative agent of bacterial blight of coffee in Kenya. *Journal of Scientific Agriculture* **2**, 76–84.

Pauling J, Röttger R, Neuner A, Salgado H, Collado-Vides J, Kalaghatgi P, Azevedo V, Tauch A, Pühler A, Baumbach J, 2012. On the trail of EHEC/EAEC—unraveling the gene regulatory networks of human pathogenic *Escherichia coli* bacteria. *Integrative Biology* **4**, 728–733.

Peng R, Zhang S, Wang S, 2003. The research progress on physiological race of *Pseudomonas syringae* pv. *tabaci*. *Journal of the Yunnan Agricultural University* **18**, 198-202.

Raimundi MK, 2017 Identificação bioquímica, molecular e patogenicidade de isolados de *Pseudomonas* spp. provenientes de cafeeiros em Minas Gerais. Tese de Doutorado. Lavras, Brazil: Universidade Federal de Lavras.

Rodrigues LMR, Almeida IMG, Patricio FRA, 2013. Mancha Aureolada do Cafeeiro Causada por *Pseudomonas syringae* pv. *garcae*. Campinas, SP, Brasil: Instituto Agrônômico de Campinas.

Rodrigues LMR, Almeida IMG, Patrício FRA, Beriam LOS, Maciel KW, Braghini MT, Guerreiro Filho O, 2017a. Aggressiveness of strains and inoculation methods for resistance assessment to Bacterial Halo Blight on coffee seedlings. *Journal of Phytopathology* **165**, 105–114.

Rodrigues LMR, Sera GH, Guerreiro Filho OG, Beriam LOS, Almeida IMG, 2017b. First report of mixed infection by *Pseudomonas syringae* pathovars *garcae* and *tabaci* on coffee plantations. *Bragantia* **76**, 543-549.

Rodrigues Neto J, Silva CHD, Beriam LOS, 2006 Mancha bacteriana do cafeeiro causada por *Pseudomonas syringae* pv *tabaci*. *Summa Phytopathologica* **32**, 85.

Sánchez L, Zapata M, Rodríguez RP, 2003. Distribución y patogenicidad de *Pseudomonas cichorii* (Swingle) Stapp en el cultivo de café em Puerto Rico. *Journal of Agriculture of the University of Puerto Rico* **87**,123–135.

Zoccoli DM, Takatsu A, Uesugi CH, 2011. Ocorrência de mancha aureolada em cafeeiros na região do Triângulo Mineiro e Alto Paranaíba. *Bragantia* **70**, 843–849.

Capítulo 1: Variabilidade e plasticidade genômica de *Pseudomonas* patogênicas ao cafeeiro em Minas Gerais, Brasil

Francisco Henrique Nunes da Silva Alves¹, Jorge Luis Badel¹

¹ Laboratório de Fitoacteriologia Molecular, Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG, Brasil.

Autor para correspondência: Jorge Luis Badel, e-mail: jorge.badel@ufv.br

RESUMO

Pseudomonas syringae pv. *garcae* (*Psgc*), *Pseudomonas syringae* pv. *tabaci* (*Psta*) e *Pseudomonas cichorii* (*Pch*), se destacam causando grande impacto econômico na cultura do café, embora as suas diversidades genéticas tenham sido pouco investigadas. O objetivo deste trabalho foi estudar a diversidade genética das espécies/patovares que infectam o cafeeiro mediante comparações genômicas entre distintas estirpes. Para tanto, as sequências genômicas de *Psgc*, *Psta* e *Pch*, obtidas do banco de dados (GenBank), foram alinhadas usando o software Mauve, os proteomas foram analisados usando CD-Hit, o pangenoma e o genoma *core* foram determinados usando a ferramenta PGAWeb e a identidade média de nucleotídeos (ANI) foi calculada usando o programa JSpeciesWS. O alinhamento das sequências possibilitou notar certa sintenia global entre as estirpes da mesma espécie/patovar, embora com a presença de inversões em algumas partes do genoma, e baixa sintenia entre *Pch* e os patovares de *P. syringae*. Mediante o estudo dos proteomas, foram identificadas 581 proteínas não redundantes exclusivas de *Psgc*, 834 de *Psta* e 1486 de *Pch*, além de 3424 proteínas, não redundantes, comuns entre os três patógenos bacterianos. A análise do pangenoma e da evolução determinou que as espécies/patovares apresentam plasticidade genômica, indicando possuírem genomas abertos. Através do ANI foi verificado que as três *Pseudomonas* patogênicas ao cafeeiro são separadas em diferentes clados. Nosso estudo serve de princípio para desenvolver estudos voltados a entender os mecanismos envolvidos na adaptação destas bactérias para infectar o cafeeiro.

Palavras-chave: *cichorii*, *Coffea*, *garcae*, genômica comparativa, mancha foliar, *tabaci*.

INTRODUÇÃO

O café é umas das commodities mais vendidas no mundo. Entre as mais de 100 espécies do gênero *Coffea*, apenas o café arábica (*Coffea arábica* L.) e o café robusta (*Coffea canephora* L.) são economicamente importantes (DaMatta & Ramalho 2004). Cinco espécies bacterianas, *Robbsia andropogonis* (Lopes-Santos et al., 2017) [sinônimo *Burkholderia andropogonis* (Smith, 1911)], *Pseudomonas syringae* (van Hall, 1902), *Pseudomonas cichorii* (Swingle 1925) e *Xylella fastidiosa* (Wells et al., 1987), têm sido reportadas como agentes fitopatogênicos de plantas de café sob condições de campo e de viveiro. Dentro de *P. syringae*, os patovares *garcae* (Amaral et al., 1956) e *tabaci* (Rodrigues Neto et al., 2006) foram relatados causando manchas foliares em plantas de café (Badel & Zambolim, 2019).

A mancha aureolada causada por *Pseudomonas syringae* pv. *garcae* (*Psgc*), foi detectada inicialmente em 1955, em cafezais localizados no município de Garça, estado de São Paulo, afetando folhas, ramos e frutos (Amaral et al., 1956). A mancha foliar bacteriana causada por *P. syringae* pv. *tabaci* (*Psta*) foi relatada somente no Brasil, principalmente afetando plantas em viveiro (Rodrigues Neto et al., 2006; 2017a; Destéfano et al., 2010; Raimundi, 2017). O crestamento foliar bacteriano causado por *P. cichorii* (*Pch*) já foi relatado no Brasil (Robbs et al., 1974; Kimura et al. 1976; Raimundi, 2017) e em diferentes localidades de Porto Rico (Sánchez et al., 2003).

Em decorrência das semelhanças dos sintomas causados por *Psgc*, *Psta* e *Pch*, atrelado à presença destes patógenos em áreas próximas e ocorrências de infecções mistas, a importância das doenças causadas por *Psta* e *Pch* podem ter sido subestimadas por erros de diagnose (Rodrigues et al., 2013; Rodrigues et al., 2017b). Embora esses patógenos causem grande impacto econômico na cultura do cafeeiro, a variabilidade genética e a patogenicidade dessas três espécies/patovares bacterianas têm sido pouco investigadas. Os poucos estudos realizados foram baseados principalmente em rep-PCR e em variações na sequência da região 16S rDNA, utilizando um número muito reduzido de estirpes, gerando resultados inconclusivos (Badel & Zambolim 2019, Maciel et al., 2018; Mwangi et al., 2018)

Atualmente, com o desenvolvimento das tecnologias de sequenciamento de alto rendimento e o aumento no número de genomas disponíveis, os estudos de genômica comparativa têm ganhado cada vez mais espaço, sendo decisivos para a compreensão dos mecanismos evolutivos e moleculares da patogenicidade bacteriana (Monteil *et al.*, 2016). Esses genomas disponíveis muitas vezes são rascunhos, compostos por centenas de *contigs*. Com a intenção de facilitar as análises comparativas, arranjar ou orientar esses *contigs* em sequências maiores (*scaffolds*) é necessário. Assim, *contigs* podem ser ordenados por comparação com um genoma completo (que não apresente *gaps*) de referência e que possua uma organização notoriamente estabelecida (Chen & Lu 2018). Softwares como o Mauve Aligner, que identifica e alinha regiões homólogas, é capaz de ordenar os *contigs* de uma espécie/patovar de acordo com um genoma de referência, fornecendo uma ótima base para estudos de genômica comparativa e de dinâmica evolutiva (Darling *et al.*, 2010; Rissman *et al.*, 2009).

A análise de pangenoma tem crescido significativamente em estudos relacionados à bactérias, sendo aplicado em vários campos de pesquisa, como patogenicidade, resistência, evolução, adaptação e estrutura populacional (Chen *et al.*, 2018; Tian *et al.*, 2016; Vernikos *et al.*, 2015; Xiao *et al.*, 2015; Rasko *et al.*, 2008). Para isto, programas que determinam as porções comuns (genoma *core*) e variáveis (genomas acessório e específico) são utilizados, incluindo o PGAWeb (Chen *et al.*, 2018), cuja vantagem, comparado com outros programas disponíveis, é poder ser utilizado com genomas rascunho. Além do pangenoma, a análise de evolução, gerada pelo PGAWeb, apresenta diversas árvores filogenéticas, que são geradas baseadas em três tipos de algoritmos (*Neighbor-Joining* (NJ), Médias Aritméticas não Ponderadas (UPGMA; *Unweighted Pair Group Method with Arithmetic mean*) e Máxima Verossimilhança (ML; *Maximum Likelihood*)) usando o robusto pacote de análise filogenética PHYLIP (Felsenstein, 1989), baseada em dois tipos de dados: matriz de ganho/perda de genes (PanBased) ou SNP (SNPBased) dos genes pertencentes ao genoma *core*, sendo o algoritmo de máxima verossimilhança usado apenas na análise com base nos SNPs do genoma *core* (Chen *et al.*, 2018).

Desde o final de 1960, quando a hibridização DNA-DNA (DDH) foi introduzida para verificar ou aperfeiçoar os agrupamentos de organismos, através de similaridades genômicas, a classificação entre procariotos é baseada em

similaridades de sequências de nucleotídeos (Ciuffo *et al.*, 2018; Wayne, 1988). Agora que as sequências genômicas estão facilmente disponíveis, os métodos comparativos diretos podem obter resultados análogos às técnicas como DDH (Konstantinidis & Tiedje, 2005). Assim, a identidade média de nucleotídeos (ANI), calculada a partir de comparações pareadas de todas as sequências compartilhadas entre quaisquer duas estirpes, é amplamente utilizada e tem sido proposta como a nova medida para definição e classificação de espécies bacterianas, podendo ser aplicada em grandes conjuntos de dados. Valores de ANI $\geq 95\%$ estão fortemente correlacionados com o valor de 70% estabelecido pela técnica de DDH para definição de espécie bacteriana (Jain *et al.*, 2018; Richter *et al.*, 2015; Varghese *et al.*, 2015; Rashid *et al.*, 2015; Chan *et al.*, 2012; Richter & Rosselón-Móra, 2009; Goris *et al.*, 2007). No entanto, estudos utilizando valores de ANI para estabelecer a posição taxonômica das bactérias que infectam ao cafeeiro não foram relatados.

Assim, o objetivo deste trabalho foi realizar uma análise de genômica comparativa, entre as espécies/patovares de *Pseudomonas* patogênicas ao cafeeiro, a fim de identificar as características comuns e específicas de cada patógeno e entender a diversidade destas espécies/patovares.

MATERIAL E MÉTODOS

Obtenção das sequências genômicas

As sequências genômicas das diferentes estirpes de *P. syringae* pv. *garcae*, *P. syringae* pv. *tabaci* e *P. cichorii*, associadas com plantas de café, foram obtidas do banco de dados (GenBank) do *National Center for Biotechnology Information* (NCBI; <http://www.ncbi.nlm.nih.gov>). Sequências curadas e depositadas como RefSeq foram utilizadas nas análises e suas informações estão contidas nas Tabelas Suplementares S1 e S2.

Ordenamento e alinhamento de genomas rascunho

O ordenamento dos *contigs* e alinhamento das sequências de *Psgc*, *Psta* e *Pch* foi executado utilizando o software Mauve Aligner v.2.4.0 (Rissman *et al.*, 2009). Devido que na atualidade não existem genomas completos de referência de *Psgc*, *Psta* e *Pch*, os *contigs* dos seus genomas rascunho retirados do Genbank foram ordenados utilizando como referência o genoma completo da estirpe bacteriana modelo *Pseudomonas syringae* pv. *tomato* DC3000 (número de acesso do NCBI ASMN780v1). Em decorrência da limitação do número de sequências disponíveis no GenBank, cinco sequências genômicas de *Psgc*, *Psta* e *Pch*, cujas montagens foram as de melhor qualidade, foram selecionadas para realizar as análises de genômica comparativa (Tabela Suplementar S1).

Predição da similaridade entre proteomas

Utilizando o servidor web CD-HIT Suite (http://weizhonglab.ucsd.edu/cdhit_suite/cgi-bin/index.cgi?cmd=cd-hit) (Huang *et al.* 2010), foram analisados os proteomas das estirpes de *Psgc*, *Psta* e *Pch*, selecionadas no ordenamento. Primeiro, os proteomas das cinco estirpes de cada patógeno bacteriano foram submetidos à ferramenta CD-HIT para eliminar proteínas redundantes, usando um limiar de 50% de identidade, e foi gerado um panproteoma para cada espécie/patovar. Após a retirada das repetições entre as cinco estirpes de cada espécie/patovar, foi gerado um único arquivo contendo todos os panproteomas não redundantes, agrupados em uma única base de dados. Essa base de dados foi submetida ao CD-HIT, com objetivo de determinar as proteínas comuns entre as espécies/patovares estudados. Os resultados obtidos foram organizados em um diagrama de Venn usando a ferramenta Jvenn (<http://jvenn.toulouse.inra.fr/app/example.html>) (Bardou *et al.*, 2014).

Pangenômica e análise de divergência

Para o estudo de pangenômica e de divergência das espécies/patovares associados ao cafeeiro, foram selecionadas trinta sequências genômicas rascunho de *Psgc* e cinco de *Pch* e *Psta* (Tabela Suplementar S1). A análise foi executada usando a ferramenta PGAWeb (<http://pgaweb.vlcc.cn/>) (Chen *et al.*, 2018).

Análise da identidade média de nucleotídeos (ANI)

Para verificar a posição taxonômica das estirpes bacterianas recuperadas do cafeeiro, os valores de ANIb foram calculados usando os genes conservados com o algoritmo JSpeciesWS (<http://jspecies.ribohost.com/jspeciesws/#home>) (Richter *et al.* 2015). Valores de ANIb foram utilizados para construir as árvores filogenéticas usando o programa PHYLIP v.3.695 empregando o algoritmo *Neighbor-Joining* (NJ) (Felsenstein, 1989). Para isso, os valores de ANIb foram convertidos em coeficientes de dissimilaridade (D) usando a equação: $D = 100 - ANIb$. As árvores foram editadas usando NJPLOT v.2.3 (Perrière & Gouy, 1996).

RESULTADOS

Genomas de bactérias que infectam o cafeeiro apresentam transposições, deleções/inserções e inversões

O alinhamento das sequências genômicas evidenciou uma sintonia global entre as estirpes da mesma espécie/patovar, embora algumas sequências apresentem transposições, deleções/inserções e inversões (Figura Suplementar S1). Esses alinhamentos revelaram que o genoma de *Psgc* aparenta ser mais estável do que os genomas dos outros dois patógenos do cafeeiro. Por sua vez, o genoma de *Pch* apresentou o maior número de rearranjos (transposições, deleções/inserções). Nesses alinhamentos, segmentos conservados do genoma, representados como blocos colineares locais (*Locally Colinear Blocks*; LCB), que são ligados por uma linha da mesma cor e que mudam de posição entre as espécies na Figura Suplementar S1, refletiram nas várias regiões de rearranjo.

Quando as sequências de melhor qualidade de cada espécie/patovar incluídas no estudo, previamente ordenadas com referência a *Pst* DC3000, foram alinhadas juntas, foi possível visualizar uma maior conservação genômica em relação a alguns LCBs entre as estirpes pertencentes à mesma espécie, *Psgc* e *Psta*, embora com a presença de inversões em algumas partes do genoma (Figura 1). A sequência genômica de *Pch* apresentou menor conservação genômica com relação àquelas dos patovares de *P. syringae*.

***Pseudomonas cichorii* apresenta o proteoma mais divergente**

Na tentativa de reforçar os resultados obtidos com o alinhamento em Mauve, comparações entre os proteomas dos mesmos patógenos bacterianos foram conduzidos utilizando CD-HIT (Figura 2). O programa identificou o número de proteínas que são exclusivas de cada espécie/patovar e aquelas compartilhadas entre eles. Sendo assim, 581 proteínas foram exclusivas de *Psgc*, 834 de *Psta* e 1486 de *Pch*. Esses resultados também indicam que 3424 são comuns entre os três patógenos bacterianos estudados. Encontrou-se que um número maior de proteínas é compartilhado entre os patovares de *P. syringae* do que em comparações com *Pch*.

Bactérias fitopatogênicas ao cafeeiro apresentam plasticidade genômica

Com a finalidade de estabelecer o número de genes compartilhados entre as estirpes (genoma *core*) da mesma espécie/patovar, foi realizado uma análise de pangenoma (Figura 3). Para as 30 sequências genômicas analisadas de *Psgc*, foi observado que, conforme a adição de sequências individuais, ocorre um decréscimo gradativo no tamanho do genoma *core* (Figura 3A). Entretanto, também foi possível observar o crescimento contínuo na curva de pangenoma (somatório dos genes *core*, acessórios e específicos). Tendências semelhantes foram observadas no pangenoma e no genoma *core* de *Psta* e *Pch* (Figura 3B e 3C). Quanto a *Psta* (Figura 3B), embora a curva de pangenoma cresça de maneira menos acentuada quando comparada com

aquelas de *Psgc* e *Pch* (Figura 3A e 3C), ainda ocorre o aumento no tamanho do pangenoma.

As árvores filogenéticas de *Psgc* geradas, baseadas na matriz de genes PanBased, utilizando UPGMA, mostrou que quatro estirpes (ASMN341584v1, ASMN341574v1, ASMN341632v1 e ASMN341634v1) são filogeneticamente mais distantes do resto, formando quatro clados distintos. As demais estirpes foram separadas em dois clados, um menor contendo cinco estirpes e o outro com 21. Variabilidade genômica adicional entre estirpes dentro destes dois clados foi observada (Figura 4A). Usando o método NJ, as estirpes ASMN341605v1 e ASMN341568 mostraram-se altamente divergentes do restante do grupo, formando clados independentes. Já os demais foram separados em seis clados diferentes, cada um com número variável de estirpes (Figura 4B). Quanto a análise filogenética baseada nos SNP do genoma *core*, as três estirpes ASMN341605v1, ASMN341568v1 e ASMN341603v1, se mostraram as mais divergentes, formando clados distintos. As outras 27 estirpes foram distribuídas em vários clados com diferentes tamanhos (Figura 5).

Com o intuito de obter informação sobre a variabilidade genômica de *Psta* e *Pch*, apesar do número reduzido de estirpes, as árvores filogenéticas geradas pelo PGAWeb também foram analisadas. Com relação a *Psta*, a árvore baseada na matriz PanBased usando o método UPGMA demonstrou a formação de dois grupos, o primeiro constituído apenas por ASMN341575v1, e o segundo grupo, que possui mais duas ramificações, com o restante das estirpes (Figura 6A). Já na árvore usando a matriz PanBased com o método NJ, formaram-se três grupos, sendo o primeiro constituído por ASMN341614v1, o segundo, que possui três estirpes (ASMN341616v1, ASMN341658v1 e ASMN341578v1) e o último formado por ASMN341575v1 (Figura 6B). A filogenia baseada em SNP com o método ML, teve a formação de dois grupos, o primeiro formado por ASMN341578v1 e o segundo pelo restante das estirpes (Figura 6C). Com relação a *Pch*, comparando as árvores geradas por UPGMA (Figura 7A) e NJ (Figura 7B), ocorre uma diferença no agrupamento das sequências, principalmente na posição das estirpes ASMN341675v1 e ASMN341610. Na árvore filogenética gerada com base em SNP pelo método ML, repete o número de agrupamentos apontado por NJ, porém ocorre uma mudança na localização da estirpe ASM341569v1 (Figura 7C).

Patovares de *P. syringae* que infectam plantas de café são separados em distintos clados

A fim de compreender como as espécies de *Pseudomonas*, que infectam plantas de café se correlacionam com as demais espécies/patovares do mesmo gênero, uma análise filogenética baseada em identidade média de nucleotídeos (ANIb) foi executada, com o programa JSpecies, utilizando sequências únicas de cada grupo, selecionadas por critérios de qualidade de montagem do genoma como N50, L50 e número de *contigs*. A árvore filogenética revelou que *P. syringae* pv. *garcae* é filogeneticamente mais próxima de *P. coronafaciens* pv. *coronafaciens*, formando um clado separado, enquanto *P. syringae* pv. *tabaci* é mais próxima de *P. amygdali* pv. *lachrymans*, formando um clado vizinho a outro formado por *P. syringae* pv. *glycinea* e *P. syringae* pv. *phaseolicola*. Por sua vez, *P. cichorii* formou um grupo externo em relação às demais espécies/patovares testadas, sendo assim o mais divergente.

DISCUSSÃO

No presente estudo, analisando sequências genômicas de estirpes de *Psgc*, *Psta* e *Pch*, utilizando o algoritmo Mauve, foi possível observar regiões conservadas do genoma (LCB), compartilhadas entre as três espécies/patovares de *Pseudomonas* patogênicas ao cafeeiro. As informações genéticas contidas em algumas dessas regiões podem contribuir para entender os possíveis mecanismos comuns associados à capacidade destas espécies/patovares causarem doença no hospedeiro. Por sua vez, as regiões variáveis podem auxiliar na compreensão dos mecanismos subjacentes à diversidade genética dentro da espécie/patovar. Em estudo de genômica comparativa das estirpes B728a, B301D e HS191 de *P. syringae* pv. *syringae*, Ravindran e colaboradores (2015), através da utilização do Mauve, puderam identificar regiões de sintonia gênica, em parte do genoma, bem como regiões de inversões e translocações que auxiliaram na compreensão da diversidade dentro da espécie. Também utilizando Mauve, Bonnin (2018), em seu estudo sobre características moleculares de estirpes de *P. aeruginosa*, identificaram regiões

comuns em um plasmídeo que confere resistência a antibióticos da classe carbapenema nas estirpes 5301-1, 5301-2 e 5301-4.

O alinhamento das sequências genômicas das diferentes espécies/patovares bacterianos patogênicos ao cafeeiro mostrou que o genoma de *Psgc* pode possuir um número consideravelmente menor de rearranjos, quando comparados com aqueles de *Psta* e *Pch*, sugerindo maior estabilidade. Esses resultados são condizentes com o fato que a gama de hospedeiros de *Psgc* é consideravelmente mais estreita (Badel & Zambolim 2019; Barta & Willis, 2005). Por sua vez, a comparação dos genomas entre essas espécies/patovares confirmou a maior divergência entre estirpes de diferentes espécies. As comparações dos proteomas entre essas espécies/patovares confirmaram essa interpretação.

As análises usando Mauve permitiram obter informações qualitativas em relação às variações dos genomas. Para entender a plasticidade dos genomas em termos quantitativos, estudos de pangenômica são mais informativos. O repertório de genes identificados em uma espécie aumenta com o número de genomas analisados. Uma espécie bacteriana possui e mantém um genoma principal (genoma *core*) e um repertório de genes dispensáveis (acessórios e específicos) que contribui para a diversidade, distribuídos entre as diferentes estirpes daquela espécie (Policarpo 2015; Dagan & Martin, 2007). Sendo assim, torna-se relevante conhecer quantos genomas deveriam ser sequenciados para representar com precisão o repertório total de genes de uma dada espécie ou patovar. Desta maneira, estudos de pangenoma tentam solucionar quantos novos genes são identificados cada vez que um genoma de uma espécie de interesse é sequenciado (Tettelin *et al.*, 2008; Medini *et al.*, 2005). Neste estudo do pangenoma de *Psgc*, *Psta* e *Pch* foi encontrado que essas espécies/patovares possuem alta variabilidade no conteúdo gênico, sendo essa uma característica marcante de espécies que possuem um genoma aberto. Este resultando é condizente com um estudo de diversidade com mais de mil genomas de espécies do gênero *Pseudomonas*, onde a análise de pangenoma indicou a presença de 1.224 famílias de genes codificadores de proteínas, pertencentes à porção do genoma *core*, indicando também que esse gênero possui um genoma aberto e que a diversidade do gênero possa ser ainda maior do que o encontrado até agora (Jun *et al.*, 2015).

As análises filogenéticas demonstraram diferenças evolutivas entre as 30 estirpes de *Psgc* de Minas Gerais, que podem auxiliar a entender como se estrutura

a população do fitopatógeno. Embora ainda muito preliminares, os resultados dessas análises igualmente sugerem que as populações de *Psta* e *Pch* possuam alta diversidade. Um único estudo realizado por Peng et al 2003, demonstrou que as populações de *Psta* são altamente diversas geneticamente e embora que *Pch* seja patógeno de outras culturas de interesse econômico, como a alface e a berinjela, possui escassos estudos reportados (Badel & Zambolim, 2019; Cottyn *et al.*, 2009; Hojo *et al.*, 2008). As implicações da diversidade e da plasticidade dos genomas dessas bactérias na sua patogenicidade ao cafeeiro são ainda desconhecidas e estudos mais aprofundados precisam ser realizados.

Usando ANIb, foi analisada a relação filogenética entre 15 sequências genômicas de espécies e patovares do gênero *Pseudomonas*. Os resultados confirmaram que *Psgc* é próxima de *P. coronafaciens* pv. *coronafaciens*, enquanto *Psta* é mais próxima de *P. amygdali* pv. *lachrymans*. Como esperado, *Pch* formou um grupo externo, sendo mais distante filogeneticamente dos patovares de *P. syringae* analisados. Esses resultados são condizentes com os encontrados por Parkinson et al (2011) e por Raimundi (2017), mediante análise filogenética com base no locus *rpoD*, e por Gardan et al. (1999), baseado em DDH. Quando combinados, os resultados de todos esses trabalhos consistentemente posicionam *Psgc* no filogrupo 4 (grupo 'coronafaciens') e *Psta* no filogrupo 2 (grupo 'amygdali'). Por conta disso, *Psgc* tem sido ultimamente referida como *P. coronafaciens* pv. *garcae*, enquanto *Psta* como *P. amygdali* pv. *tabaci*.

Por fim, nossos resultados indicam uma alta diversidade genética das bactérias patogênicas ao cafeeiro. Essa diversidade genética é possivelmente o resultado tanto de rearranjos genômicos quanto de variações no tamanho dos genomas acessório e específico. Nosso estudo serve de princípio para desenvolver estudos voltados a entender os mecanismos envolvidos na adaptação dessas bactérias para infectar o cafeeiro.

REFERÊNCIAS

- Amaral JF, Do Teixeira C, Pinheiro ED, 1956. A bactéria causadora da mancha aureolada do cafeeiro. *Arquivos do Instituto Biológico* **23**, 151-155.
- Badel JL, Zambolim L, 2019. Coffee bacterial diseases: a plethora of scientific opportunities. *Plant Pathology* **68**, 411- 425.
- Bardou P, Mariette J, Escudie F, Djemiel C, Klopp C, 2014. Jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* **15**, 293.
- Barta TM, Willis DK, 2005. Biological and molecular evidence that *Pseudomonas syringae* pathovars *coronafaciens*, *striaefaciens* and *garcae* are likely the same pathovar. *Journal of Phytopathology* **153**, 492–499.
- Bonnin RA, Bogaerts P, Girlich D, Huang DT, Dortet L, Glupczynski Y, Naas T, 2018. Molecular characterization of OXA-198 carbapenemase producing *Pseudomonas aeruginosa* clinical isolates. *Antimicrobial Agents and Chemotherapy* **62**, e02496-17.
- Chan JZ-M, Halachev MR, Loman NJ, Constantinidou C, Pallen MJ, 2012. Defining bacterial species in the genomic era: insights from the genus *Acinetobacter*. *BMC Microbiology* **12**, 302.
- Chen K, Lu CL, 2018. CSAR-web: a web server of contig scaffolding using algebraic rearrangements. *Nucleic Acids Research* **46**, 55-59.
- Chen X, Zhang Y, Zhang Z, Zhang Y, Sun C, Yang M, Wang J, Liu Q, Zhang B, Chen M, Yu J, Wu J, Jin Z and Xiao J, 2018. PGAweb: a web server for bacterial pan-genome analysis. *Frontiers in Microbiology* **9**, 1910.
- Ciufo S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M, 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *International Journal of Systematic and Evolutionary Microbiology* **68**, 2386-2392.
- Cottyn B, Heylen K, Heyrman J, Vanhouteghem K, Pauwelyn E, Bleyaert P, Vaerenbergh JV, Höfte M, Vos PD, Maes M, 2009. *Pseudomonas cichorii* as the causal agent of midrib rot, an emerging disease of greenhouse-grown butterhead lettuce in Flanders. *Systematic and Applied Microbiology* **32**, 211–225.
- DaMatta FH, Ramalho JDC, 2004. Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology* **16**, 55-81.

Dagan T, Martin W, 2007. Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. *Proceedings of the National Academy of Sciences* **104**, 870-875.

Darling AE, Mau B, Perna NT, 2010. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* **5**, e11147.

Destéfano SAL, Rodrigues LMR, Beriam LOS, Patrício FRA, Thomaziello RA, Rodrigues-Neto J, 2010. Bacterial leaf spot of coffee caused by *Pseudomonas syringae* pv. *tabaci* in Brazil. *Plant Pathology* **59**, 1162–1163.

Felsenstein J, 1989. PHYLIP – Phylogeny Inference Package. *Cladistics* **5**, 164-166.

Gardan L, Shafik H, Belouin S, Broch R, Grimont F, 1999. DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). *International Journal of Systematic Bacteriology* **49**, 469–478.

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM, 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology* **57**, 81-91.

Hojo H, Koyanagi M, Tanaka M, 2008. The *hrp* genes of *Pseudomonas cichorii* are essential for pathogenicity on eggplant but not on lettuce. *Microbiology* **154**, 2920–2928.

Huang Y, Niu B, Gao Y, Fu L, Li W, 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**: 680-682.

Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S, 2018. High throughput ANI analysis of 90k prokaryotic genomes reveals clear species boundaries. *Nature Communications* **9**, 5114.

Jun S, Wassenaar TM, Nookaew I, Hauser L, Wanchai V, Land M, Timm CM, Lu TS, Schadt CW, Doktycz MJ, Pelletier DA, Ussery DW, 2015. Diversity of *Pseudomonas* genomes, including *Populus*-associated isolates, as revealed by comparative genome analysis. *Applied and Environmental Microbiology* **82**, 375-383.

Kimura O, Robbs CF, Ferrari JAR, 1976. Algunas observaciones relacionadas con las bacteriosis de cafeiro. In: *Congresso de Pesquisas Cafeeiras 4.*, Guarapari, Rio de Janeiro, 104.

Konstantinidis KT, Tiedje JM, 2009. Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences U. S. A.* **102**, 2567-2572.

Lopes-Santos L, Castro DBA, Ferreira-Tonin M, Corrêa DBA, Weir BS, Park D, Ottoboni LMM, Neto JR, Destéfano SAL, 2017. Reassessment of the taxonomic position of *Burkholderia andropogonis* and description of *Robbsia andropogonis* gen. nov., com. nov. *Antonie Van Leeuwenhoek* **110**, 727-736.

Maciel KW, Destefano SAL, Beriam LOS, Almeida IMG, Patrício FRA, Rodrigues LMR, Guerreiro Filho O, 2018. Bacterial halo blight of coffee crop: aggressiveness and genetic diversity of strains. *Bragantia* **77**, 96–106.

Medini D, Donati C, Tettelin H, Massignani V, Rappuoli R, 2005. The microbial pan-genome. *Current Opinion in Genetics and Development* **6**, 589-594.

Monteil CL, Yahara K, Studholme DJ, Megeiros L, Méric G, Swingle B, Morris CE, Vinatzer BA, Sheppard SK, 2016. Population-genomic insights into emergence, crop-adaptation, and dissemination of *Pseudomonas syringae* pathogens. *Microbial Genomics* **21**, e000089.

Mwangi IJ, Gichuru EK, Masanga J, Alakonya AE, 2018. A Comparative analysis associated to virulence of *Pseudomonas syringae* pv. *garcae*, the causative agent of bacterial blight of coffee in Kenya. *Journal of Scientific Agriculture* **2**, 76–84.

Parkinson N, Bryant R, Bew J, Elphinstone J, 2011. Rapid phylogenetic identification of members of the *Pseudomonas syringae* species complex using the *rpoD* locus. *Plant Pathology* **60**, 338-344.

Peng R, Zhang S, Wang S, 2003. The research progress on physiological race of *Pseudomonas syringae* pv. *tabaci*. *Journal of the Yunnan Agricultural University* **18**, 198-202.

Perrière G, Gouy M, 1996. WWW-Query: an on-line retrieval system for biological sequence banks. *Biochimie*, **78**, 364-369.

Policarpo AD, 2015. Análise do pangenoma de *Streptococcus pneumoniae* e comparação de genomas de serótipos 1 e 3. Dissertação de Mestrado. Universidade de Lisboa.

Raimundi MK, 2017 Identificação bioquímica, molecular e patogenicidade de isolados de *Pseudomonas* spp. provenientes de cafeeiros em Minas Gerais. Lavras, Tese de Doutorado. Brazil: Universidade Federal de Lavras.

Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebahia M, Thomson NR, Chaudhuri R, Henderson IR, Sperandio V, Ravel J, 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *Journal of Bacteriology* **190**, 6881-6893.

Rashid MR, Young JP, Overall I, Clercx P, Willems A, Braun MS, Wink M, 2015. Average nucleotide identity of genome sequences supports the description of *Rhizobium lentis* sp. nov., *Rhizobium bangladeshense* sp. nov. and *Rhizobium binae* sp. nov. from lentil (*Lens culinaris*) nodules. *International Journal of Systematic and Evolutionary Microbiology* **65**, 3037-3045.

Ravindran A, Jalan N, Yuan JS, Wang DC, 2015. Comparative genomics of *Pseudomonas syringae* pv. *syringae* strains B301D and HS191 and insights into intrapathovar traits associated with plant pathogenesis. *Microbiology Open* **4**, 553-573.

Richter M, Rosselló-Móra R, 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences U. S. A.* **106**, 19126-19131.

Richter M, Rosselló-Móra R, Glockner FO, Peplies J, 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* **32**, 929-931.

Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT, 2009. Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics* **25**, 2071-2073.

Robbs CF, Kimura O, Ribeiro RLD, Oyadomari, L. C. 1974. Bacterial leaf blight: a new disease of coffee (*Coffea arabica* L.) caused by *Pseudomonas cichorii* (Swingle) Stapp. *Arquivos da Universidade Federal Rural do Rio de Janeiro* **4**, 1–5.

Rodrigues LMR, Almeida IMG, Patricio FRA, 2013. Mancha Aureolada do Cafeeiro Causada por *Pseudomonas syringae* pv. *garcae*. Campinas, SP, Brasil: *Instituto Agrônômico de Campinas*.

Rodrigues LMR, Almeida IMG, Patrício FRA, Beriam LOS, Maciel KW, Braghini MT, Guerreiro Filho O, 2017a. Aggressiveness of strains and inoculation methods for resistance assessment to Bacterial Halo Blight on coffee seedlings. *Journal of Phytopathology* **165**, 105–114.

Rodrigues LMR, Sera GH, Filho OG, Beriam LOS, Almeida IMG, 2017b. First report of mixed infection by *Pseudomonas syringae* pathovars *garcae* and *tabaci* on coffee plantations. *Bragantia* **76**, 543-549.

Rodrigues Neto J., Silva C. H. D., Beriam, LOS, 2006. Mancha bacteriana do cafeeiro causada por *Pseudomonas syringae* pv *tabaci*. *Summa Phytopathologica* **32**, 8.

Sánchez L, Zapata M, Rodríguez RP, 2003. Distribución y patogenicidad de *Pseudomonas cichorii* (Swingle) Stapp en el cultivo de café em Puerto Rico. *Journal of Agriculture of the University of Puerto Rico* **87**, 123-135.

Smith EF, 1911. *Bacteria in Relation to Plant Diseases*. Washington DC: *Carnegie Institute Publications*,

Tettelin H, Riley D, Cattuto C, Medini D, 2008. Comparative genomics: the bacterial pan-genome. *Current Opinion in Microbiology* **11**, 472-477.

Tian X, Zhang Z, Yang T, Chen M, Li J, Chen F, Yang J, Li W, Zhang B, Zhang Z, Wu J, Zhang C, Long L, Xiao J, 2016. Comparative genomics analysis of *Streptomyces* species reveals their adaptation to the marine environment and their diversity at the genomic level. *Frontiers in Microbiology* **7**, 998.

Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpides NC, Pati A, 2015. Microbial species delineation using whole genome sequences. *Nucleic Acids Research* **43**, 6761-6771.

Vernikos G, Medini D, Riley DR, Tettelin H, 2015. Ten years of pan-genome analyses. *Current Opinion in Microbiology*. **23**, 148–154.

Wells JM, Raju BC, Hung HY, Weisburg WG, Mandelcopaul L, Brenner DJ, 1987. *Xylella fastidiosa* gen-nov. sp-nov. - Gram negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. *International Journal of Systematic Bacteriology* **37**, 136–143.

Xiao J, Zhang Z, Wu J, Yu J, 2015. A brief review of software tools for pangenomics. *Genomics Proteomics Bioinformatics* **13**, 73–76.

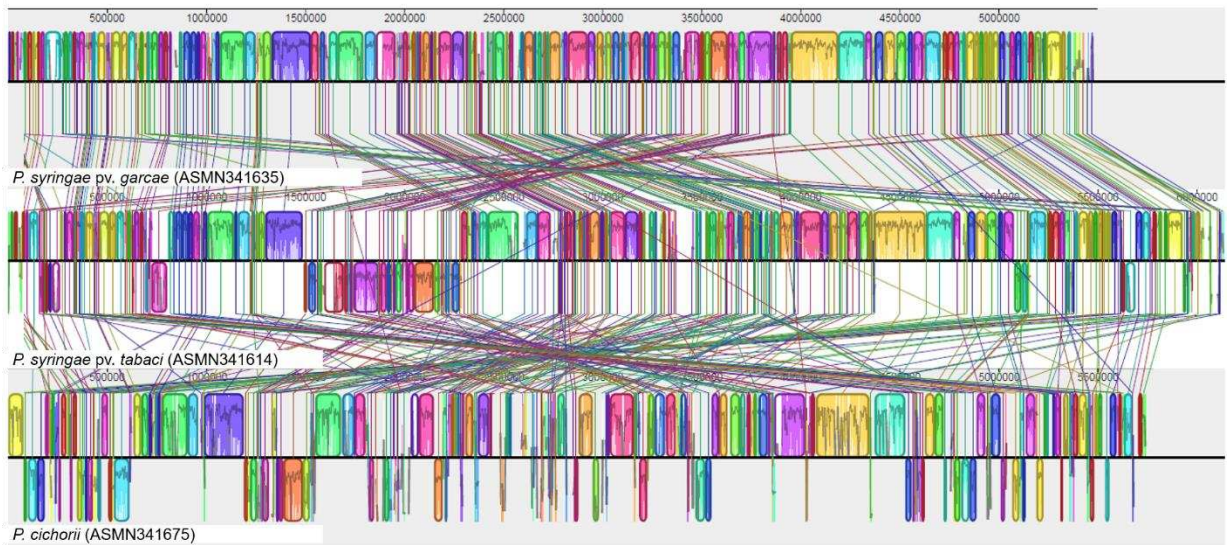
FIGURAS

Figura 1. Alinhamento das sequências genômicas de *Pseudomonas syringae* pv. *garcae*, *P. syringae* pv. *tabaci* e *Pseudomonas cichorii*, usando Mauve. As sequências rascunho RefSeq obtidas do GenBank foram primeiro alinhadas com o genoma de referência (*P. syringae* pv. *tomato* DC3000) e em seguida, as sequências com os *contigs* ordenados foram alinhadas. Os números indicam as posições em pares de bases dos genomas.

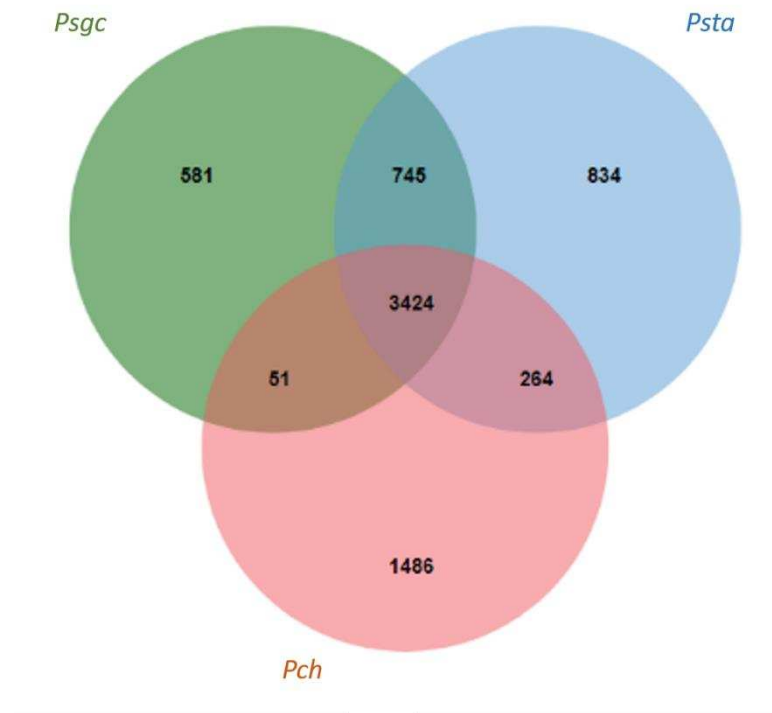


Figura 2. Diagrama de Venn demonstrando as proteínas compartilhadas e exclusivas de *Pseudomonas syringae* pv. *garcae*, *P. syringae* pv. *tabaci* e *Pseudomonas cichorii*, determinadas usando CD-HIT. Os números nas interseções indicam as proteínas que são compartilhadas entre duas ou três espécies/patovares. O número de proteínas de cada patógeno é agrupado por uma cor e indicam as proteínas exclusivas de cada espécie/patovar.

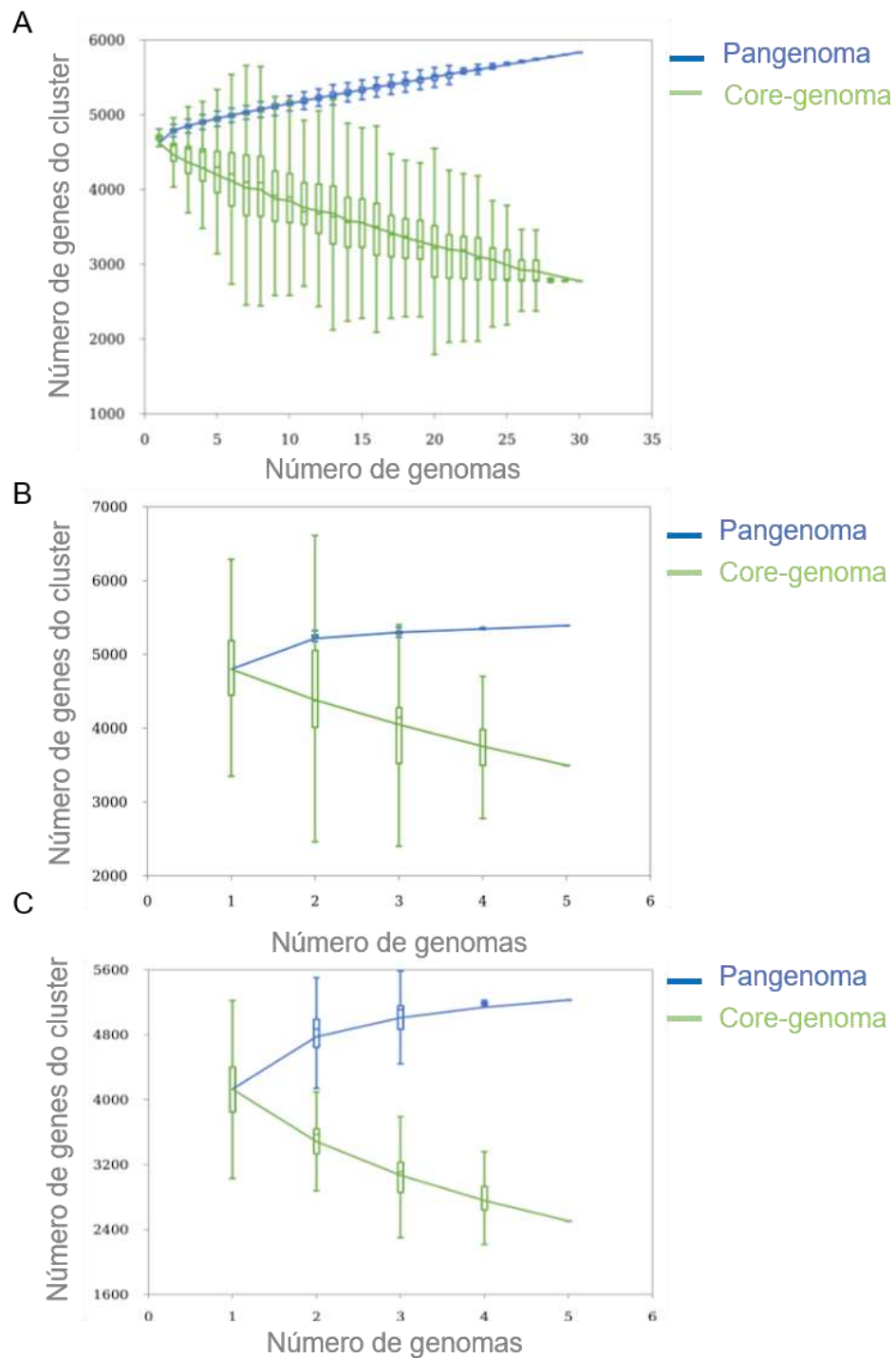


Figura 3. Modelos funcionais do genoma *core* e do pangenoma de (A) *Pseudomonas syringae* pv. *garca*, (B) *P. syringae* pv. *tabaci* e (C) *Pseudomonas cichorii*. O número de genes de cada *cluster* (pangenoma ou *core*) foram plotados em função do número de genomas adicionados. A cada adição foi calculada uma nova média e mediana, representada pelo *boxplot* em cada ponto.

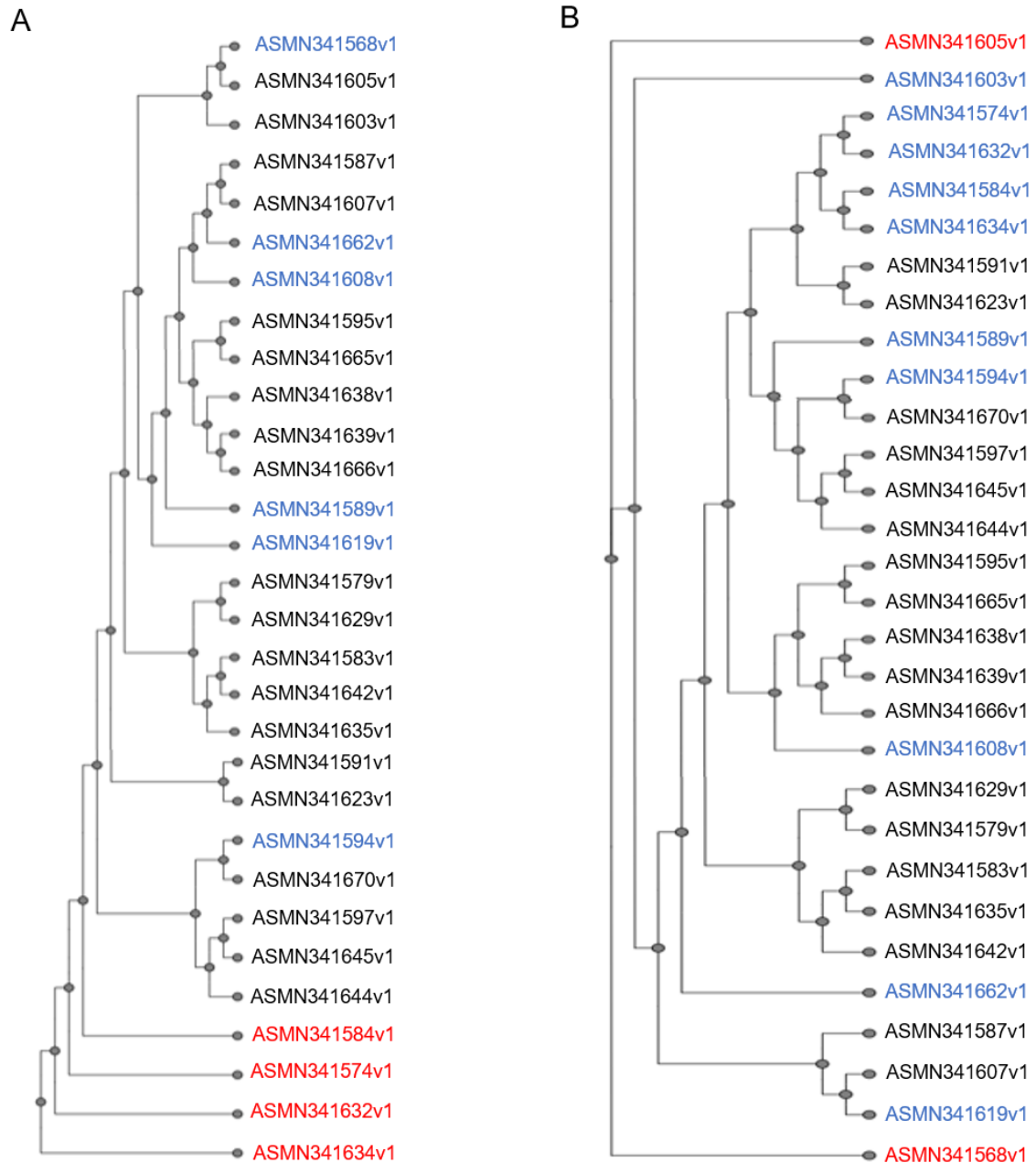


Figura 4. Árvores filogenéticas geradas utilizando 30 seqüências genômicas de estirpes de *Pseudomonas syringae* pv. *garcae*, baseadas na matriz de genes PanBased. As árvores foram construídas utilizando os métodos de (A) Médias Aritimédicas não Ponderadas (UPGMA) e (B) *Neighbor-Joining* (NJ). As estirpes mais divergentes são destacados em vermelho. As estirpes que apresentaram agrupamentos diferentes, segundo o método utilizado, são destacadas em azul.

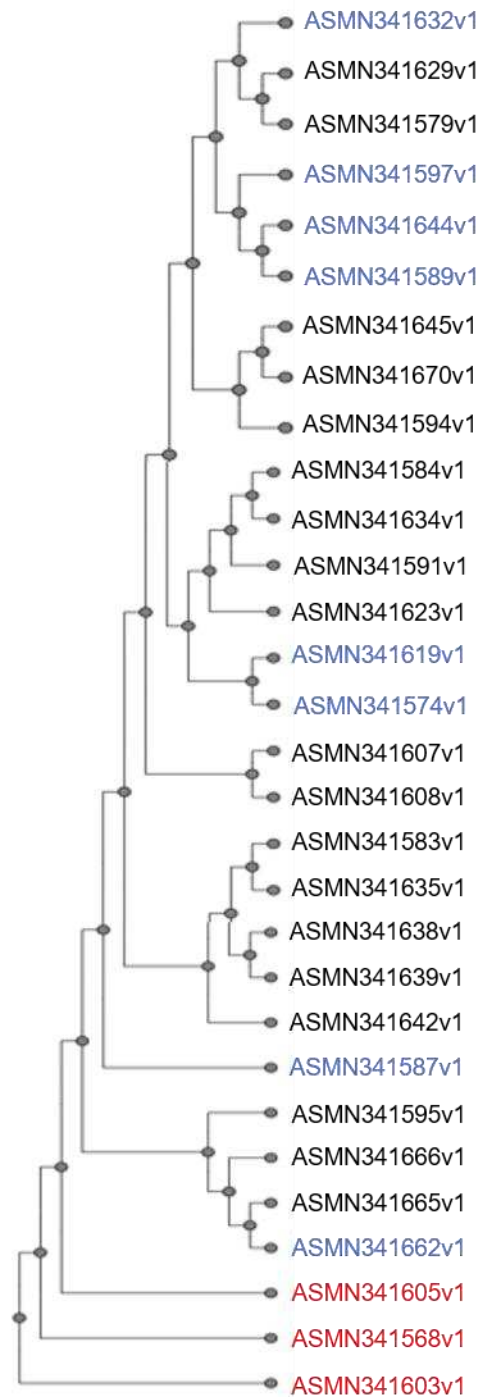


Figura 5. Árvore filogenética construída utilizando 30 sequências genômicas de estirpes de *Pseudomonas syringae* pv. *garcae*, baseadas em SNP do genoma *core*, utilizando o método de máxima verossimilhança (ML). As estirpes mais divergentes são destacadas em vermelho. As estirpes que apresentaram agrupamentos diferentes, com relação a árvore baseada no pangenoma usando o método NJ, são destacadas em azul.

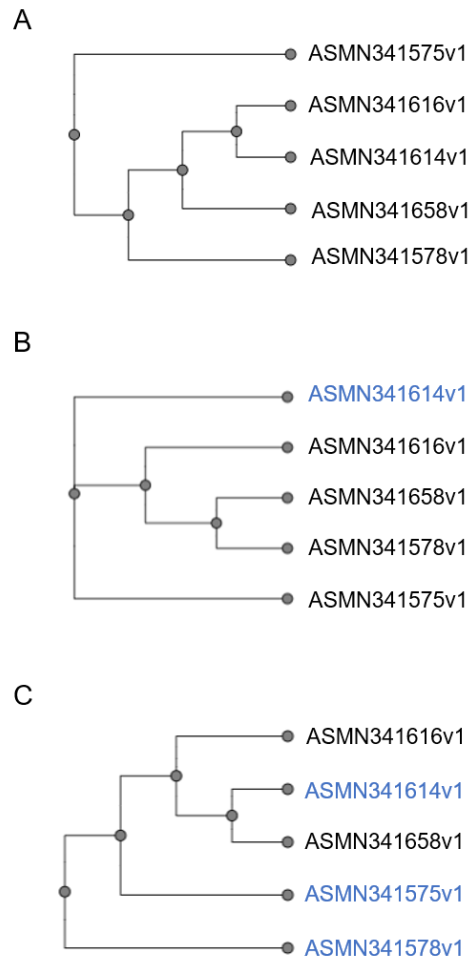


Figura 6. Árvores filogenéticas construídas utilizando cinco sequências genômicas de estirpes de *Pseudomonas syringae* pv. *tabaci*, baseadas na matriz de genes PanBased (A e B) e em SNP do genoma core. Os métodos de agrupamento utilizados foram (A) Médias Aritimédicas não Ponderadas (UPGMA), (B) *Neighbor-Joining* (NJ) e (C) Máxima Verossimilhança (ML). As estirpes que tiveram o agrupamento diferente, dependendo do método utilizado, são destacadas em azul.

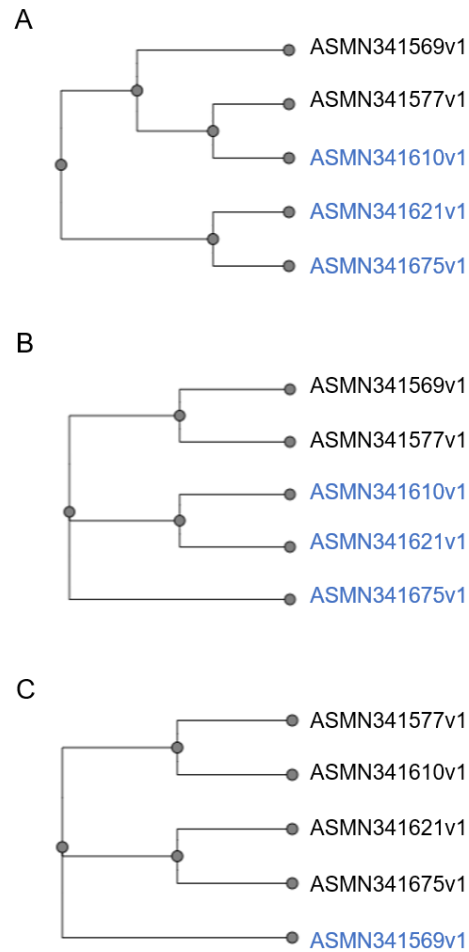


Figura 7. Árvores filogenéticas construídas utilizando cinco sequências genômicas de estirpes de *Pseudomonas cichorii*, baseadas na matriz de genes PanBased (A e B) e em SNP do genoma *core*. Os métodos de agrupamento utilizados foram (A) Médias Aritimédicas não Ponderadas (UPGMA), (B) *Neighbor-Joining* (NJ) e (C) Máxima Verossimilhança (ML). As estirpes que tiveram o agrupamento diferente dependendo do método utilizado são destacadas em azul.

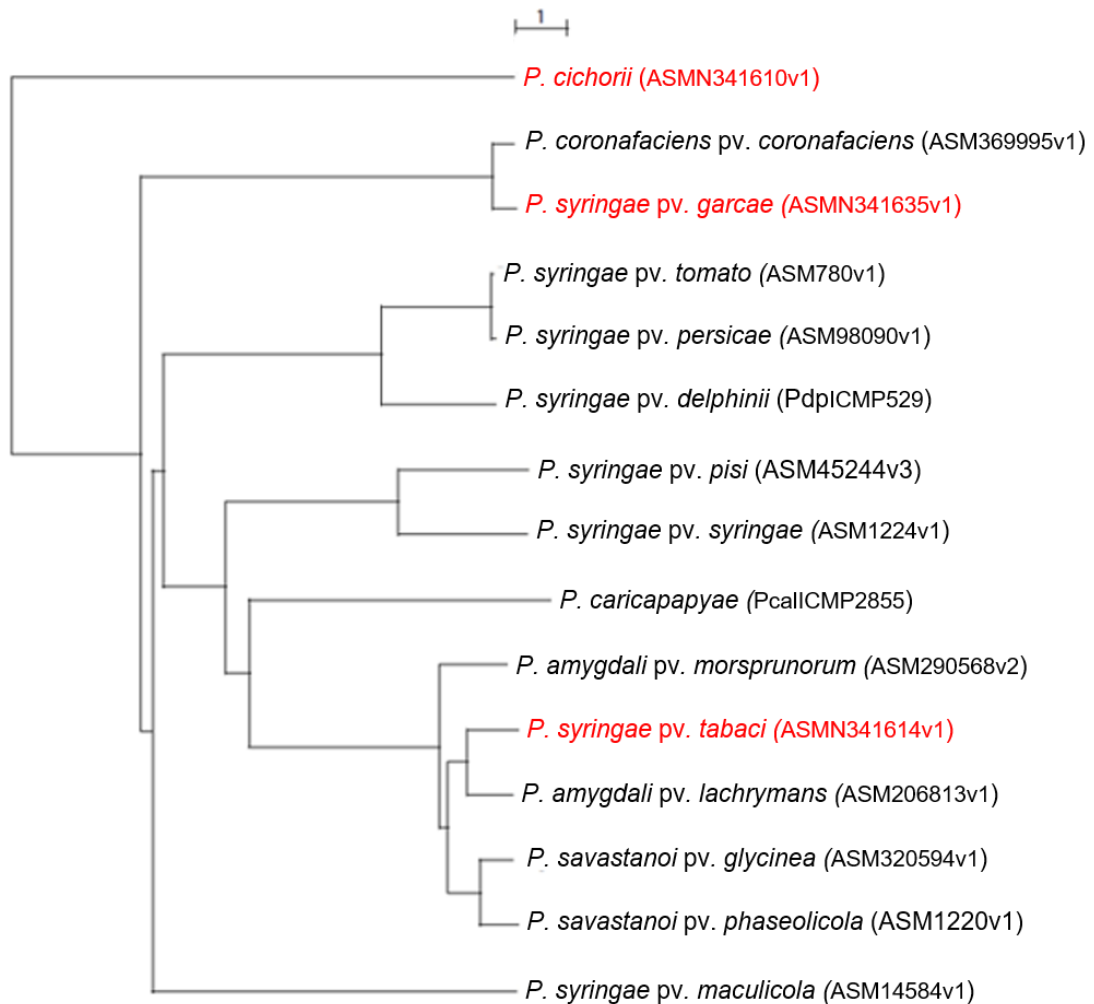


Figura 8. Árvore filogenética de espécies/patovares de *Pseudomonas*, baseadas em valores de ANIb, construídas usando o método *Neighbor-Joining* (NJ). Os valores de ANIb foram calculados com o programa JSpecies, a árvore foi construída com PHYLIP e editada em NJPLOT. O número de acesso ao GenBank é indicado entre parênteses. As bactérias patogênicas ao cafeeiro são indicadas em vermelho.

MATERIAL SUPLEMENTAR

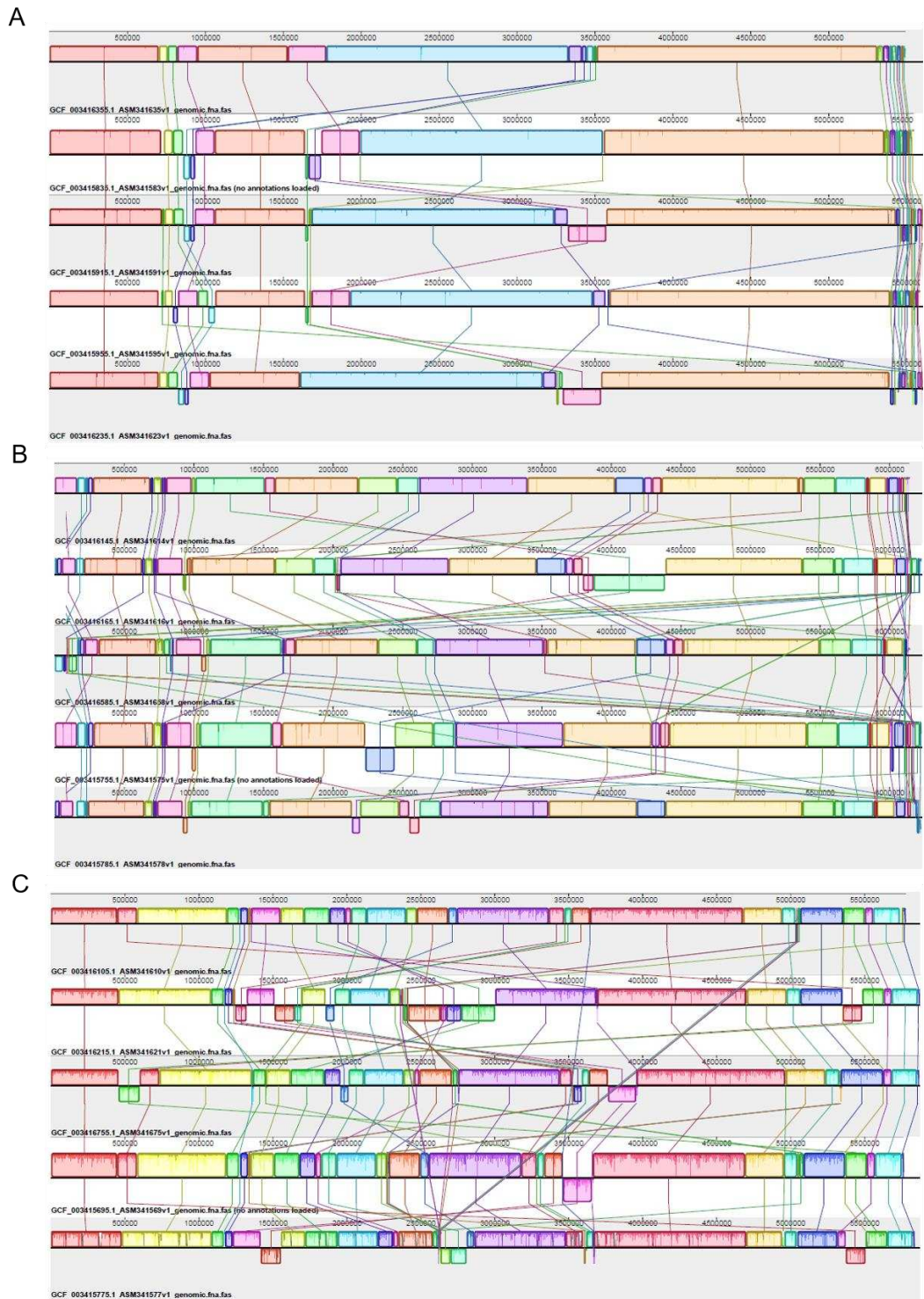


Figura Suplementar S1. Alinhamento das cinco seqüências genômicas de cada patógeno bacteriano do cafeeiro. (A) *Pseudomonas syringae* pv. *garcae*. (B) *P. syringae* pv. *tabaci*. (C) *Pseudomonas cichorii*. Os *contigs* das seqüências RefSeq recuperadas do GenBank foram inicialmente alinhadas com o genoma de referência (*P. syringae* pv. *tomato* DC3000). Blocos colineares locais são conectados por uma linha e identificados pela mesma cor. Os números indicam as posições nos genomas, em pares de bases,

Tabela Suplementar S1. Genomas de estirpes de *Pseudomonas* obtidas de plantas de *Coffea arabica* no estado de Minas Gerais, recuperados do GenBank.

Espécie	Patovar	Estirpe	Número de acesso	Comprimento da sequência	Nº de Contigs	N50	L50
<i>coronafaciens</i>	<i>garcae</i>	462*	ASM341635v1	5,495,479	42	259,852	6
<i>coronafaciens</i>	<i>garcae</i>	471*	ASM341623v1	5,597,954	48	301,655	7
<i>coronafaciens</i>	<i>garcae</i>	445*	ASM341591v1	5,616,225	51	303,153	6
<i>coronafaciens</i>	<i>garcae</i>	460*	ASM341583v1	5,541,917	53	262,266	6
<i>coronafaciens</i>	<i>garcae</i>	438*	ASM341595v1	5,579,882	53	293,898	6
<i>coronafaciens</i>	<i>garcae</i>	436	ASM341665v1	5,583,686	54	309,917	6
<i>coronafaciens</i>	<i>garcae</i>	457	ASM341642v1	5,551,542	56	282,795	7
<i>coronafaciens</i>	<i>garcae</i>	456	ASM341584v1	5,541,938	58	239,887	7
<i>coronafaciens</i>	<i>garcae</i>	484	ASM341607v1	5,653,368	59	303,157	5
<i>coronafaciens</i>	<i>garcae</i>	451	ASM341589v1	5,605,581	61	313,463	6
<i>coronafaciens</i>	<i>garcae</i>	455	ASM341587v1	5,656,459	62	291,81	7
<i>coronafaciens</i>	<i>garcae</i>	465	ASM341629v1	5,591,140	62	303,164	5
<i>coronafaciens</i>	<i>garcae</i>	459	ASM341639v1	5,626,802	62	235,348	7
<i>coronafaciens</i>	<i>garcae</i>	487	ASM341603v1	5,698,373	64	239,921	7
<i>coronafaciens</i>	<i>garcae</i>	483	ASM341608v1	5,592,424	65	239,837	8
<i>coronafaciens</i>	<i>garcae</i>	461	ASM341638v1	5,622,834	65	303,136	6
<i>coronafaciens</i>	<i>garcae</i>	475	ASM341574v1	5,649,569	66	272,99	6
<i>coronafaciens</i>	<i>garcae</i>	443	ASM341594v1	5,561,375	67	259,911	6
<i>coronafaciens</i>	<i>garcae</i>	485	ASM341605v1	5,698,025	67	275,139	7
<i>coronafaciens</i>	<i>garcae</i>	434	ASM341666v1	5,635,253	67	235,045	7
<i>coronafaciens</i>	<i>garcae</i>	466	ASM341579v1	5,590,450	69	291,81	6
<i>coronafaciens</i>	<i>garcae</i>	454	ASM341644v1	5,681,222	69	239,853	6
<i>coronafaciens</i>	<i>garcae</i>	486	ASM341568v1	5,706,553	71	239,837	6
<i>coronafaciens</i>	<i>garcae</i>	432	ASM341670v1	5,655,450	73	235,298	5
<i>coronafaciens</i>	<i>garcae</i>	463	ASM341632v1	5,683,688	74	173,302	10
<i>coronafaciens</i>	<i>garcae</i>	476	ASM341619v1	5,647,280	75	303,167	7
<i>coronafaciens</i>	<i>garcae</i>	450	ASM3416445v1	5,676,804	76	272,632	6

<i>coronafaciens</i>	<i>garcae</i>	437	ASM341662v1	5,683,912	76	283,4	7
<i>coronafaciens</i>	<i>garcae</i>	464	ASM341634v1	5,609,724	77	262,224	7
<i>coronafaciens</i>	<i>garcae</i>	150	ASM41597v1	5,728,017	78	249,481	7
<i>amygdali</i>	<i>tabaci</i>	480*	ASM341614v1	6,138,565	90	124,493	15
<i>amygdali</i>	<i>tabaci</i>	440*	ASM341658v1	6,150,227	91	132,939	15
<i>amygdali</i>	<i>tabaci</i>	469*	ASM341578v1	6,222,674	99	136,569	15
<i>amygdali</i>	<i>tabaci</i>	478*	ASM341575v1	6,232,055	101	151,325	13
<i>amygdali</i>	<i>tabaci</i>	479*	ASM341616v1	6,223,414	101	122,192	15
<i>cichorii</i>	-	481*	ASM341610v1	5,778,505	38	312,803	6
<i>cichorii</i>	-	482*	ASM341569v1	5,765,860	49	256,289	9
<i>cichorii</i>	-	473*	ASM341577v1	5,844,590	50	262,056	9
<i>cichorii</i>	-	474*	ASM341621v1	5,871,845	52	182,316	9
<i>cichorii</i>	-	136*	ASM341675v1	5864,912	78	174,522	10

(*) Estirpes que foram selecionadas para alinhamento e análise de proteômica, com base em parâmetros como número de *contigs*, N50 e L50.

Tabela Suplementar S2. Estirpes de *Pseudomonas* cujos genomas foram obtidos do banco de dados GenBank e utilizados na análise filogenética baseada em ANIb.

Espécie	Patovar	Estirpe	Número da amostra	Nª de contigs	N50	L50
<i>coronafaciens</i>	<i>garcae</i>	462	ASM341635v1	42	259,85	6
<i>amygdali</i>	<i>tabaci</i>	480	ASM341614v1	90	124,49	15
<i>cichorii</i>	-	481	ASM341610v1	38	312,80	6
<i>syringae</i>	<i>tomato</i>	DC3000	ASM780v1	-	-	-
<i>savastanoi</i>	<i>phaseolicola</i>	1448A	ASM1220v1	-	-	-
<i>syringae</i>	<i>syringae</i>	B728A	ASM1224v1	-	-	-
<i>syringae</i>	<i>maculicola</i>	ES4326	ASM14584v1	878	17,22	101
<i>syringae</i>	<i>pisi</i>	PP1	ASM45244v3	-	-	-
<i>caricapapayae</i>	-	ICMP2855	PcaICMP2855	198	78,66	26
<i>amygdali</i>	<i>lachrymans</i>	NM002	ASM206813v1	-	-	-
<i>amygdali</i>	<i>morsprunorum</i>	R15244	ASM290568v2	-	-	-
<i>savastanoi</i>	<i>glycinea</i>	2411	ASM320594v1	800	57,99	32
<i>coronafaciens</i>	<i>coronafaciens</i>	ICMP 10052	ASM369995v1	202	120,08	13
<i>syringae</i>	<i>delphinii</i>	ICMP529	PdpICMP529	428	27,13	65
<i>syringae</i>	<i>persicae</i>	NCPPB 2254	ASM98090v1	201	69,35	31

(-) Sequências com o genoma completo.

Capítulo 2: Divergência e conservação gênica dos sistemas de secreção e efetores de tipo III em *Pseudomonas* patogênicas ao cafeeiro de Minas Gerais

Francisco Henrique Nunes da Silva Alves¹, Jorge Luis Badel¹

¹ Laboratório de Fitobacteriologia Molecular, Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG, Brasil.

Autor para correspondência: Jorge Luis Badel, e-mail: jorge.badel@ufv.br

RESUMO

As manchas foliares causadas pelas bactérias fitopatogênicas *P. syringae* pv. *garcae* (*Psgc*), *P. syringae* pv. *tabaci* (*Psta*) e *Pseudomonas cichorii* (*Pch*) em plantas de café ocasionam grandes perdas econômicas na produção de mudas em viveiro e também no campo. Apesar da grande importância do café, não existem trabalhos relacionados aos mecanismos moleculares associados à capacidade de *Psgc*, *Psta* e *Pch* em causarem doenças neste hospedeiro. Assim, este trabalho teve como objetivo identificar os *clusters* de genes *hrp/hrc* e prever os efetores de tipo III (T3SE) presentes em genomas de *Psgc*, *Psta* e *Pch*, potencialmente envolvidos na capacidade destes patógenos causarem doença. Para tanto, as sequências genômicas das *Pseudomonas* patogênicas ao cafeeiro, obtidas do banco de dados (GenBank), tiveram os *cluster* de genes *hrp/hrc* de cada estirpe mapeados usando o visualizador de genomas CLC e alinhadas usando Mauve. A predição do T3SE foi realizada usando Blastp e *scripts* em Python, localizando a presença de sequências assinaturas presentes nas regiões promotoras (caixa *hrp*) e sinal de secreção na extremidade N-terminal das proteínas. Alta sintonia gênica foi observada entre os *clusters* de genes *hrp/hrc* dos diferentes patovares de *P. syringae*. Entretanto, inversões foram encontradas nos genes *hrpA*, *hrpB*, *hrpF*, *hrpG*, *hrpV*, *hrpL* e deleções dos genes *hrpD*, *hrpF*, *hrpT*, *hrpQb* e *hrpL* em pelo menos um dos patovares, o que sugere que os *clusters* sofreram alterações durante processo evolutivo. Comparações entre os pansecretomas do T3SE demonstrou que *Psgc* possui 46 famílias de efetores, sendo 10 possíveis novos efetores, *Psta* possui 43 famílias de

efetores, sendo 13 novos candidatos a efetores e *Pch* contendo 20 famílias de efetores, sendo 10 prováveis novos efetores. Foi constatado que *avrE* foi a única família de efetor compartilhado entre as três espécies/patovares bacterianas. Assim, indicando que o efetor *avrE* pode possuir papel fundamental nas interações dessas bactérias com o cafeeiro. Os resultados deste estudo são a base para a caracterização funcional dos genes candidatos e o entendimento dos mecanismos de patogenicidade destas bactérias fitopatogênicas no futuro.

Palavras-chave: *avrE*, caixa *hrp*, *cichorii*, *Coffea*, *garcae*, *hrp/hrc*, patogenicidade, *tabaci*.

INTRODUÇÃO

A cultura do cafeeiro é de extrema importância para economia brasileira, uma vez que o café arábica (*Coffea arábica* L.) e o café robusta (*Coffea canephora* L.) são umas das commodities agrícolas mais vendidas do mundo (DaMatta & Ramalho, 2004). Porém, perdas significativas na produção são causadas por pragas, e diferentes patógenos como nematoides, fungos e bactérias (Rodrigues *et al.*, 2013). Dentre as doenças bacterianas, *P. syringae* pv. *garcae* (*Psgc*) (Amaral *et al.*, 1956), *P. syringae* pv. *tabaci* (*Psta*) (Rodrigues Neto *et al.*, 2006) e *Pseudomonas cichorii* (*Pch*) (Swingle 1925) merecem destaque pelas perdas de mudas em viveiro e em campo.

Apesar da grande importância da cultura do cafeeiro e os consideráveis danos que essas bactérias fitopatogênicas causam na cultura do café, não existem trabalhos relacionados aos mecanismos moleculares que permitam que essas bactérias causem doença neste hospedeiro (Badel & Zambolim 2019).

Sabe-se que os patógenos de plantas desenvolveram mecanismos sofisticados para subverter a resposta do sistema imunológico da planta e adquirir os nutrientes (Badel *et al.*, 2013). Como primeira resposta do sistema imune das plantas, receptores no hospedeiro identificam os padrões moleculares associados ao patógeno (*Pathogen-Associated Molecular Patterns*, PAMPs), ativando uma resposta conhecida como

imunidade desencadeada por PAMPs (*PAMP-Triggered Immunity*; PTI) que promove o fluxo de íons através da membrana plasmática, a geração de espécies reativas de oxigênio, a deposição de calose e a sinalização e indução de genes relacionados à defesa (Jones & Dangl 2006; Zipfel & Robatzek 2010). Como resultado, os patógenos liberam proteínas efetoras que bloqueiam o reconhecimento dos PAMPs ou silenciam os sinais gerados pelo PTI (Badel *et al.*, 2013; Gohre & Robatzek 2008). Alguns efetores, direta ou indiretamente, são reconhecidos pela planta desencadeando uma segunda linha de resposta, conhecida como imunidade desencadeada por efetores (*Effector-Triggered Immunity*; ETI), culminando na reação de hipersensibilidade (HR) (Klement & Goodman 1967). Em resposta ao ETI, os patógenos desenvolveram proteínas efetoras que evitam o seu reconhecimento pelas plantas e suprimem os sinais associados ao ETI (Badel *et al.*, 2013; Jones & Dangl 2006).

Em bactérias fitopatogênicas Gram-negativas, o sistema de secreção do tipo III (T3SS) é codificado pelos genes *hrp* (*hypersensitive response and pathogenicity*) e *hrc* (*hrp conserved*) (Zwiesler-Vollick *et al.*, 2002), os quais são responsáveis pela montagem do pilus HrpA e injeção de proteínas efetoras diretamente na célula do hospedeiro, onde modulam a fisiologia da planta para favorecimento do patógeno (Roine *et al.*, 1997; Jin & He, 2001; Zwiesler-Vollick *et al.*, 2002). Além da proteína HrpA que forma o pilus, o T3SS é constituído por proteínas que se localizam nas membranas da célula bacteriana, incluindo HrcV e HrcC, cuja função é essencial para a secreção de efetores (Collmer *et al.*, 2000; Cornelis and Van Gijsegem, 2000; Stuber *et al.*, 2003).

Três proteínas reguladoras, HrpR, HrpS e HrpL, são requeridas para a expressão dos genes *hrp/hrc* e dos genes que codificam o T3SE em *P. syringae* (Collmer *et al.*, 2000; Grimm *et al.*, 1995; Hutcheson *et al.*, 2001; Xiao & Hutcheson, 1994; Zwiesler-Vollick *et al.*, 2002). Os reguladores HrpR e HrpS ativam a expressão dos genes *hrpL*, cuja proteína codificada por sua vez ativa os genes *hrp*, *hrc* e *hop* (*hrp outer proteins*; codificam proteínas secretadas pelo T3SS) ligando-se às regiões promotoras através do reconhecimento da caixa *hrp* (Grimm *et al.*, 1995; Lindeberg *et al.* 2005; Vencato *et al.*, 2006; Xiao & Hutcheson, 1994). Uma vez que os genes sejam expressados e as proteínas sintetizadas, estas são reconhecidas pelo T3SS devido à presença de padrões presentes nas extremidades N-terminais, constituídas principalmente por padrões de

hidrofobicidade/hidrofilicidade, ao invés de uma sequência de aminoácidos conservada (Bhattacharjee *et al.* 2006; Dou *et al.* 2008).

Considerando-se que *Psgc*, *Psta* e *Pch* causam sintomas semelhantes em plantas de café, comparações genômicas relacionadas ao T3SS e aos efetores secretados por ele, podem conduzir na identificação do conjunto de genes de efetores que sejam conservados e essenciais à patogenicidade destas espécies/patovares. Essa identificação auxiliará a compreensão da especificidade destas espécies/patovares e dos aspectos evolutivos envolvidos na adaptação destes patógenos com relação ao hospedeiro (Studholme *et al.*, 2002; Zwiesler-Vollick *et al.*, 2002).

Sendo assim, este trabalho visou identificar os *clusters* de genes *hrp/hrc* e predizer os T3SE presentes em genomas de *Psgc*, *Psta* e *Pch*. Para essa análise, ferramentas de bioinformática e *scripts* customizados foram desenvolvidos para identificar os genes com base em similaridade de sequência, características da região promotora dos genes e identificação de padrões de sinais de secreção de proteínas efetoras.

MATERIAL E MÉTODOS

Obtenção das sequências genômicas

Para obtenção das sequências genômicas de *P. syringae* pv. *garcae*, *P. syringae* pv. *tabaci* e *P. cichorii*, patogênicas ao cafeeiro, foram utilizadas as sequências curadas e depositadas como RefSeq, obtidas do banco de dados (GenBank) do *National Center for Biotechnology Information* (NCBI, <http://www.ncbi.nlm.nih.gov>).

Análise comparativa do *cluster* de genes *hrp/hrc*

Para obter as sequências de nucleotídeos do *cluster* de genes *hrp/hrc* de cada estirpe, as sequências dos genes *hrpK* e *hrpS* que definem as extremidades desse *cluster*

em *Pst* DC3000 (Alfano *et al.*, 2000) foram localizados nas sequências genômicas usando o visualizador de genomas CLC v. 8.0 (<https://www.qiagenbioinformatics.com/products/clc-sequence-viewer/>). A sintenia dos genes que compõem o *cluster* de genes *hrp/hrc* das estirpes tipo ou representativas de *Psgc*, *Psta* e *Pch* foi determinada usando o alinhador de sequências múltiplas Mauve v. 2.4.0 (<http://darlinglab.org/mauve/mauve.html>) (Rissman *et al.*, 2009). A identidade de nucleotídeos e amino ácidos entre grupos de genes e proteínas pertencentes à mesma família foi determinada usando Blast (Altschul *et al.*, 1990).

Predição e genômica comparativa de complementos de T3SE

A predição do complemento de T3SE das cinco estirpes selecionadas de cada bactéria fitopatogênica para o estudo comparativo foi feita usando Balstp (Altschul *et al.*, 1990). Para isto, as sequências de proteínas preditas nos genomas de *Psgc*, *Psta* e *Pch* foram comparadas com as sequências de efetores conhecidas de *P. syringae*, depositadas no *Pseudomonas syringae* Genome Resource (<http://www.pseudomonas-syringae.org/>), um banco de dados que contém informações a respeito da biologia de genomas de mais de 50 patovares de *P. syringae* e mais de 80 sequências de efetores conhecidos.

Adicionalmente, para a identificação de T3SE de *Psgc*, *Psta* e *Pch* foram elaborados *scripts* em Python para localizar padrões presentes nas regiões promotoras (caixa *hrp*) dos genes *hrp/hrc* e *hop* que codificam proteínas efetoras de *Pseudomonas* (Kajihara *et al.*, 2012) (Tabela 1). Um *script* encontra genes que possuem caixas *hrp* na sua região promotora, além dos dois genes seguintes no sentido *downstream*. Filtragens foram realizadas para eliminar genes que codificam proteínas que fazem parte do T3SS, chaperonas, acessórios localizados na membrana e genes que se encontram na fita complementar. O outro *script* encontra T3SE com base no sinal de secreção presentes na extremidade N-terminal das proteínas Hop que são reconhecidas pelo T3SS (Hung *et al.*, 2014). As características previamente descritas utilizadas neste estudo foram a presença de $\geq 10\%$ de serina nos primeiros 50 aminoácidos, presença de isoleucina,

leucina, valina ou prolina na terceira ou quarta posição e ausência de ácido aspártico ou glutamato na 12^a posição na sequência da proteína analisada (Schechter *et al.*, 2012; Vencato *et al.*, 2006). Os resultados são apresentados em diagrama de Venn, usando a ferramenta Jvenn (<http://jvenn.toulouse.inra.fr/app/example.html>) (Bardou *et al.*, 2014).

RESULTADOS

Pseudomonas cichorii* possui um *hrp/hrc* cluster divergente dos patovares de *Pseudomonas syringae

Foi possível observar que os patovares de *P. syringae* possuem as sequências dos *hrp/hrc* clusters muito similares, enquanto a espécie *Pch* possui uma sequência mais divergente, sendo mais extensa, com uma grande inserção contendo 21 ORFs entre os genes *hrcC* e *hrcU*. Essa inserção se encontra entre as posições 7000 e 32000 das sequências ordenadas dos clusters *hrp/hrc*, sendo conservada dentro das cinco estirpes de *Pch* e ausentes nos patovares de *P. syringae* (Figura 1). Essa grande inserção corresponde a uma região enriquecida em proteínas reguladoras, fatores de transcrição e permeases do tipo ABC (Tabela Suplementar S1)

Apesar dos genes dos clusters *hrp/hrc* das espécies/patovares que infectam o cafeeiro terem a mesma anotação nos seus genomas do que os genes do cluster de *Pst* DC3000, alguns deles não apresentaram similaridade significativa (Figura 2). Assim, dos membros pertencentes ao cluster *hrp/hrc* de *Psgc*, apenas *hrpA*, *hrpD*, *hrpF*, *hrpG*, *hrcQa* e *hrpJ* não possuem identidade quando comparados com os genes do cluster de referência, enquanto os demais membros identificados possuem identidades que variam de 36 a 98%, não sendo identificados, na região genômica extraída com CLC, os genes *hrpT*, *hrcQb* e *hrpL* (Figura 2A e 2B).

Analisando a sequência de *Psta*, os membros *hrcJ*, *hrpG*, *hrcS*, *hrpP*, *hrpQ* e *hrpJ* não possuem similaridade significativa com os genes presentes no cluster *hrp/hrc* de *Pst* DC3000. Os genes *hrpD*, *hrpF*, *hrpT* e *hrcQb* não foram encontrados na região genômica

selecionada. Os genes do *cluster hrp/hrc* de *Psta*, localizados no estudo, possuem identidades que variam entre 41 a 94% quando comparados com os genes do *hrp/hrc cluster* de referência (Figura 2A e 2C).

Dentre os genes do *cluster hrp/hrc* de *Pch*, *hrpD*, *hrpJ*, *hrpG*, *hrcU*, *hrcT*, *hrcR*, *hrpP*, *hrcN* e *hrcV* não tiveram identidade com os respectivos membros do *cluster hrp/hrc* de *Pst*, enquanto *hrpR*, *hrpA*, *hrpV*, *hrcQa*, *hrpT*, *hrcQb* e *hrpL* não foram localizados no genoma de *Pch*. Os demais componentes identificados tiveram as menores identidades com os genes homólogos de *Pst* DC3000 quando comparados com *Psgc* e *Psta*, variando entre 27 a 58% (Figura 2A e 2D).

Diversidade de efetores novos e conhecidos entre *Pseudomonas* que infectam o cafeeiro

A busca de efetores nas cinco sequências genômicas de *Psgc*, baseada em similaridade de sequências de amino ácidos, permitiu a identificação de 37 a 48 possíveis efetores. Analisando a presença da caixa *hrp* na região promotora foi localizado de 21 a 28 efetores candidatos (Tabela Suplementar S2 e S3). No entanto, nestes grupos de possíveis efetores foram identificadas proteínas pertencentes à mesma família, assim, os resultados foram filtrados de modo que cada família de efector seja representada apenas por um membro. Adicionalmente, os resultados da busca por similaridade de sequências e presença de caixa *hrp* foram reunidos para obter o pansecretoma de tipo III de *Psgc*, utilizando as informações obtidas das cinco estirpes incluídas no estudo. Os resultados mostraram que *Psgc* possui efetores candidatos de 46 famílias, distribuídos entre as cinco estirpes avaliadas (Figura 3). Das 46 famílias de T3SE identificadas em *Psgc*, foram encontrados 13 possíveis novos efetores, através do *script* desenvolvido em Python, que analisa assinaturas referentes à caixa *hrp* na região promotora de genes regulados por HrpL.

A pesquisa de similaridade de sequências de aminoácidos, analisando as cinco sequências genômicas de *Psta*, localizou a presença de 24 a 30 efetores candidatos, enquanto analisando a presença da caixa *hrp* na região promotora identificou a presença

de 22 a 26 possíveis efetores (Tabela Suplementar S4 e S5). Após a retirada das repetições de efetores pertencentes à mesma família, os dados de similaridade e caixa *hrp* reunidos demonstraram que *Psta* possui 43 famílias de efetores candidatos, compartilhados entre as cinco estirpes analisadas (Figura 4). Das 43 famílias de efetores identificados em *Psta*, usando o *script* em Python, foi detectado a presença de 13 possíveis novos efetores, por meio da assinatura referente à caixa *hrp*.

Em *Pch*, a pesquisa de similaridade de sequências de aminoácidos nas cinco sequências genômicas, localizou a presença de três efetores candidatos, analisando a presença da caixa *hrp* na região promotora, identificou a existência de 5 a 11 possíveis efetores (Tabela Suplementar S6 e S7). Posterior a retirada das repetições de efetores pertencentes à mesma família, os dados agrupados de similaridade e caixa *hrp* demonstraram que *Pch* possui 14 famílias de efetores candidatos, distribuídos entre as cinco estirpes avaliadas (Figura 5). Entre as 14 famílias de efetores identificados em *Pch*, o *script* em Python detectou a presença de 10 possíveis novos efetores, com base na assinatura na região promotora dos genes regulados por HrpL.

Um efetor altamente conservado e importante para a patogenicidade é comum entre bactérias patogênicas ao cafeeiro

Comparações entre os pansecretomas de tipo III das diferentes espécies/patovares mostrou que das 46 famílias de efetores de *Psgc*, as famílias *avrB*, *avrPto*, *hopY*, *hopZ*, *hopAI*, *hopAU*, *hopBI*, *hopAJ*, *hopH*, *hopAF*, *hopBF* e 10 possíveis novos efetores foram exclusivos de *Psgc*. Em *Psta*, das 43 famílias de efetores encontradas, foram exclusivos os efetores *hopI*, *hopM*, *hopQ*, *hopAB*, *hopAZ* e 13 possíveis novos efetores. As famílias de efetores, *hopA*, *hopB*, *hopBN* e 10 candidatos a novos efetores foram exclusivos de *Pch*. Dos possíveis efetores de *Psgc*, três (*Psgc_ef07*, *Psgc_ef08* e *Psgc_ef13*) são compartilhados com *Psta*. Foi observado que as três espécies/patovares compartilham apenas uma família de efetores, identificado como *avrE* (Figura 6). As funções previamente relatadas para efetores encontrados no

pansecretoma do T3SS de *Psgc*, *Psta* e *Pch* são diversas e podem ser observadas na Tabela 2.

DISCUSÃO

No presente trabalho, os genes que constituem os *clusters* de genes *hrp/hrc*, e as proteínas efetoras secretadas pelos T3SS de patógenos bacterianos que causam manchas foliares no cafeeiro foram preditos. Analisando os *cluster* de genes *hrp/hrc* extraídos das sequências genômicas rascunho de estirpes representativas de *Pst*, *Psgc*, *Psta* e *Pch*, foi verificada a presença de homólogos da maioria dos genes do *cluster* de referência de *Pst* DC3000 nos *clusters* de *Psgc* e *Psta*, indicando uma alta conservação do T3SS entre os patovares de *P. syringae* analisados. Com relação ao *hrp/hrc cluster* de *Pst* DC3000, os *clusters* dos patovares *Psgc* e *Psta* apresentaram alta sintonia gênica e os genes mostraram alta similaridade de sequência. Quando comparado com o *cluster hrp/hrc* de *Pst* DC3000, além de apresentar uma menor conservação de genes, foi observada uma grande inserção no *cluster* de *Pch*, a qual está presente apenas nas estirpes dessa espécie. Esses resultados ratificam a ideia que o T3SS é um componente importante para a capacidade de espécies de *Pseudomonas* em causar doença nas plantas hospedeiras.

No entanto, estudos sobre os mecanismos moleculares pelos quais *Psgc* causa doença em plantas de café não foram reportados (Badel & Zambolim 2019), e o requerimento do T3SS para sua patogenicidade não foi claramente estabelecido. Devido ao fato que foi demonstrado que a estirpe 1202 de *Psgc* foi capaz de injetar quimeras entre um efector de *Hemileia vastatrix* e o sinal de secreção de tipo III de efetores bacterianos, acredita-se que *Psgc* possua um T3SS funcional (Maia *et al.*, 2017). Este estudo identificou em *Psgc* homólogos de genes presentes no *cluster hrp/hrc* de *Pst* DC3000, cujo T3SS é essencial para causar doença em plantas suscetíveis e HR em plantas resistentes (Collmer *et al.*, 2000).

Contudo, alguns genes do *cluster hrp/hrc* de *Psgc* não apresentaram similaridade significativa com genes do *cluster* de *Pst* DC3000, como *hrpL*, um regulador da expressão

dos genes *hrp/hrc* e *hop*, que não foi encontrado nas regiões genômicas selecionadas. A eficiência do *script* foi demonstrada pela sua capacidade de localizar genes que fazem parte dos *clusters* de genes *hrp/hrc*, assim como de encontrar genes que codificam Hops, proteínas acessórias e chaperonas preditos e caracterizados funcionalmente em estudos prévios (dados não demonstrados) (Collmer *et al.*, 2000; Xiao & Hutcheson 1994; Grimm *et al.*, 1995; Hutcheson *et al.*, 2001; Zwiesler-Vollick *et al.*, 2002). Genes homólogos de *hrpL* frequentemente estão presentes entre *hrcV* e *hrpK* em diversos patovares de *P. syringae*, como observado nos *clusters* de genes *hrp/hrc* de *Pst* e *Psta* neste estudo. Análises mais detalhadas com a utilização de genomas completos poderiam confirmar que genes tão importantes como o *hrpL* de fato não estão presentes em *Psgc*. A presença de genes homólogos de *hrpR*, *hrpS* e *hrcV* (Xiao & Hutcheson, 1994; Grimm *et al.*, 1995; Hutcheson *et al.*, 2001; Zwiesler-Vollick *et al.*, 2002), que são reguladores globais dos genes *hrp/hrc* sugere que o sistema de secreção de *Psgc* apresenta similaridade funcional com aquele de *Pst* DC3000.

Alta sintonia génica foi observada entre os *clusters hrp/hrc* dos diferentes patovares de *P. syringae*. Entretanto, inversões foram encontradas nos genes *hrpA*, *hrpB*, *hrpF*, *hrpG*, *hrpV*, *hrpL* e deleções dos genes *hrpD*, *hrpF*, *hrpT*, *hrpQb* e *hrpL* em pelo menos um dos patovares, o que sugere que os *clusters* têm sido sujeitos a rápidas alterações durante a história evolutiva. A implicação desses rearranjos genômicos na função do T3SS é desconhecida. Nossos resultados são consistentes com relatos prévios onde foi determinado que *hrcV*, que codifica para uma proteína do T3SS que se localiza na membrana interna, é um dos genes mais conservados do *cluster* de genes *hrp/hrc* (Cornelis and Van Gijsegem, 2000; Stuber *et al.*, 2003).

Como em *Psgc*, acredita-se que o T3SS de *Psta* é funcional e importante para sua interação com o hospedeiro, baseado nas seguintes evidências: (1) a expressão do efetor *hopA1* (antigamente *hopPsyA* ou *hrmA*) tornou a estirpe *Psta* 11528 avirulenta em *Nicotiana tabacum* cv. Xanthi, que é suscetível à estirpe de tipo selvagem (Alfano *et al.* 1997); (2) a expressão do effector *hopQ1-1* na mesma estirpe a tornou incompatível em *Nicotiana benthamiana* (Wei *et al.*, 2007); (3) um mutante da estirpe 6506 de *Psta* com uma deleção no gene *hrcC* (codifica para uma proteína ancorada na membrana) não causou HR em *Arabidopsis* (Zhang *et al.*, 2010); e (4) enquanto em *Psta* 11528, de tipo

selvagem, ativou as respostas de defesa de *N. benthamiana*, uma estirpe com uma mutação em *hrcV* não conseguiu, provavelmente devido a não secretar efetores reconhecidos pela planta (Oh & Collmer, 2005).

Os resultados deste estudo demonstram que o *cluster* de genes *hrp/hrc* de *Pch* é divergente daqueles dos patovares de *P. syringae*. Essa divergência se reflete tanto no número de genes conservados quanto na baixa similaridade dos genes, onde a máxima similaridade observada foi de 58% para *hrpS*. Somente nove dos 18 genes do *cluster hrp/hrc* extraídos dos genomas rascunho apresentam similaridade significativa com aqueles do *cluster hrp/hrc* de *Pst* DC3000.

Em estudos prévios, foi demonstrado que a função dos genes *hrcG*, *hrcT*, *hrpL* e *hrpS* em *Pch* SPC9018 são necessários para causar doenças em plantas de berinjela (*Solanum melongera*), salsação (*Apium graveolens*) e quiabo (*Abelmoschus esculentus*), porém aparentemente não são essenciais para a sua patogenicidade em alguns hospedeiros como a alface (*Lactuca sativa* L. cv. Success), indicando que *Pch* possui diferentes mecanismos para causar doença em diferentes hospedeiros (Badel & Zambolim 2019; Hojo *et al.*, 2008; Ramkumar *et al.*, 2015). Essas observações confirmam que é necessário desenvolver estudos funcionais, criando mutantes em genes essenciais para a função do T3SS em estirpes coletadas do cafeeiro e testar a sua capacidade de causar doença nessa espécie hospedeira.

A proximidade filogenética entre os patógenos bacterianos que causam mancha foliar no cafeeiro e a similaridade dos sintomas causados, sugere a possibilidade que mecanismos comuns possam ser utilizados por esses patógenos. Os resultados deste estudo, relacionados ao pansecretoma do T3SS das *Pseudomonas* patogênicas ao cafeeiro relevou a presença de 27 efetores *core* em *Psgc*, formado por oito possíveis novos efetores encontrados através da presença de caixas *hrp* nas regiões promotoras dos genes. Em *Psta*, foram encontrados 28 efetores *core*, constituídos por nove novos possíveis efetores, e em *Pch*, dos cinco efetores *core* encontrados, três são possíveis novos efetores.

Foi possível observar que *Pch* possui um pansecretoma do T3SS menor do que aqueles de *Psgc* e *Psta*. O número de efetores de uma espécie/patovar bacteriano é altamente variável, segundo a estirpe (Badel & Zambolim, 2019). Uma única estirpe de

P. syringae tipicamente pode codificar mais de 30 efetores diferentes (Lindeberg *et al.*, 2008). Contudo, diferentes estirpes de espécies/patovares de *Pseudomonas* podem exibir diversos conjuntos de efetores. Assim, uma determinada espécie/patovar de *Pseudomonas* exibe um conjunto de efetores fixos, presente em todas as estirpes, e que possuem papel fundamental no processo da doença (Studholme *et al.*, 2002). Efetores específicos aparentemente têm efeitos redundantes com outros efetores do *core* e são individualmente dispensáveis, com perdas pequenas ou sem efeito detectável na virulência do patógeno quando modificados (Studholme *et al.*, 2002; Grant *et al.*, 2006).

Este trabalho detectou que apenas genes da família de efetores *avrE* estiveram presentes em todas as espécies/patovares de *Pseudomonas* que infectam plantas de café. Genes da família *avrE* têm sido encontrados em todas as bactérias de importância agrícola que possuem um T3SS, incluindo membros das classes betaproteobacteria (*Ralstonia*) e gammaproteobacteria (*Xanthomonas* e *Pseudomonas*). Vários estudos já relataram que alguns genes da família *avrE* contribuem significativamente para a virulência em plantas hospedeiras (Degrave *et al.*, 2015), sendo requeridos no início e durante o processo infeccioso, conferindo a capacidade de supressão das respostas da planta dependentes de ácido salicílico, como deposição de calose ou expressão dos genes relacionados a patogênese 1 (*Pathogenesis Related 1; PR1*) (Boreau *et al.*, 2006; Deng *et al.*, 2003; Hogan *et al.*, 2013; Kim *et al.*, 2011; Pester *et al.*, 2012). A mutação de genes que codificam proteínas da família AvrE pode variar desde a diminuição da virulência (Badel *et al.* 2006) até a perda da patogenicidade (Bogdanove *et al.*, 1998).

No presente estudo, tentou-se identificar T3SE com base no sinal de secreção presente na porção N-terminal. Diversas abordagens foram utilizadas para prever os efetores com base nessa característica, incluindo o uso de um *script* customizado em Python, e os programas disponíveis on line pEffect (<https://services.bromberglab.org/peffect/>) (Goldberg *et al.*, 2016) e EffectiveDB (<https://effectors.csb.univie.ac.at/>) (Eichinger *et al.*, 2016). Um número exagerado de efetores (> 500) foi predito utilizando cada abordagem. Acredita-se que os modelos baseados na identificação dos sinais de secreção do tipo III não estão ainda corretamente acurados. Portanto, os resultados dessas análises foram desconsiderados.

Por fim, a descrição dos pansecretomas de tipo III de *Psgc*, *Psta* e *Pch* neste estudo servem como ponto de partida para trabalhos futuros visando a caracterização funcional dos efetores identificados. O conhecimento do repertório dos efetores das espécies/patovares patogênicas ao cafeeiro pode auxiliar na compreensão dos mecanismos que estas utilizam para subjugar o hospedeiro. Assim, este trabalho indica que o efector *avrE* pode possuir papel fundamental nas interações dessas bactérias com o cafeeiro. Por outro lado, a descoberta de novos efetores pode ajudar a entender como cada espécie/patovar se adaptou ao hospedeiro e as características exclusivas de cada patossistema. Trabalhos que contribuam para melhor entender a função e a diversidade genética associada ao repertório do T3SS das *Pseudomonas* patogênicas ao cafeeiro são de extrema importância para o desenvolvimento de estratégias que visem o controle eficiente das doenças causadas por essas bactérias.

REFERÊNCIAS

Alfano JR, Charkowski AO, Deng W-L, Badel JL, Petnicki-Ocwieja T, Van DK, Collmer A, 2000. The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. *Proceedings of the National Academy of Sciences of the U. S. A.* **9**, 4856-4861.

Alfano JR, Kim H-S, Delaney TP, Collmer A, 1997. Evidence that the *Pseudomonas syringae* pv. *syringae* hrp-linked *hrmA* gene encodes an Avr-like protein that acts in an hrp-dependent manner within tobacco cells. *Molecular Plant-Microbe Interactions* **10**, 580-588.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, 1990. Basic Local Alignment search tool. *Journal of Molecular Biology* **215**, 403-4010.

Amaral JF, Teixeira C, Pinheiro EDA, 1956. Bactéria causadora da mancha aureolada do cafeeiro. *Arquivos do Instituto Biológico* **23**, 151-155.

Araki H, Innan H, Kreitman M, Bergelson H, 2007. Molecular evolution of pathogenicity-island genes in *Pseudomonas viridiflava*. *Genetics* **2**, 1031-1041.

Badel JL, Piquerez SJM, Greenshields D, Rallapalli G, Fabro G, Ishaque N, Jones JDG, 2013. In *Planta* Effector Competition Assays detect *Hyaloperonospora arabidopsidis* effectors that contribute to virulence and localize to different plant subcellular compartments. *Molecular Plant-Microbe Interactions* **7**, 745-757.

Badel JL, Shimizu R, Oh H-S, Collmer A, 2006. A *Pseudomonas syringae* pv. *tomato* *avrE1/hopM1* mutant is severely reduced in growth and lesion formation in tomato. *Molecular Plant-Microbe Interactions* **19**, 99-111.

Badel JL, Zambolim L, 2019. Coffea bacterial diseases: a plethora of scientific opportunities. *Plant Pathology* **68**, 411- 425.

Bardou P, Mariette J, Escudie F, Djemiel C, Klopp C, 2014. Jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* **15**, 1-7.

Bhattacharjee S, Hiller NL, Liolios K, Win J, Kanneganti T-D, Young C, Kamoun S, Haldar K, 2006. The malarial host-targeting signal is conserved in the irish potato famine pathogen. *Plos Pathogens* **5**, 453-465.

Bogdanove AJ, Kim JF, Wei Z, Kolchinsky P, Charkowski AO, Conlin AK, Collmer A, Beer SV, 1998. Homology and functional similarity of an *hrp*-linked pathogenicity locus, *dspEF*, of *Erwinia amylovora* and the avirulence locus *avrE* of *Pseudomonas syringae* pathovar *tomato*. *Proceedings of the National Academy of Sciences of the U. S. A.* **95**, 1325-1330.

Boureau T, ElMaarouf-Bouteau H, Garnier A, Brisset M-N, Perino C, Pucheu I, Barny M-A, 2006. DspA/E, a type III effector essential for *Erwinia amylovora* pathogenicity and growth *in planta*, induces cell death in host apple and nonhost tobacco plants. *Molecular Plant-Microbe Interactions* **19**, 16-24.

Castaneda-Ojeda MP, López-Solanilla E, Ramos C, 2017. Differential modulation of plant immune responses by diverse members of the *Pseudomonas savastanoi* pv. *savastanoi* HopAF type III effector family. *Molecular Plant Pathology* **5**, 625-634.

Collmer A, Badel JL, Charkowski AO, Deng W-L, Fouts DE, Ramos AR, Rehm AH, Anderson DM, Scheewind O, van Dijk K, Alfano JR, 2000. *Pseudomonas syringae* Hrp type III secretion system and effector proteins. *Proceedings of the National Academy of Sciences of the U. S. A.* **97**, 8770-8777.

Cornelis GR, Van Gijsegem F, 2000. Assembly and function of type III secretory systems. *Annual Review of Microbiology* **54**, 735-774.

Cui H, Wang Y, Xue L, Chu J, Yan C, Fu J, Chen M, Innes RW, Zhou J-M, 2010. *Pseudomonas syringae* effector protein AvrB perturbs *Arabidopsis* hormone signaling by activating MAP kinase 4. *Cell Host and Microbe* **7**, 164-175.

DaMatta FH, Ramalho JDC, 2004. Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology* **16**, 55-81.

Degrave A, Moreau M, Launay A, Barny M-A, Brisset M-N, Patrit O, Taconnat L, Vedel R, Fagard M, 2013. The bacterial effector DspA/E is toxic in *Arabidopsis thaliana* and is required for multiplication and survival of fire blight pathogen: DspA/E is required for *Ea* growth in *A. thaliana*. *Molecular Plant Pathology* **14**, 506-517.

Degrave A, Siamer S, Boureau T, Barny M-A, 2015. The AvrE superfamily: ancestral type III effectors involved in suppression of pathogen-associated molecular pattern-triggered immunity. *Molecular Plant Pathology* **16**, 899-905.

Deng W-L, Rehm AH, Charkowski AO, Rojas CM, Collmer A, 2003. *Pseudomonas syringae* Exchangeable Effector Loci: sequence diversity in representative pathovar and

virulence function in *P. syringae* pv. *syringae* B728a. *Journal of Bacteriology* **8**, 2592-2602.

Dou D, Kale SD, Wang X, Jiang RHY, Bruce NA, Arredondo FD, Zhang X, Tyler BM, 2008. RXLR-mediated entry of *Phytophthora sojae* effector *Avr1b* into soybean cells does not require pathogen-encoded machinery. *The Plant Cell* **20**, 1930-1947.

Eichinger V, Nussbaumer T, Platzer A, Jehl M-A, Arnold R, Rattei T, 2016. EffectiveDB- updates and novel features for a better annotation of bacterial secreted proteins and type III, IV, VI secretion systems. *Nucleic Acids Research* **44**, 669-674.

Ferrante P, Clarke CR, Cavanaugh KA, Michelmore RW, Buonauro R, Vinatzer BA 2009. Contributions of the effector gene *hopQ1-1* to differences in host range between *Pseudomonas syringae* pv. *phaseolicola* and *P. syringae* pv. *tabaci*. *Molecular Plant Pathology* **10**, 837–842.

Gohre V, Robatzek S, 2008. Breaking the barriers: microbial effector molecules subvert plant immunity. *Annual Review of Phytopathology* **46**, 189-215.

Goldberg T, Rost B, Bromberg Y, 2016. Computational prediction shines light on type III secretion origins. *Scientific Reports* **6**, 1-10.

Grant SR, Fisher EJ, Chang JH, Mole BM, Dangl JL, 2006. Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria. *Annual Review of Microbiology* **60**, 425-449.

Grimm C, Aufsatz W, Panopoulos NJ, 1995. The *hrpRS* locus of *Pseudomonas syringae* pv. *phaseolicola* constitutes a complex regulatory unit. *Molecular Microbiology* **15**, 155–165.

Hogan CS, Mole BM, Grant SR, Willis DK, Charkowski AO, 2013. The type III secreted effector DspE is required early in *Solanum tuberosum* leaf infection by *Pectobacterium carotovorum* to cause cell death, and requires Wx(3-6)D/E motifs. *Plos ONE* **8**, e65534.

Hojo H, Koyanagi M, Tanaka M, Kajihara S, Ohnishi K, Kiba A, Hikichi Y, 2008. The *hrp* genes of *Pseudomonas cichorii* are essential for pathogenicity on eggplant but not on lettuce. *Microbiology* **154**, 2920-2928.

Hulin MT, Mansfield Jw, Brain P, Xu X, Jackson RW, Harrison RJ, 2008. Characterisation of the pathogenicity of strains of *Pseudomonas syringae* towards cherry and plum. *Plant Pathology* **67**, 1177-1193.

- Hung NB, Ramkumar G, Lee YH, 2014. An effector gene *hopA1* influences on virulence, host specificity, and lifestyles of *Pseudomonas cichorii* JBC1. *Research in Microbiology* **165**, 620–629.
- Hutcheson SW, Bretz J, Sussan T, Park K, 2001. Enhancer-binding proteins HrpR and HrpS interact to regulate *hrp*-encoded type III protein secretion in *Pseudomonas syringae* strains. *Journal of Bacteriology* **183**, 5589-5598.
- Jelenska J, Yao N, Vinatzer BA, Wright CM, Brodsky JL, Greenberg JT, 2007. A J domain virulence effector of *Pseudomonas syringae* remodels host chloroplasts and suppresses defenses. *Current Biology* **6**, 499-508.
- Jin QL, He SY, 2001. Role of the Hrp pilus in type III secretion in *Pseudomonas syringae*. *Science* **294**, 2556–2558.
- Jones JD, Dangl JL, 2006. The plant immune system. *Nature* **444**, 323-329.
- Kajihara S, Hojo H, Koyanagi M, 2012. Implication of *hrpW* in virulence of *Pseudomonas cichorii*. *Plant Pathology* **61**, 335-363.
- Kim HS., Thammarat P, Lommel SA, Hogan CS, Charkowski AO, 2011. *Pectobacterium carotovorum* elicits plant cell death with DspE/F but the *P. carotovorum* DspE does not suppress callose or induce expression of plant genes early in plant–microbe interactions. *Molecular Plant–Microbe Interactions* **24**, 773–786.
- Klement Z, Goodman RN, 1967. The hypersensitive reaction to infection by bacterial plant pathogens. *Annual Review of Phytopathology* **5**, 17-44.
- Lindeberg M, Myers CR, Collmer A, Schneider DJ, 2008. Roadmap to new virulence determinants in *Pseudomonas syringae*: insights from comparative genomics and genome organization. *Molecular Plant-Microbe Interactions* **21**, 685-700.
- Lindeberg M, Stavrinides J, Chang JH, Alfano JR, Collmer A, Dangl JL, Greenberg JT, Mansfield JW, Guttman DS, 2005. Proposed guidelines for a unified nomenclature and phylogenetic analysis of type III Hop effector proteins in the plant pathogen *Pseudomonas syringae*. *Molecular Plant-Microbe Interactions* **18**, 275-282.
- Lorang JM, Keen NT, 1995. Characterization of *avrE* from *Pseudomonas syringae* pv. *tomato*: a *hrp*-linked avirulence locus consisting of at least two transcriptional units. *Molecular Plant–Microbe Interactions* **8**, 49–57.

Maia T, Badel JL, Ramirez-Marin G, Rocha CM, Fernandes MB, Silva JCF, Azevedo-Junior GM, Brommonschenkel SH, 2017. The *Hemileia vastatrix* effector HvEC-016 suppresses bacterial blight symptoms in coffee genotypes with the *S_H1* rust resistance gene. *New Phytologist* **213**, 1315-1329.

Matas IM, Castanêda-Ojeda MP, Aragón IM, Antúnez-Lamas M, Murillo J, Rodríguez-Palenzuela P, López-Solanilla E, Ramos C, 2014. Translocation and functional analysis of *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 type III secretion system effectors reveals two novel effector families of the *Pseudomonas syringae* complex. *Molecular Plant-Microbe Interactions* **5**, 424-436.

Oh H-S, Collmer A, 2005. Basal resistance against bacteria in *Nicotiana benthamiana* leaves is accompanied by reduced vascular staining and suppressed by multiple *Pseudomonas syringae* type III secretion system effector proteins. *The Plant Journal* **44**, 348-359.

Oh H-S, Kvitko BH, Morello JE, Collmer A, 2007. *Pseudomonas syringae* lytic transglycosylases coregulated with the type III secretion system contribute to the translocation of effector proteins into plant cells. *Journal of Bacteriology* **22**, 8277-8289.

Pester D, Milcevicová R, Schaffer J, Wilhelm E, Blümel S, 2012. *Erwinia amylovora* expresses fast and simultaneously *hrp/dsp* virulence genes during flower infection on apple trees. *PLoS ONE* **7**, e32583.

Ramkumar G, Lee SW, Weon H, Kim B, Lee YH, 2015. First report on the whole genome sequence of *Pseudomonas cichorii* strain JBC1 and comparison with other *Pseudomonas* species. *Plant Pathology* **64**, 63-70.

Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT, 2009. Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics* **25**, 2071-2073.

Rodrigues LMR, Almeida IMG, Patrício FRA, 2013. Mancha Aureolada do Cafeeiro Causada por *Pseudomonas syringae* pv. *garcae*. Campinas, SP, Brasil: Instituto Agrônômico de Campinas.

Rodrigues Neto J, Silva CHD, Beriam LOS, 2006. Mancha bacteriana do cafeeiro causada por *Pseudomonas syringae* pv. *tabaci*. *Summa Phytopathologica* **32**, 85.

Roine E, Wei W, Yuan J, Nurmiäho-Lassila EL, Kalkkinen N, Romantschuk M, He SY, 1997. Hrp pilus: an *hrp*-dependent bacterial surface appendage produced by

Pseudomonas syringae pv. *tomato* DC3000. *Proceedings of the National Academy of Sciences of the U. S. A.* **94**, 3459–3464.

Schechter LM, Valenta JC, Schneider DJ, Collmer A, Sakk E, 2012. Functional and computational analysis of amino acid patterns predictive of type III secretion system substrates in *Pseudomonas syringae*. *Plos ONE* **7**, e36038.

Shen H, Keen NT, 1993. Characterization of the promoter of avirulence gene D from *Pseudomonas syringae* pv. *tomato*. *Journal of Bacteriology* **18**, 5916-5924.

Stuber K, Frey J, Burnens AP, Kuhnert P, 2003. Detection of type III secretion genes as a general indicator of bacterial virulence. *Molecular and Cellular Probes*, **17**, 25–32.

Studholme DJ, Ibanez SG, Maclean D, Dangl JL, Chang JH, Rathjen JP, 2002. A draft genome sequence and functional screen reveals the repertoire of type III secreted proteins of *Pseudomonas syringae* pathovar *tabaci* 11528. *BMC Genomics* **10**, 395.

Vencato M, Tian F, Alfano JR, Buell CR, Cartinhour S, DeClerck GA, Guttman DS, Stavrinides J, Joardar V, Lindeberg M, Bronstein PA, Mansfield JW, Myers CR, Collmer A, Schneider DJ, 2006. Bioinformatics-enabled identification of the HrpL regulon and type III secretion system effector proteins of *Pseudomonas syringae* pv. *phaseolicola* 1448A. *Molecular Plant-Microbe Interactions* **19**, 1193-1206.

Wei C-F, Kvitko BH, Shimizu R, Crabill E, Alfano JR, Lin N-C, Martin GB, Huang H-C, Collmer A, 2007. A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*. *The Plant Journal* 51:32-46.

Wei H-L, Chakravarthy S, Mathieu J, Swingle B, Martin GB, Collmer A, 2015. *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion effector polymutants reveal an interplay between HopAD1 and AvrPtoB. *Cell Host and Microbe* **6**, 752-762.

Xiang T, Zong N, Zou Y, Wu Y, Zhang J, Xing W, Li Y, Tang X, Zhu L, Chai J, Zhou JM, 2008. *Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases. *Current Biology* **18**, 74–80.

Xiao Y, Hutcheson S, 1994. A single promoter sequence recognized by a newly identified alternate sigma factor directs expression of pathogenicity and host range determinants in *Pseudomonas syringae*. *Journal of Bacteriology* **176**, 3089–3091.

Zhang J, Shao F, Li Y, Cui H, Chen L, Li H, Zou Y, Long C, Lan L, Chai J, Chen S, Tang X, Zhou JM, 2007. A *Pseudomonas syringae* effector inactivates MAPKs to suppress PAMP-induced immunity in plants. *Cell Host and Microbe* **1**, 175–185.

Zhang H, Lu H, Li X, Li Y, Cui H, Wen C-K, Tang X, Su Z, 2010. Effector-triggered and pathogen-associated molecular pattern-triggered immunity differentially contribute to basal resistance to *Pseudomonas syringae*. *Molecular Plant-Microbe Interactions* **23**, 940-948.

Zhou H, Lin J, Johnson A, Morgan RL, Zhong W, Ma W, 2011. *Pseudomonas syringae* type III effector HopZ1 targets a host enzyme to suppress isoflavone biosynthesis and promote infection in soybean. *Cell Host and Microbe* **3**, 177-186.

Zipfel C, Robatzek S, 2010. Pathogen-associated molecular pattern-triggered immunity. *Plant Physiology* **154**, 551-554.

Zwiesler-Vollick J, Plovanich-Jones AE, Nomura K, Brandyopadhyay S, Joardar V, Kunkel BN, He SY, 2002. Identification of novel *hrp*-regulated genes through functional genomic analysis of the *Pseudomonas syringae* pv. *tomato* DC3000 genome. *Molecular Microbiology* **45**, 1207-1218.

FIGURAS



Figura 1. Alinhamento em Mauve das seqüências dos *clusters* de genes *hrp/hrc*, selecionadas por CLC, de *Pseudomonas syringae* pv. *tomato* (DC3000), *P. syringae* pv. *garcae* (ASMN341635), *P. syringae* pv. *tabaci* (ASMN341614) com cinco seqüências de *Pseudomonas cichorii* (ASMN341675, ASMN341577, ASMN341621, ASMN341610).

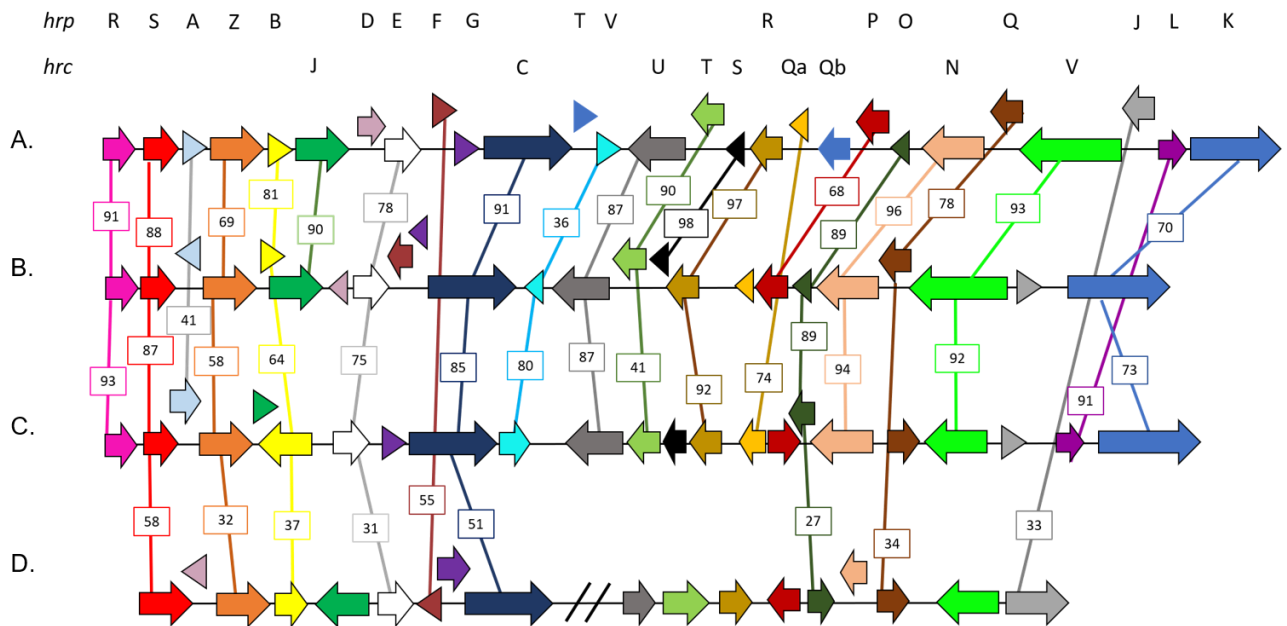


Figura 2. Sequências de *clusters* de genes *hrp/hrc* selecionadas em CLC. (A) *Pseudomonas syringae* pv. *tomato*, sequência de referência. (B) *P. syringae* pv. *garcae*. (C) *P. syringae* pv. *tabaci*. (D) *Pseudomonas cichorii*. Cada ORF é representada por uma seta, as setas com as mesmas cores obtiveram a mesma anotação nas sequências genômicas. Setas ligadas entre linha identificam os genes que mostraram similaridade significativa com genes da sequência de referência. Os números identificam a porcentagem de identidade com relação aos genes do *hrp/hrc cluster* do genoma de referência.

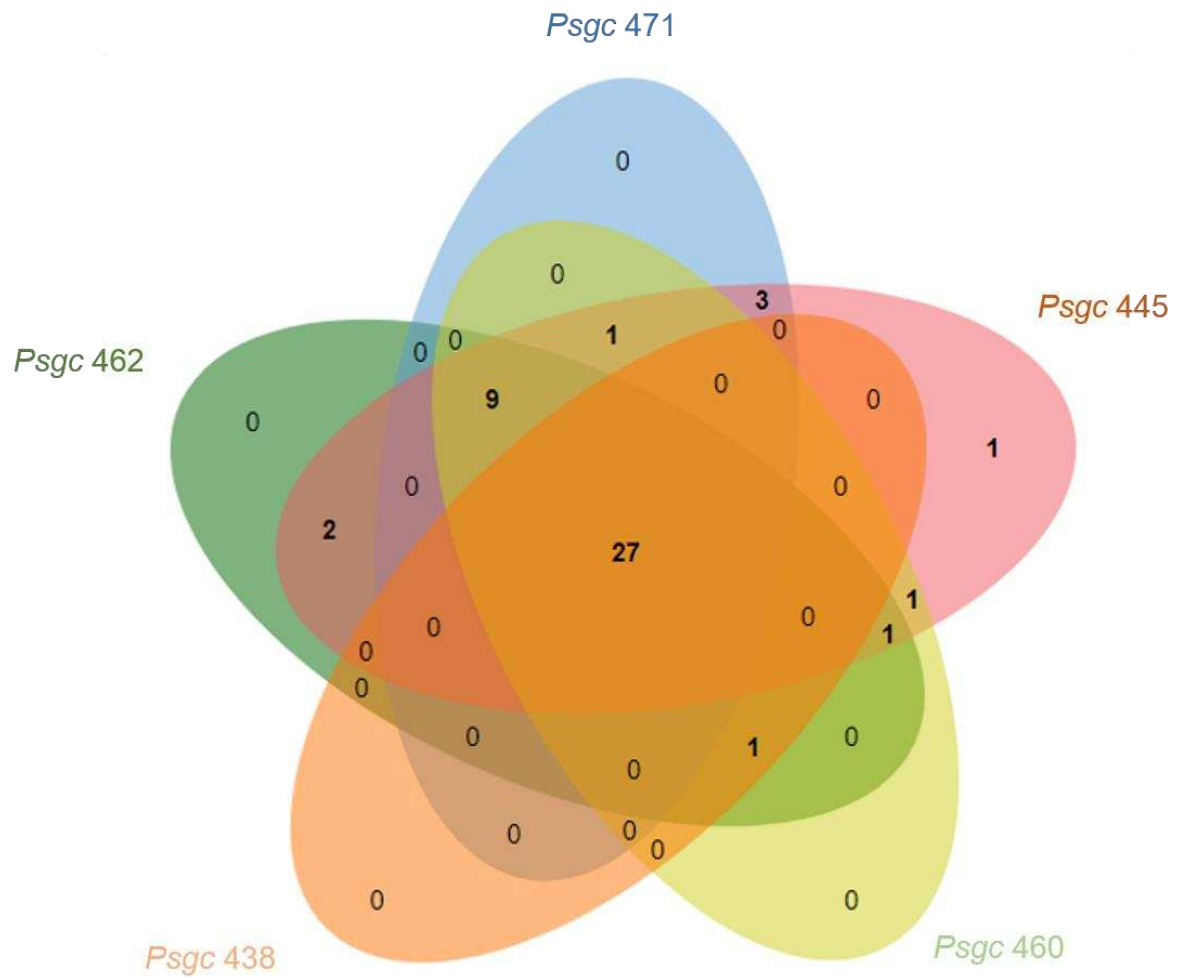


Figura 3. Diagrama de Venn do pansecretoma do T3SS de *Pseudomonas syringae* pv. *garcae* obtidos mediante Blastp com o banco de dados depositados no *P. syringae* Genome Resource e do *script* customizado em Python com base na presença de caixa *hrp* na região promotora. O número de efetores de cada estirpe é agrupado por uma cor diferente. Os números nas interserções indicam os efetores que são compartilhados entre as estirpes.

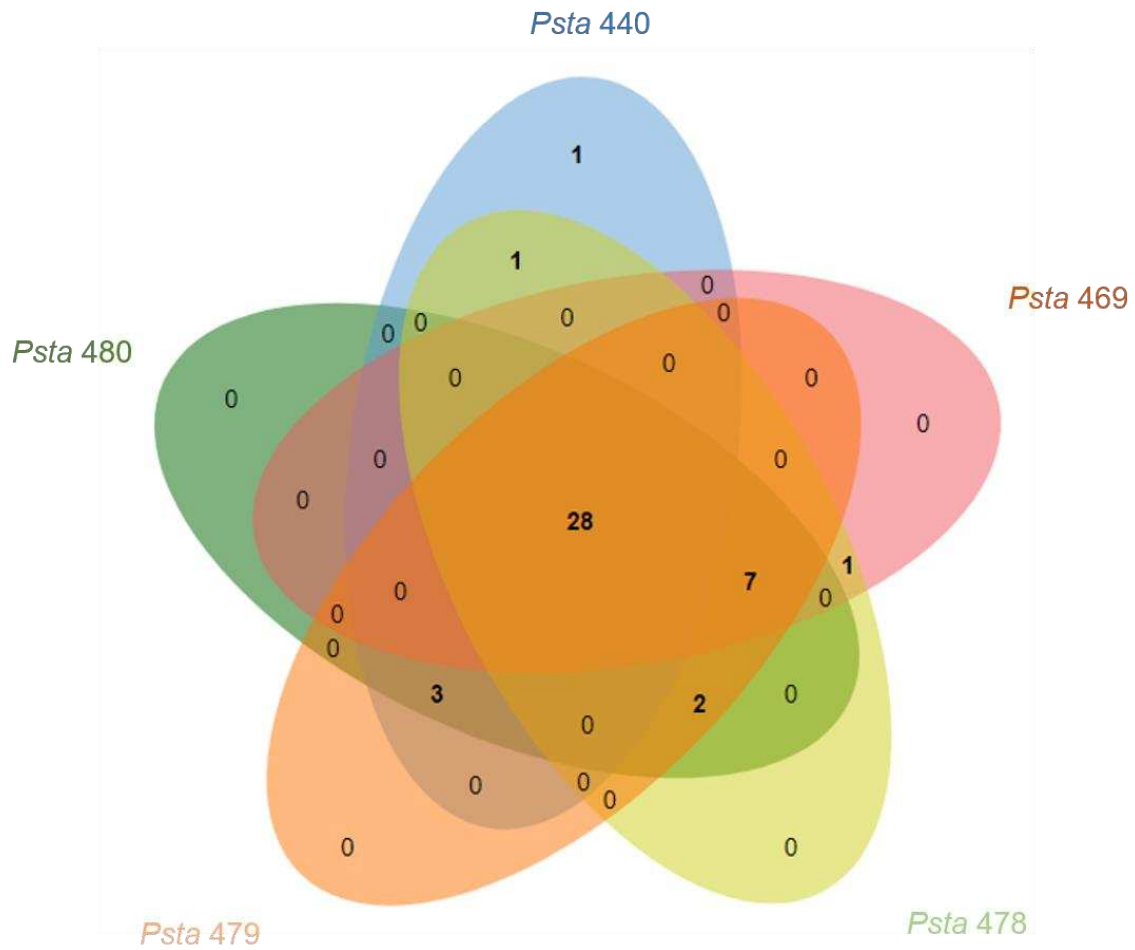


Figura 4. Diagrama de Venn do pansecretoma do T3SS de *Pseudomonas syringae* pv. *tabaci* obtidos mediante Blastp com o banco de dados depositados no *P. syringae* Genome Resource e do *script* customizado em Python com base na presença de caixa *hrp* na região promotora. O número de efetores de cada estirpe é agrupado por uma cor diferente. Os números nas interserções indicam os efetores que são compartilhados entre as estirpes.

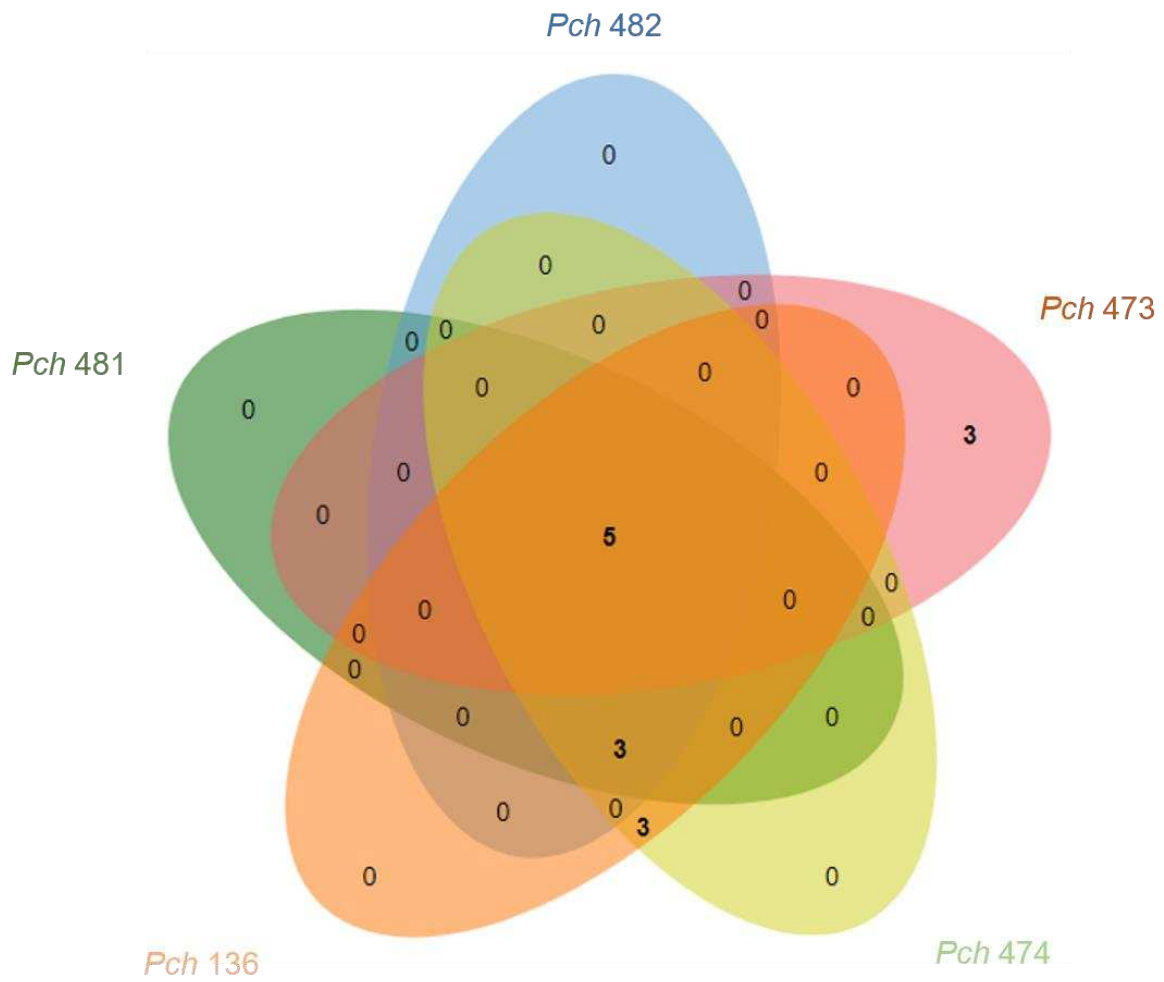


Figura 5. Diagrama de Venn do pansecretoma do T3SS de *Pseudomonas cichorii* obtidos mediante Blastp com o banco de dados depositados no *P. syringae* Genome Resource e do *script* customizado em Python com base na presença de caixa *hrp* na região promotora. O número de efetores de cada estirpe é agrupado por uma cor diferente. Os números nas interseções indicam os efetores que são compartilhados entre as estirpes.

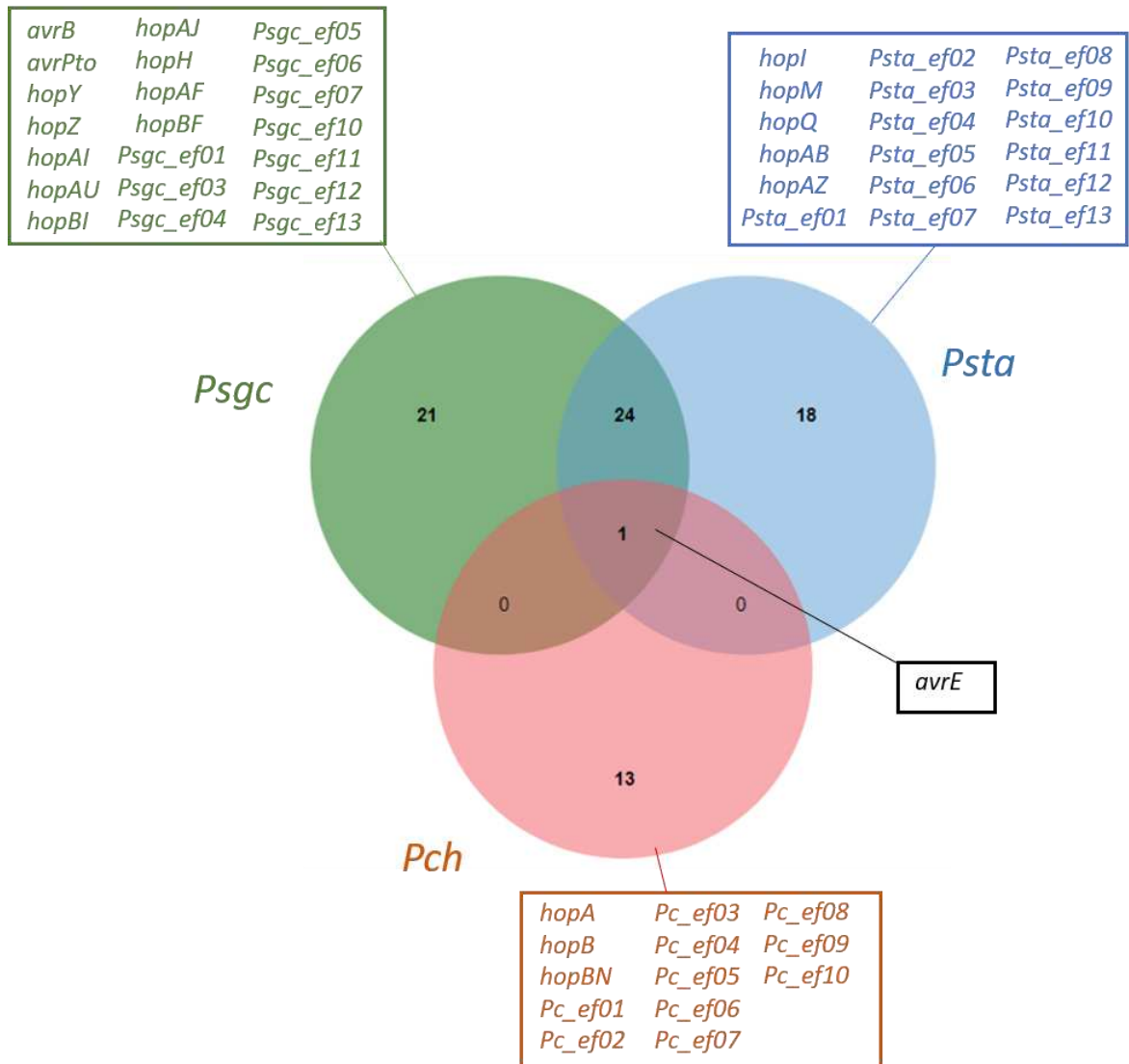


Figura 6. Diagrama de Venn demonstrando os efetores dos pansecretomas compartilhados e exclusivos de *Pseudomonas syringae* pv. *garcae*, *P. syringae* pv. *tabaci* e *Pseudomonas cichorii*. O número de efetores de cada espécie/patovar é agrupado por uma cor diferente. Os números nas interseções indicam os efetores que são compartilhados entre duas ou três espécies/patovares.

TABELAS

Tabela 1. Sequências da caixa *hrp* utilizadas pelo *script* em Python para identificação de efetores de tipo III.

Sequência	Referência
GGAACC-N ₁₅ -CCAC-2-A	Hojo <i>et al.</i> (2008)
GGAACC-N ₁₆ -CCAC-2-A	Hojo <i>et al.</i> (2008)
GGAACT-N ₁₅ -CCAC-2-A	Vencato <i>et al.</i> (2006)
GGAACT-N ₁₆ -CCAC-2-A	Vencato <i>et al.</i> (2006)
GGAATT-N ₁₅ -CCAC-2-A	Araki <i>et al.</i> (2007)
GGAATT-N ₁₆ -CCAC-2-A	Araki <i>et al.</i> (2007)
GGAACC-N ₁₅ -ACAC-2-A	Shen & Keen (1993)
GGAACC-N ₁₇ -CCAC-2-A	Xiao & Hutcheson (1994)

Tabela 2. Efeitores encontrados no pansecretoma do T3SS de *Pseudomonas syringae* pv. *garcae*, *P. syringae* pv. *tabaci* e *Pseudomonas cichorii*.

Efeto	Função	Patógeno	Hospedeiro	Referência
<i>avrB</i>	Perturbação hormonal e aumento de susceptibilidade	<i>P. syringae</i>	<i>Arabidopsis</i>	Cui <i>et al.</i> (2010)
	Supressão da via do ácido salicílico	<i>E. amylovora</i>	<i>Malus domestica</i>	Boureau <i>et al.</i> (2006)
	Reação de hipersensibilidade em não hospedeiros	<i>E. amylovora</i>	<i>Nicotiana tabacum</i>	Boureau <i>et al.</i> (2006)
<i>avrE</i>	Morte celular por toxicidade	<i>P. syringae</i> pv. <i>glycinea</i>	<i>Glycine max</i>	Lorang <i>et al.</i> (1995)
	Promoção de crescimento bacteriano <i>in planta</i>	<i>P. carotovorum</i>	<i>Nicotiana benthamiana</i>	Kim <i>et al.</i> (2011)
	Inibição de síntese proteica	-	<i>A. thaliana</i>	Degrave <i>et al.</i> (2013)
<i>avrF</i>	Chaperona	<i>P. viridiflava</i>	-	Araki <i>et al.</i> (2007)
<i>avrPto</i>	Bloqueio das respostas do sistema imune	<i>P. syringae</i>	<i>Solanum lycopersicum</i>	Xiang <i>et al.</i> (2008)
<i>hopA</i>	Virulência e especificidade de hospedeiro	<i>P. cichorii</i>	<i>Solanum lycopersicum</i>	Hung <i>et al.</i> (2014)
<i>hopAB</i>	Redundante do efeto <i>avrPto</i>	-	-	Hulin <i>et al.</i> (2008)
<i>hopAE</i>	Redução dos níveis de calose	<i>P. savastanoi</i> pv. <i>savastanoi</i>	<i>Nicotiana tabacum</i>	Castaneda-Ojeda <i>et al.</i> (2017)
	Inibidor da reação de hipersensibilidade	<i>P. savastanoi</i> pv. <i>savastanoi</i>	<i>Nicotiana tabacum</i>	Castaneda-Ojeda <i>et al.</i> (2017)
<i>hopAI</i>	Inativação da MAP-quinase	<i>P. syringae</i>	<i>Arabidopsis</i>	Zhang <i>et al.</i> (2007)
<i>hopAJ</i>	Translocação de efetores	<i>P. syringae</i>	<i>Solanum lycopersicum</i>	Oh <i>et al.</i> (2007)
<i>hopAU</i>	-	-	-	-
<i>hopAZ</i>	Inibidor de ETI	<i>P. savastanoi</i> pv. <i>savastanoi</i>	<i>Nicotiana tabacum</i>	Matas <i>et al.</i> (2014)
	Supressão da deposição de calose	<i>P. savastanoi</i> pv. <i>savastanoi</i>	<i>Nicotiana tabacum</i>	Matas <i>et al.</i> (2014)
<i>hopB</i>	-	-	-	-
<i>hopBF</i>	-	-	-	-
<i>hopBI</i>	Inibidor de ETI	<i>P. savastanoi</i> pv. <i>savastanoi</i>	<i>Nicotiana tabacum</i>	Matas <i>et al.</i> (2014)
<i>hopBN</i>	-	-	-	-

<i>hopH</i>	-	-	-	-
<i>hopI</i>	Supressão do acúmulo de ácido salicílico	<i>P. syringae</i>	<i>A. thaliana</i>	Jelenska <i>et al.</i> (2007)
<i>hopM</i>	-	-	-	-
<i>hopQ</i>	Indução de morte celular e especificidade de hospedeiro	<i>P. syringae</i>	<i>Nicotiana benthamiana</i>	Wei <i>et al.</i> (2015)
	Aumento dos sintomas da doença	<i>P. syringae pv. tabaci</i>	<i>Phaseolus vulgaris</i>	Ferrante <i>et al.</i> (2009)
<i>hopV</i>	-	-	-	-
<i>hopY</i>	-	-	-	-
<i>hopZ</i>	Aumento da multiplicação bacteriana	<i>P. syringae</i>	<i>Glycine max</i>	Zhou <i>et al.</i> (2011)

(-) Efeitores que não foram encontrados na literatura relatando as suas funções.

MATERIAL SUPLEMENTAR

Tabela Suplementar S1. Proteínas presentes na região conservada do *cluster* de genes *hrp/hrc* de *Pseudomonas cichorii* e ausentes nas sequências de *Pseudomonas syringae* pv. *tomate*, *P. syringae* pv. *garcae*, e *P. syringae* pv. *tabaci*. As sequências foram obtidas usando o programa CLC e foram indicadas através do Blastx com o banco de dados de proteínas do NCBI. O maior valor de identidade e cobertura para todas as estirpes analisadas é apresentado.

Anotação	Cobertura (%)	E-value	Identidade (%)	ID da Proteína
NAD(+) difosfatase	87	0	100	WP_117164763.1
Crotanase/enoyl-CoA	68	0	99,63	WP_117164764.1
Proteína circadiana KaiC	98	0	99,8	WP_117164765.1
Regulador de resposta	99	0	99,85	WP_117164766.1
Regulador de resposta	95	3,00E-78	100	WP_025259857.1
Regulador de resposta	93	0	100	WP_117164767.1
Proteína CheB	83	3,00E-145	100	WP_117164768.1
Glutamato-O-metil-transferase	95	0	100	WP_117164769.1
Regulador de resposta	96	0	100	WP_117164770.1
Regulador de resposta	76	2,00E-67	100	WP_117164771.1
Glicosiltransferase	95	0	98,1	WP_122370847.1
Pecorrina-6A sintase	72	1,00E-180	100	WP_117164772.1
Regulador transcricional AraC	98	0	100	WP_117164773.1
Ativador transcricional MltR	53	1,00E-15	71,67	RMM76127.1
Transportador do tipo ABC	99	0	100	WP_117164774.1
Transportador permease do tipo ABC	96	0	100	WP_117164775.1
Transportador permease do tipo ABC	97	2,00E-122	100	WP_117164776.1
Transportador permease do tipo ABC	93	3,00E-61	98,99	RMS22821.1
Manitol desidrogenase	94	0	100	WP_117164778.1

Tabela Suplementar S2. Efeitores de *Pseudomonas syringae* pv. *garcae* identificados por Blastp contra o banco de dados de efetores depositados no *Pseudomonas syringae* Genome Resource.

Estirpe	Número de acesso	Efetor	Cobertura		Identidade	
			(%)	E-value	(%)	ID da Proteína
462	ASMN341635v1	<i>avrB1</i>	100	0.0	85	WP_080267796.1
462	ASMN341635v1	<i>avrB2</i>	87	4 e-82	47	WP_080267796.1
462	ASMN341635v1	<i>avrB3</i>	93	5 e-21	35	WP_080267796.1
462	ASMN341635v1	<i>avrB4-1</i>	99	8 e-42	31	WP_080267796.1
462	ASMN341635v1	<i>avrB4-2</i>	99	3 e-41	31	WP_080267796.1
462	ASMN341635v1	<i>avrD1</i>	100	0.0	98	WP_024671645.1
462	ASMN341635v1	<i>avrE1</i>	100	0.0	75	WP_117144245.1
462	ASMN341635v1	<i>avrE1</i>	90	0.0	68	WP_117195718.1
462	ASMN341635v1	<i>avrPto1</i>	98	9 e-55	53	WP_005892600.1
462	ASMN341635v1	<i>avrPto1</i>	98	2 e-34	40	WP_081023840.1
462	ASMN341635v1	<i>avrPto2</i>	93	4 e-35	44	WP_005892600.1
462	ASMN341635v1	<i>avrPto2</i>	93	1 e-19	36	WP_081023840.1
462	ASMN341635v1	<i>avrPto3</i>	98	5 e-24	35	WP_005892600.1
462	ASMN341635v1	<i>avrPto3</i>	99	1 e-21	34	WP_081023840.1
462	ASMN341635v1	<i>avrPto4</i>	100	2 e-98	86	WP_081023840.1
462	ASMN341635v1	<i>avrPto4</i>	100	6 e-29	40	WP_005892600.1
462	ASMN341635v1	<i>avrRpm1</i>	96	6 e-79	58	WP_117144539.1
462	ASMN341635v1	<i>avrRpm2</i>	88	1 e-29	39	WP_117144539.1
462	ASMN341635v1	<i>avrRps4</i>	100	3 e-117	74	WP_024668838.1
462	ASMN341635v1	<i>hopD1</i>	100	0.0	93	WP_117144314.1
462	ASMN341635v1	<i>hopH1</i>	95	9 e-30	39	WP_024689564.1
462	ASMN341635v1	<i>hopH2</i>	100	1 e-157	100	WP_024689564.1
462	ASMN341635v1	<i>hopR1</i>	100	0.0	82	WP_081023907.1
462	ASMN341635v1	<i>hopV1</i>	100	0.0	98	WP_081023882.1
462	ASMN341635v1	<i>hopW1</i>	39	6 e-72	50	WP_055004147.1
462	ASMN341635v1	<i>hopX1</i>	100	0.0	83	WP_081000186.1
462	ASMN341635v1	<i>hopY1</i>	99	1 e-121	64	WP_055004655.1
462	ASMN341635v1	<i>hopAA1-1</i>	99	0.0	79	WP_081023867.1
462	ASMN341635v1	<i>hopAA1-1</i>	100	0.0	70	WP_024671270.1
462	ASMN341635v1	<i>hopAA1-2</i>	100	0.0	99	WP_081023867.1
462	ASMN341635v1	<i>hopAA1-2</i>	99	0.0	68	WP_024671270.1
462	ASMN341635v1	<i>hopAE1</i>	99	0.0	73	WP_055004147.1
462	ASMN341635v1	<i>hopAF1</i>	61	1 e-120	92	WP_117144039.1

462	ASMN341635v1	<i>hopAG1</i>	100	0.0	64	WP_005895613.
462	ASMN341635v1	<i>hopAH1</i>	100	0.0	78	WP_081023866.1
462	ASMN341635v1	<i>hopAH1</i>	62	1 e-130	67	WP_055002768.1
462	ASMN341635v1	<i>hopAI1</i>	100	6 e-151	75	WP_024690843.1
462	ASMN341635v1	<i>hopAQ1</i>	60	1 e-13	63	WP_024668838.1
462	ASMN341635v1	<i>hopAS1</i>	92	0.0	88	WP_055004241.1
462	ASMN341635v1	<i>hopAU1</i>	100	0.0	99	WP_055004219.1
462	ASMN341635v1	<i>hopBD1</i>	99	0.0	97	WP_055004698.1
462	ASMN341635v1	<i>hopBD2</i>	84	2 e-82	50	WP_055004698.1
462	ASMN341635v1	<i>hopBF1</i>	100	2 e-152	100	WP_005886680.1
462	ASMN341635v1	<i>hopBH1</i>	100	0.0	99	WP_024691349.1
462	ASMN341635v1	<i>hopBI1</i>	75	0.0	99	WP_024670895.1
471	ASMN341623v1	<i>avrB1</i>	100	0.0	85	WP_080267796.1
471	ASMN341623v1	<i>avrB2</i>	87	4 e-82	47	WP_080267796.1
471	ASMN341623v1	<i>avrB3</i>	93	5 e-21	35	WP_080267796.1
471	ASMN341623v1	<i>avrB4-1</i>	99	8 e-42	31	WP_080267796.1
471	ASMN341623v1	<i>avrB4-2</i>	99	3 e-41	31	WP_080267796.1
471	ASMN341623v1	<i>avrD1</i>	100	0.0	98	WP_024671645.1
471	ASMN341623v1	<i>avrE1</i>	100	0.0	75	WP_117148032.1
471	ASMN341623v1	<i>avrE1</i>	88	0.0	68	WP_117148037.1
471	ASMN341623v1	<i>avrPto1</i>	98	9 e-55	53	WP_005892600.1
471	ASMN341623v1	<i>avrPto1</i>	98	2 e-34	40	WP_081023840.1
471	ASMN341623v1	<i>avrPto2</i>	93	4 e-35	44	WP_005892600.1
471	ASMN341623v1	<i>avrPto2</i>	93	1 e-19	36	WP_081023840.1
471	ASMN341623v1	<i>avrPto3</i>	98	5 e-24	35	WP_005892600.1
471	ASMN341623v1	<i>avrPto3</i>	99	1 e-21	34	WP_081023840.1
471	ASMN341623v1	<i>avrPto4</i>	100	2 e-98	86	WP_081023840.1
471	ASMN341623v1	<i>avrPto4</i>	100	6 e-29	40	WP_005892600.1
471	ASMN341623v1	<i>avrRps4</i>	100	3 e-117	74	WP_024668838.1
471	ASMN341623v1	<i>hopD1</i>	100	0.0	93	WP_117144314.1
471	ASMN341623v1	<i>hopH1</i>	95	9 e-30	39	WP_024689564.1
471	ASMN341623v1	<i>hopH2</i>	100	1 e-157	100	WP_024689564.1
471	ASMN341623v1	<i>hopk1</i>	43	3 e-70	76	WP_024668838.1
471	ASMN341623v1	<i>hopR1</i>	100	0.0	82	WP_081023907.1
471	ASMN341623v1	<i>hopV1</i>	100	0.0	98	WP_117147998.1
471	ASMN341623v1	<i>hopW1</i>	39	6 e-72	50	WP_055004147.1
471	ASMN341623v1	<i>hopX1</i>	100	0.0	83	WP_081000186.1

471	ASMN341623v1	<i>hopY1</i>	99	1 e-121	64	WP_055004655.1
471	ASMN341623v1	<i>hopZ1</i>	93	3 e-124	70	WP_081023836.1
471	ASMN341623v1	<i>hopZ2</i>	88	1 e-23	27	WP_081023836.1
471	ASMN341623v1	<i>hopAA1-1</i>	99	0.0	79	WP_081023867.1
471	ASMN341623v1	<i>hopAA1-1</i>	100	0.0	70	WP_024671270.1
471	ASMN341623v1	<i>hopAA1-2</i>	100	0.0	99	WP_081023867.1
471	ASMN341623v1	<i>hopAA1-2</i>	99	0.0	68	WP_024671270.1
471	ASMN341623v1	<i>hopAE1</i>	99	0.0	73	WP_055004147.1
471	ASMN341623v1	<i>hopAF1</i>	61	1 e-120	92	WP_117144039.1
471	ASMN341623v1	<i>hopAG1</i>	100	0.0	64	WP_005895613.
471	ASMN341623v1	<i>hopAH1</i>	100	0.0	78	WP_081023866.1
471	ASMN341623v1	<i>hopAH1</i>	65	1 e-130	67	WP_055002768.1
471	ASMN341623v1	<i>hopAI1</i>	100	6 e-151	75	WP_024690843.1
471	ASMN341623v1	<i>hopAQ1</i>	60	1 e-13	63	WP_024668838.1
471	ASMN341623v1	<i>hopAR1</i>	61	2 e-12	28	WP_005735297.1
471	ASMN341623v1	<i>hopAS1</i>	92	0.0	88	WP_055004241.1
471	ASMN341623v1	<i>hopAU1</i>	100	0.0	99	WP_055004219.1
471	ASMN341623v1	<i>hopAY1</i>	84	0.0	99	WP_005735297.1
471	ASMN341623v1	<i>hopBF1</i>	100	2 e-152	100	WP_005886680.1
471	ASMN341623v1	<i>hopBH1</i>	100	0.0	99	WP_024691349.1
471	ASMN341623v1	<i>hopBI1</i>	75	0.0	99	WP_024670895.1
445	ASMN341591v1	<i>avrB1</i>	100	0.0	85	WP_080267796.1
445	ASMN341591v1	<i>avrB2</i>	87	4 e-82	47	WP_080267796.1
445	ASMN341591v1	<i>avrB3</i>	93	5 e-21	35	WP_080267796.1
445	ASMN341591v1	<i>avrB4-1</i>	99	8 e-42	31	WP_080267796.1
445	ASMN341591v1	<i>avrB4-2</i>	99	3 e-41	31	WP_080267796.1
445	ASMN341591v1	<i>avrD1</i>	100	0.0	98	WP_024671645.1
445	ASMN341591v1	<i>avrE1</i>	100	0.0	75	WP_117144245.1
445	ASMN341591v1	<i>avrE1</i>	90	0.0	75	WP_117150633.1
445	ASMN341591v1	<i>avrPto1</i>	98	9 e-55	53	WP_005892600.1
445	ASMN341591v1	<i>avrPto1</i>	98	2 e-34	40	WP_081023840.1
445	ASMN341591v1	<i>avrPto2</i>	93	4 e-35	44	WP_005892600.1
445	ASMN341591v1	<i>avrPto2</i>	93	1 e-19	36	WP_081023840.1
445	ASMN341591v1	<i>avrPto3</i>	98	5 e-24	35	WP_005892600.1
445	ASMN341591v1	<i>avrPto3</i>	99	1 e-21	34	WP_081023840.1
445	ASMN341591v1	<i>avrPto4</i>	100	2 e-98	86	WP_081023840.1
445	ASMN341591v1	<i>avrPto4</i>	100	6 e-29	40	WP_005892600.1

445	ASMN341591v1	<i>avrRps4</i>	100	3 e-117	74	WP_024668838.1
445	ASMN341591v1	<i>hopD1</i>	100	0.0	93	WP_117144314.1
445	ASMN341591v1	<i>hopH1</i>	95	9 e-30	39	WP_024689564.1
445	ASMN341591v1	<i>hopH2</i>	100	1 e-157	100	WP_024689564.1
445	ASMN341591v1	<i>hopk1</i>	43	3 e-70	76	WP_024668838.1
445	ASMN341591v1	<i>hopR1</i>	100	0.0	82	WP_081023907.1
445	ASMN341591v1	<i>hopV1</i>	100	0.0	98	WP_081023882.1
445	ASMN341591v1	<i>hopW1</i>	39	6 e-72	50	WP_055004147.1
445	ASMN341591v1	<i>hopX1</i>	100	0.0	83	WP_081000186.1
445	ASMN341591v1	<i>hopY1</i>	99	1 e-121	64	WP_055004655.1
445	ASMN341591v1	<i>hopZ1</i>	93	3 e-124	70	WP_081023836.1
445	ASMN341591v1	<i>hopZ2</i>	88	1 e-23	27	WP_081023836.1
445	ASMN341591v1	<i>hopAA1-1</i>	99	0.0	79	WP_081023867.1
445	ASMN341591v1	<i>hopAA1-1</i>	100	0.0	70	WP_024671270.1
445	ASMN341591v1	<i>hopAA1-2</i>	100	0.0	99	WP_081023867.1
445	ASMN341591v1	<i>hopAA1-2</i>	99	0.0	68	WP_024671270.1
445	ASMN341591v1	<i>hopAE1</i>	99	0.0	73	WP_055004147.1
445	ASMN341591v1	<i>hopAF1</i>	61	1 e-120	92	WP_117144039.1
445	ASMN341591v1	<i>hopAG1</i>	100	0.0	64	WP_005895613.1
445	ASMN341591v1	<i>hopAH1</i>	100	0.0	78	WP_081023866.1
445	ASMN341591v1	<i>hopAH1</i>	65	1 e-130	67	WP_055002768.1
445	ASMN341591v1	<i>hopAI1</i>	100	6 e-151	75	WP_024690843.1
445	ASMN341591v1	<i>hopAQ1</i>	60	1 e-13	63	WP_024668838.1
445	ASMN341591v1	<i>hopAR1</i>	61	2 e-12	28	WP_005735297.1
445	ASMN341591v1	<i>hopAS1</i>	92	0.0	88	WP_055004241.1
445	ASMN341591v1	<i>hopAU1</i>	100	0.0	99	WP_055004219.1
445	ASMN341591v1	<i>hopAY1</i>	84	0.0	99	WP_005735297.1
445	ASMN341591v1	<i>hopBD1</i>	99	0.0	97	WP_055004698.1
445	ASMN341591v1	<i>hopBD2</i>	84	2 e-82	50	WP_055004698.1
445	ASMN341591v1	<i>hopBF1</i>	100	2 e-152	100	WP_005886680.1
445	ASMN341591v1	<i>hopBH1</i>	100	0.0	99	WP_024691349.1
445	ASMN341591v1	<i>hopBI1</i>	75	0.0	99	WP_024670895.1
460	ASMN341583v1	<i>avrB1</i>	100	0.0	85	WP_080267796.1
460	ASMN341583v1	<i>avrB2</i>	87	4 e-82	47	WP_080267796.1
460	ASMN341583v1	<i>avrB3</i>	93	5 e-21	35	WP_080267796.1
460	ASMN341583v1	<i>avrB4-1</i>	99	8 e-42	31	WP_080267796.1
460	ASMN341583v1	<i>avrB4-2</i>	99	3 e-41	31	WP_080267796.1

460	ASMN341583v1	<i>avrD1</i>	100	0.0	98	WP_024671645.1
460	ASMN341583v1	<i>avrE1</i>	100	0.0	75	WP_117207739.1
460	ASMN341583v1	<i>avrE1</i>	90	0.0	75	WP_117207745.1
460	ASMN341583v1	<i>avrPto1</i>	98	9 e-55	53	WP_005892600.1
460	ASMN341583v1	<i>avrPto1</i>	98	2 e-34	40	WP_081023840.1
460	ASMN341583v1	<i>avrPto2</i>	93	4 e-35	44	WP_005892600.1
460	ASMN341583v1	<i>avrPto2</i>	93	1 e-19	36	WP_081023840.1
460	ASMN341583v1	<i>avrPto3</i>	98	4 e-24	35	WP_005892600.1
460	ASMN341583v1	<i>avrPto3</i>	99	1 e-21	34	WP_081023840.1
460	ASMN341583v1	<i>avrPto4</i>	100	2 e-98	86	WP_081023840.1
460	ASMN341583v1	<i>avrPto4</i>	100	6 e-29	40	WP_005892600.1
460	ASMN341583v1	<i>avrRpm1</i>	96	6 e-79	58	WP_117144539.1
460	ASMN341583v1	<i>avrRpm2</i>	88	1 e-29	39	WP_117144539.1
460	ASMN341583v1	<i>avrRps4</i>	100	3 e-117	74	WP_024668838.1
460	ASMN341583v1	<i>hopD1</i>	100	0.0	93	WP_117144314.1
460	ASMN341583v1	<i>hopH1</i>	95	9 e-30	39	WP_024689564.1
460	ASMN341583v1	<i>hopH2</i>	100	1 e-157	100	WP_024689564.1
460	ASMN341583v1	<i>hopk1</i>	43	3 e-70	76	WP_024668838.1
460	ASMN341583v1	<i>hopR1</i>	100	0.0	82	WP_081023907.1
460	ASMN341583v1	<i>hopV1</i>	100	0.0	98	WP_081023882.1
460	ASMN341583v1	<i>hopW1</i>	39	6 e-72	50	WP_055004147.1
460	ASMN341583v1	<i>hopX1</i>	100	0.0	83	WP_081000186.1
460	ASMN341583v1	<i>hopY1</i>	99	1 e-121	64	WP_055004655.1
460	ASMN341583v1	<i>hopAA1-1</i>	99	0.0	79	WP_081023867.1
460	ASMN341583v1	<i>hopAA1-1</i>	100	0.0	70	WP_024671270.1
460	ASMN341583v1	<i>hopAA1-2</i>	100	0.0	99	WP_081023867.1
460	ASMN341583v1	<i>hopAA1-2</i>	99	0.0	68	WP_024671270.1
460	ASMN341583v1	<i>hopAE1</i>	99	0.0	73	WP_055004147.1
460	ASMN341583v1	<i>hopAF1</i>	61	1 e-120	92	WP_117144039.1
460	ASMN341583v1	<i>hopAG1</i>	100	0.0	64	WP_005895613.1
460	ASMN341583v1	<i>hopAH1</i>	100	0.0	78	WP_081023866.1
460	ASMN341583v1	<i>hopAH1</i>	65	1 e-130	67	WP_055002768.1
460	ASMN341583v1	<i>hopAl1</i>	100	6 e-151	75	WP_024690843.1
460	ASMN341583v1	<i>hopAQ1</i>	60	1 e-13	63	WP_024668838.1
460	ASMN341583v1	<i>hopAS1</i>	92	0.0	88	WP_055004241.1
460	ASMN341583v1	<i>hopAU1</i>	100	0.0	99	WP_055004219.1
460	ASMN341583v1	<i>hopBD1</i>	99	0.0	97	WP_055004698.1

460	ASMN341583v1	<i>hopBD2</i>	84	2 e-82	50	WP_055004698.1
460	ASMN341583v1	<i>hopBF1</i>	100	2 e-152	100	WP_005886680.1
460	ASMN341583v1	<i>hopBH1</i>	100	0.0	99	WP_024691349.1
460	ASMN341583v1	<i>hopBI1</i>	75	0.0	99	WP_024670895.1
438	ASMN341595v1	<i>avrB1</i>	100	0.0	85	WP_080267796.1
438	ASMN341595v1	<i>avrB2</i>	87	4 e-82	47	WP_080267796.1
438	ASMN341595v1	<i>avrB3</i>	93	5 e-21	35	WP_080267796.1
438	ASMN341595v1	<i>avrB4-1</i>	99	8 e-42	31	WP_080267796.1
438	ASMN341595v1	<i>avrB4-2</i>	99	3 e-41	31	WP_080267796.1
438	ASMN341595v1	<i>avrE1</i>	100	0.0	75	WP_117144573.1
438	ASMN341595v1	<i>avrE1</i>	88	0.0	68	WP_117149654.1
438	ASMN341595v1	<i>avrPto1</i>	98	9 e-55	53	WP_005892600.1
438	ASMN341595v1	<i>avrPto1</i>	98	2 e-34	40	WP_081023840.1
438	ASMN341595v1	<i>avrPto2</i>	93	4 e-35	44	WP_005892600.1
438	ASMN341595v1	<i>avrPto2</i>	93	1 e-19	36	WP_081023840.1
438	ASMN341595v1	<i>avrPto3</i>	98	4 e-24	35	WP_005892600.1
438	ASMN341595v1	<i>avrPto3</i>	99	1 e-21	34	WP_081023840.1
438	ASMN341595v1	<i>avrPto4</i>	100	2 e-98	86	WP_081023840.1
438	ASMN341595v1	<i>avrPto4</i>	100	6 e-29	40	WP_005892600.1
438	ASMN341595v1	<i>avrRpm1</i>	96	6 e-79	58	WP_117144539.1
438	ASMN341595v1	<i>avrRpm2</i>	88	1 e-29	39	WP_117144539.1
438	ASMN341595v1	<i>hopH1</i>	95	9 e-30	39	WP_024689564.1
438	ASMN341595v1	<i>hopH2</i>	100	1 e-157	100	WP_024689564.1
438	ASMN341595v1	<i>hopR1</i>	100	0.0	82	WP_081023907.1
438	ASMN341595v1	<i>hopV1</i>	100	0.0	98	WP_081023882.1
438	ASMN341595v1	<i>hopW1</i>	39	6 e-72	50	WP_055004147.1
438	ASMN341595v1	<i>hopX1</i>	100	0.0	83	WP_081000186.1
438	ASMN341595v1	<i>hopY1</i>	99	1 e-121	64	WP_055004655.1
438	ASMN341595v1	<i>hopAA1-1</i>	99	0.0	79	WP_081023867.1
438	ASMN341595v1	<i>hopAA1-1</i>	100	0.0	70	WP_024671270.1
438	ASMN341595v1	<i>hopAA1-2</i>	100	0.0	99	WP_081023867.1
438	ASMN341595v1	<i>hopAA1-2</i>	99	0.0	68	WP_024671270.1
438	ASMN341595v1	<i>hopAE1</i>	99	0.0	73	WP_055004147.1
438	ASMN341595v1	<i>hopAF1</i>	61	1 e-120	92	WP_117144039.1
438	ASMN341595v1	<i>hopAG1</i>	100	0.0	64	WP_005895613.1
438	ASMN341595v1	<i>hopAH1</i>	100	0.0	78	WP_081023866.1
438	ASMN341595v1	<i>hopAH1</i>	65	1 e-130	67	WP_055002768.1

438	ASMN341595v1	<i>hopA11</i>	100	6 e-151	75	WP_024690843.1
438	ASMN341595v1	<i>hopAS1</i>	92	0.0	88	WP_055004241.1
438	ASMN341595v1	<i>hopBF1</i>	100	2 e-152	100	WP_005886680.1
438	ASMN341595v1	<i>hopBH1</i>	100	0.0	99	WP_024691349.1
438	ASMN341595v1	<i>hopBI1</i>	75	0.0	99	WP_024670895.1

Tabela Suplementar S3. Efeitores identificados nas cinco estirpes de *Pseudomonas syringae* pv. *garcae* usando o *script* customizado em Python que localiza sequências assinaturas presentes nas regiões promotoras (caixa *hrp*).

Estirpe	Número de acesso	Qualificador	Efeitor
438	WP_081023907.1	Efeitor de tipo III de <i>Psgc</i>	<i>hopR1</i>
438	WP_080267796.1	Proteína hipotética	<i>avrB</i>
438	WP_024670046.1	Proteína de ligação a quitina	<i>Psgc_EF1</i>
438	WP_005886680.1	Proteína hipotética	<i>hopBF1</i>
438	WP_024671432.1	Proteína hipotética	<i>Psgc_EF2</i>
438	WP_024689564.1	Proteína hipotética	<i>hopH1</i>
438	WP_117144573.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE1</i>
438	WP_117149978.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE1</i>
438	WP_117149654.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE1</i>
438	WP_117149978.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE1</i>
438	WP_117149654.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE1</i>
438	WP_055004147.1	Endonuclease de restrição	<i>hopAE1</i>
438	WP_005894932.1	Proteína hipotética	<i>Psgc_EF3</i>
438	WP_081023840.1	Efeitor de tipo III de <i>Psgc</i>	<i>avrPto1</i>
438	WP_024670898.1	Proteína com domínio de metiltransferase	<i>Psgc_EF4</i>
438	WP_055002524.1	N-acetiltransferase	<i>Psgc_EF5</i>
438	WP_117143995.1	Oxidoreductase de ligação a FAD	<i>Psgc_EF6</i>
438	WP_081023826.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
438	WP_005887538.1	NAD(P)(+) transidrogenase Si-específica	<i>Psgc_EF8</i>
438	WP_117144539.1	Efeitor de tipo III de <i>Psgc</i>	<i>avrRpm1</i>
438	WP_024671270.1	Proteína hipotética	<i>hopAA1</i>
445	WP_005894932.1	Proteína hipotética	<i>Psgc_EF3</i>
445	WP_081023836.1	Efeitor do sistema de secreção de tipo III da família YopJ de <i>Psgc</i>	<i>hopZ1</i>
445	WP_024670898.1	Proteína com domínio de metiltransferase	<i>Psgc_EF4</i>
445	WP_055002524.1	N-acetiltransferase	<i>Psgc_EF5</i>
445	WP_054998401.1	Oxidoreductase de ligação a FAD	<i>Psgc_EF6</i>
445	WP_080267796.1	Proteína hipotética	<i>avrB</i>
445	WP_024670046.1	Proteína de ligação a quitina	<i>Psgc_EF1</i>
445	WP_005886680.1	Proteína hipotética	<i>hopBF1</i>
445	WP_024691349.1	Proteína hipotética	<i>Psgc_EF9</i>
445	WP_117144314.1	Proteína efetora AvrPphD de <i>Psgc</i>	<i>avrPphD</i>
445	WP_117144038.1	Proteína hipotética	<i>hopAF1</i>
445	WP_117144039.1	Efeitor de tipo III de <i>Psgc</i>	<i>hopAF1</i>

445	WP_055003882.1	Proteína <i>helper</i> de tipo III HopAK1	<i>hopAK1</i>
445	WP_005888957.1	Proteína SmpB de ligação a SsrA	<i>Psgc_EF10</i>
445	WP_024668778.1	Proteína hipotética	<i>Psgc_EF11</i>
445	WP_117144245.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE</i>
445	WP_081023907.1	Efector de tipo III de <i>Psgc</i>	<i>hopR1</i>
445	WP_024671432.1	Proteína hipotética	<i>Psgc_EF2</i>
445	WP_024689564.1	Proteína hipotética	<i>hopH1</i>
445	WP_055004601.1	Transglicosilase lítica de mureína	<i>hopAJ1</i>
445	WP_055004602.1	Proteína com domínio DUF1440	<i>Psgc_EF12</i>
445	WP_081023840.1	Efector de tipo III de <i>Psgc</i>	<i>avrPto1</i>
445	WP_024668838.1	Proteína hipotética	<i>avrRps4</i>
445	WP_024671645.1	Proteína biosintética de siringolídeo AvrD1	<i>avrD1</i>
445	WP_080268302.1	Proteína da família histidina fosfatase	<i>Psgc_EF13</i>
445	WP_055004147.1	Endonuclease de restrição	<i>hopAE1</i>
445	WP_081023826.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
445	WP_005887538.1	NAD(P)(+) transidrogenase Si-específica	<i>Psgc_EF8</i>
445	WP_024671270.1	Proteína hipotética	<i>hopAA1</i>
460	WP_024671270.1	Proteína hipotética	<i>hopAA1</i>
460	WP_005894932.1	Proteína hipotética	<i>Psgc_EF3</i>
460	WP_024670898.1	Proteína com domínio de metiltransferase	<i>Psgc_EF4</i>
460	WP_055002524.1	N-acetiltransferase	<i>Psgc_EF5</i>
460	WP_117143995.1	Oxidoreductase de ligação a FAD	<i>Psgc_EF6</i>
460	WP_117144539.1	Efector de tipo III de <i>Psgc</i>	<i>avrRpm1</i>
460	WP_055004147.1	Endonuclease de restrição	<i>hopAE1</i>
460	WP_024691349.1	Proteína hipotética	<i>Psgc_EF9</i>
460	WP_055004601.1	Transglicosilase lítica de mureína	<i>hopAJ1</i>
460	WP_055004602.1	Proteína com domínio DUF1440	<i>Psgc_EF13</i>
460	WP_081023907.1	Efector de tipo III de <i>Psgc</i>	<i>hopR1</i>
460	WP_080267796.1	Proteína hipotética	<i>avrB</i>
460	WP_024670046.1	Proteína de ligação a quitina	<i>Psgc_EF1</i>
460	WP_005886680.1	Proteína hipotética	<i>hopBF1</i>
460	WP_117146000.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE</i>
460	WP_117207739.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE</i>
460	WP_117144314.1	Proteína efetora AvrPphD de <i>Psgc</i>	<i>avrPphD</i>
460	WP_081023826.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
460	WP_005887538.1	NAD(P)(+) transidrogenase Si-específica	<i>Psgc_EF8</i>
460	WP_055003882.1	Proteína <i>helper</i> de tipo III HopAK1	<i>hopAK1</i>

460	WP_117144038.1	Proteína hipotética	<i>hopAF1</i>
460	WP_117144039.1	Efector de tipo III de <i>Psgc</i>	<i>hopAF1</i>
460	WP_024671432.1	Proteína hipotética	<i>Psgc_EF2</i>
460	WP_024689564.1	Proteína hipotética	<i>hopH1</i>
460	WP_024668838.1	Proteína hipotética	<i>avrRps4</i>
460	WP_024671645.1	Proteína biosintética de siringolídeo AvrD1	<i>avrD1</i>
460	WP_080268302.1	Proteína da família histidina fosfatase	<i>Psgc_EF13</i>
462	WP_005894932.1	Proteína hipotética	<i>Psgc_EF3</i>
462	WP_055003882.1	Proteína <i>helper</i> de tipo III HopAK1	<i>hopAK1</i>
462	WP_117144038.1	Proteína hipotética	<i>hopAF1</i>
462	WP_117144039.1	Efector de tipo III de <i>Psgc</i>	<i>hopAF1</i>
462	WP_005888957.1	Proteína SmpB de ligação a SsrA	<i>Psgc_EF10</i>
462	WP_024668778.1	Proteína hipotética	<i>Psgc_EF11</i>
462	WP_024670898.1	Proteína com domínio de metiltransferase	<i>Psgc_EF4</i>
462	WP_055002524.1	N-acetiltransferase	<i>Psgc_EF5</i>
462	WP_117143995.1	Oxidoreductase de ligação a FAD	<i>Psgc_EF6</i>
462	WP_117144314.1	Proteína efetora AvrPphD de <i>Psgc</i>	<i>avrPphD</i>
462	WP_024691349.1	Proteína hipotética	<i>Psgc_EF9</i>
462	WP_081023840.1	Efector de tipo III de <i>Psgc</i>	<i>avrPto1</i>
462	WP_081023826.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
462	WP_005887538.1	NAD(P)(+) transidrogenase Si-específica	<i>Psgc_EF8</i>
462	WP_081023907.1	Efector de tipo III de <i>Psgc</i>	<i>hopR1</i>
462	WP_080267796.1	Proteína hipotética	<i>avrB</i>
462	WP_024670046.1	Proteína de ligação a quitina	<i>Psgc_EF1</i>
462	WP_005886680.1	Proteína hipotética	<i>hopBF1</i>
462	WP_055004147.1	Endonuclease de restrição	<i>hopAE1</i>
462	WP_024668838.1	Proteína hipotética	<i>avrRps4</i>
462	WP_024671645.1	Proteína biosintética de siringolídeo AvrD1	<i>avrD1</i>
462	WP_080268302.1	Proteína da família histidine fosfatase	<i>Psgc_EF13</i>
462	WP_024671432.1	Proteína hipotética	<i>Psgc_EF2</i>
462	WP_024689564.1	Proteína hipotética	<i>hopH1</i>
462	WP_024671270.1	Proteína hipotética	<i>hopAA1</i>
462	WP_117144245.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE</i>
462	WP_117195730.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE</i>
462	WP_117195730.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE</i>
462	WP_117144539.1	Efector de tipo III de <i>Psgc</i>	<i>avrRpm1</i>
471	WP_024671645.1	Proteína biosintética de siringolídeo AvrD1	<i>avrD1</i>

471	WP_080268302.1	Proteína da família histidine fosfatase	<i>Psgc_EF13</i>
471	WP_024668838.1	Proteína hipotética	<i>avrRps4</i>
471	WP_081023826.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
471	WP_005887538.1	NAD(P)(+) transidrogenase Si-específica	<i>Psgc_EF8</i>
471	WP_055004147.1	Endonuclease de restrição	<i>hopAE1</i>
471	WP_081023907.1	Efetor de tipo III de <i>Psgc</i>	<i>hopR1</i>
471	WP_024671270.1	Proteína hipotética	<i>hopAA1</i>
471	WP_024691349.1	Proteína hipotética	<i>Psgc_EF9</i>
471	WP_081023840.1	Efetor de tipo III de <i>Psgc</i>	<i>avrPto1</i>
471	WP_117144314.1	Proteína efetora AvrPphD de <i>Psgc</i>	<i>avrPphD</i>
471	WP_081023836.1	Efetor do sistema de secreção de tipo III da família YopJ de <i>Psgc</i>	<i>hopZ1</i>
471	WP_080267796.1	Proteína hipotética	<i>avrB</i>
471	WP_024670046.1	Proteína de ligação a quitina	<i>Psgc_EF1</i>
471	WP_005886680.1	Proteína hipotética	<i>hopBF1</i>
471	WP_024670898.1	Proteína com domínio de metiltransferase	<i>Psgc_EF4</i>
471	WP_055002524.1	N-acetiltransferase	<i>Psgc_EF5</i>
471	WP_054998401.1	Oxidoreductase de ligação a FAD	<i>Psgc_EF6</i>
471	WP_055003882.1	Proteína <i>helper</i> de tipo III HopAK1	<i>hopAK1</i>
471	WP_117144038.1	Proteína hipotética	<i>hopAF1</i>
471	WP_117144039.1	Efetor de tipo III de <i>Psgc</i>	<i>hopAF1</i>
471	WP_005894932.1	Proteína hipotética	<i>Psgc_EF3</i>
471	WP_117148031.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE</i>
471	WP_117148032.1	Proteína da família AvrE de efetores de tipo 3 de <i>Psgc</i>	<i>avrE</i>
471	WP_024671432.1	Proteína hipotética	<i>Psgc_EF2</i>
471	WP_024689564.1	Proteína hipotética	<i>hopH1</i>

Tabela Suplementar S4. Efeitores de *Pseudomonas syringae* pv. *tabaci* identificados por Blastp contra o banco de dados de efetores depositados no *Pseudomonas syringae* Genome Resource.

Estirpe	Número de acesso	Efeitor	Cobertura (%)	E-value	Identidade (%)	ID da proteína
480	ASMN341614v1	<i>avrD1</i>	100	0.0	98	WP_024671645.1
480	ASMN341614v1	<i>avrE1</i>	100	0.0	84	WP_117140309.1
480	ASMN341614v1	<i>avrRpm1</i>	96	5 e-80	59	WP_117140700.1
480	ASMN341614v1	<i>avrRpm2</i>	64	2 e-29	43	WP_117140700.1
480	ASMN341614v1	<i>avrRps4</i>	100	4 e-117	74	WP_024668838.1
480	ASMN341614v1	<i>hopD1</i>	100	0.0	93	WP_117140212.1
480	ASMN341614v1	<i>hopI1</i>	79	2 e-123	82	WP_057404518.1
480	ASMN341614v1	<i>hopK1</i>	43	3 e-70	76	WP_024668838.1
480	ASMN341614v1	<i>hopM1</i>	100	0.0	82	WP_117140307.1
480	ASMN341614v1	<i>hopR1</i>	100	0.0	86	WP_080953444.1
480	ASMN341614v1	<i>hopV1</i>	74	4 e-172	77	WP_005777017.1
480	ASMN341614v1	<i>hopW1</i>	39	4 e-79	49	WP_117140040.1
480	ASMN341614v1	<i>hopX1</i>	98	3 e-170	68	WP_080726063.1
480	ASMN341614v1	<i>hopAA1-1</i>	100	0.0	83	WP_117140304.1
480	ASMN341614v1	<i>hopAA1-2</i>	99	0.0	82	WP_117140304.1
480	ASMN341614v1	<i>hopAB1</i>	100	6 e-132	47	WP_117139356.1
480	ASMN341614v1	<i>hopAB2</i>	100	0.0	68	WP_117139356.1
480	ASMN341614v1	<i>hopAB3</i>	100	0.0	63	WP_117139356.1
480	ASMN341614v1	<i>hopAE1</i>	98	0.0	72	WP_117140040.1
480	ASMN341614v1	<i>hopAG1</i>	100	0.0	81	WP_010196193.1
480	ASMN341614v1	<i>hopAH1</i>	100	0.0	83	WP_081023459.1
480	ASMN341614v1	<i>hopAH1</i>	64	1 e-135	69	WP_039847471.1
480	ASMN341614v1	<i>hopAQ1</i>	60	1 e-13	63	WP_024668838.1
480	ASMN341614v1	<i>hopAR1</i>	100	2 e-158	78	WP_117140095.1
480	ASMN341614v1	<i>hopAS1</i>	92	0.0	95	WP_117140549.1
480	ASMN341614v1	<i>hopAY1</i>	71	7 e-12	28	WP_117140095.1
480	ASMN341614v1	<i>hopAZ1</i>	100	3 e-123	73	WP_024663017.1
480	ASMN341614v1	<i>hopBD1</i>	99	0.0	97	WP_055004698.1
480	ASMN341614v1	<i>hopBD2</i>	84	2 e-82	50	WP_055004698.1
480	ASMN341614v1	<i>hopBH1</i>	35	3 e-18	39	WP_117139931.1
440	ASMN341658v1	<i>avrE1</i>	100	0.0	84	WP_117140309.1
440	ASMN341658v1	<i>avrRpm1</i>	96	5 e-80	59	WP_117140700.1
440	ASMN341658v1	<i>avrRpm2</i>	64	2 e-29	43	WP_117140700.1

440	ASMN341658v1	<i>hopI1</i>	79	2 e-123	82	WP_057404518.1
440	ASMN341658v1	<i>hopM1</i>	100	0.0	83	WP_117182490.1
440	ASMN341658v1	<i>hopQ1-1</i>	100	0.0	97	WP_117182356.1
440	ASMN341658v1	<i>hopQ1-2</i>	100	4 e-86	88	WP_117182356.1
440	ASMN341658v1	<i>hopR1</i>	100	0.0	86	WP_080953444.1
440	ASMN341658v1	<i>hopV1</i>	74	4 e-172	77	WP_005777017.1
440	ASMN341658v1	<i>hopW1</i>	39	4 e-79	49	WP_117140040.1
440	ASMN341658v1	<i>hopX1</i>	98	3 e-170	68	WP_080726063.1
440	ASMN341658v1	<i>hopAA1-1</i>	100	0.0	83	WP_117140304.1
440	ASMN341658v1	<i>hopAA1-2</i>	99	0.0	82	WP_117140304.1
440	ASMN341658v1	<i>hopAB1</i>	100	7 e-132	47	WP_117139356.1
440	ASMN341658v1	<i>hopAB2</i>	100	0.0	68	WP_117139356.1
440	ASMN341658v1	<i>hopAB3</i>	100	0.0	63	WP_117139356.1
440	ASMN341658v1	<i>hopAE1</i>	98	0.0	72	WP_117140040.1
440	ASMN341658v1	<i>hopAG1</i>	100	0.0	81	WP_010196193.1
440	ASMN341658v1	<i>hopAH1</i>	100	0.0	83	WP_081023459.1
440	ASMN341658v1	<i>hopAH1</i>	64	1 e-135	69	WP_039847471.1
440	ASMN341658v1	<i>hopAR1</i>	100	2 e-158	78	WP_117140095.1
440	ASMN341658v1	<i>hopAS1</i>	92	0.0	95	WP_117140549.1
440	ASMN341658v1	<i>hopAY1</i>	71	7 e-12	28	WP_117140095.1
440	ASMN341658v1	<i>hopAZ1</i>	100	3 e-123	73	WP_024663017.1
440	ASMN341658v1	<i>hopBH1</i>	35	3 e-18	39	WP_117139931.1
469	ASMN341578v1	<i>avrD1</i>	100	0.0	98	WP_024671645.1
469	ASMN341578v1	<i>avrE1</i>	100	0.0	84	WP_117140309.1
469	ASMN341578v1	<i>avrRpm1</i>	96	5 e-80	59	WP_117140700.1
469	ASMN341578v1	<i>avrRpm2</i>	64	2 e-29	43	WP_117140700.1
469	ASMN341578v1	<i>avrRps4</i>	100	4 e-117	74	WP_024668838.1
469	ASMN341578v1	<i>hopI1</i>	79	2 e-123	82	WP_057404518.1
469	ASMN341578v1	<i>hopk1</i>	43	3 e-70	76	WP_024668838.1
469	ASMN341578v1	<i>hopM1</i>	100	0.0	82	WP_117140307.1
469	ASMN341578v1	<i>hopR1</i>	100	0.0	86	WP_080953444.1
469	ASMN341578v1	<i>hopV1</i>	74	4 e-172	77	WP_005777017.1
469	ASMN341578v1	<i>hopW1</i>	39	4 e-79	49	WP_117140040.1
469	ASMN341578v1	<i>hopX1</i>	98	3 e-170	68	WP_080726063.1
469	ASMN341578v1	<i>hopAA1-1</i>	100	0.0	83	WP_117140304.1
469	ASMN341578v1	<i>hopAA1-2</i>	99	0.0	82	WP_117140304.1
469	ASMN341578v1	<i>hopAB1</i>	100	7 e-132	47	WP_117139356.1

469	ASMN341578v1	<i>hopAB2</i>	100	0.0	68	WP_117139356.1
469	ASMN341578v1	<i>hopAB3</i>	100	0.0	63	WP_117139356.1
469	ASMN341578v1	<i>hopAE1</i>	98	0.0	72	WP_117140040.1
469	ASMN341578v1	<i>hopAG1</i>	100	0.0	81	WP_010196193.1
469	ASMN341578v1	<i>hopAH1</i>	100	0.0	83	WP_081023459.1
469	ASMN341578v1	<i>hopAH1</i>	64	1 e-135	69	WP_039847471.1
469	ASMN341578v1	<i>hopAQ1</i>	60	2 e-13	63	WP_024668838.1
469	ASMN341578v1	<i>hopAR1</i>	100	2 e-158	78	WP_117140095.1
469	ASMN341578v1	<i>hopAS1</i>	92	0.0	95	WP_117140549.1
469	ASMN341578v1	<i>hopAY1</i>	71	7 e-12	28	WP_117140095.1
469	ASMN341578v1	<i>hopAZ1</i>	100	3 e-123	73	WP_024663017.1
469	ASMN341578v1	<i>hopBD1</i>	99	0.0	97	WP_055004698.1
469	ASMN341578v1	<i>hopBD2</i>	84	2 e-82	50	WP_055004698.1
469	ASMN341578v1	<i>hopBH1</i>	35	3 e-18	39	WP_117139931.1
478	ASMN341575v1	<i>avrD1</i>	100	0.0	98	WP_024671645.1
478	ASMN341575v1	<i>avrE1</i>	100	0.0	84	WP_117140309.1
478	ASMN341575v1	<i>avrRpm1</i>	96	5 e-80	59	WP_117140700.1
478	ASMN341575v1	<i>avrRpm2</i>	64	2 e-29	43	WP_117140700.1
478	ASMN341575v1	<i>avrRps4</i>	100	4 e-117	74	WP_024668838.1
478	ASMN341575v1	<i>hopD1</i>	100	0.0	93	WP_117140212.1
478	ASMN341575v1	<i>hopI1</i>	79	2 e-123	82	WP_057404518.1
478	ASMN341575v1	<i>hopk1</i>	43	3 e-70	76	WP_024668838.1
478	ASMN341575v1	<i>hopM1</i>	100	0.0	82	WP_117140307.1
478	ASMN341575v1	<i>hopR1</i>	100	0.0	86	WP_080953444.1
478	ASMN341575v1	<i>hopV1</i>	74	4 e-172	77	WP_005777017.1
478	ASMN341575v1	<i>hopW1</i>	39	4 e-79	49	WP_117140040.1
478	ASMN341575v1	<i>hopX1</i>	98	3 e-170	68	WP_080726063.1
478	ASMN341575v1	<i>hopAA1-1</i>	100	0.0	83	WP_117140304.1
478	ASMN341575v1	<i>hopAA1-2</i>	99	0.0	82	WP_117140304.1
478	ASMN341575v1	<i>hopAB1</i>	100	7 e-132	47	WP_117139356.1
478	ASMN341575v1	<i>hopAB2</i>	100	0.0	68	WP_117139356.1
478	ASMN341575v1	<i>hopAB3</i>	100	0.0	63	WP_117139356.1
478	ASMN341575v1	<i>hopAE1</i>	98	0.0	72	WP_117140040.1
478	ASMN341575v1	<i>hopAG1</i>	100	0.0	81	WP_010196193.1
478	ASMN341575v1	<i>hopAH1</i>	100	0.0	83	WP_081023459.1
478	ASMN341575v1	<i>hopAH1</i>	64	1 e-135	69	WP_039847471.1
478	ASMN341575v1	<i>hopAQ1</i>	60	2 e-13	63	WP_024668838.1

478	ASMN341575v1	<i>hopAR1</i>	100	2 e-158	78	WP_117140095.1
478	ASMN341575v1	<i>hopAS1</i>	92	0.0	95	WP_117140549.1
478	ASMN341575v1	<i>hopAY1</i>	71	7 e-12	28	WP_117140095.1
478	ASMN341575v1	<i>hopAZ1</i>	100	3 e-123	73	WP_024663017.1
478	ASMN341575v1	<i>hopBD1</i>	99	0.0	97	WP_055004698.1
478	ASMN341575v1	<i>hopBD2</i>	84	2 e-82	50	WP_055004698.1
478	ASMN341575v1	<i>hopBH1</i>	35	3 e-18	39	WP_117139931.1
479	ASMN341616v1	<i>avrD1</i>	100	0.0	98	WP_024671645.1
479	ASMN341616v1	<i>avrE1</i>	100	0.0	84	WP_117140309.1
479	ASMN341616v1	<i>avrRpm1</i>	96	5 e-80	59	WP_117140700.1
479	ASMN341616v1	<i>avrRpm2</i>	64	2 e-29	43	WP_117140700.1
479	ASMN341616v1	<i>avrRps4</i>	100	4 e-117	74	WP_024668838.1
479	ASMN341616v1	<i>hopD1</i>	100	0.0	93	WP_117140212.1
479	ASMN341616v1	<i>hopI1</i>	79	2 e-123	82	WP_057404518.1
479	ASMN341616v1	<i>hopk1</i>	43	3 e-70	76	WP_024668838.1
479	ASMN341616v1	<i>hopM1</i>	100	0.0	82	WP_117140307.1
479	ASMN341616v1	<i>hopR1</i>	100	0.0	86	WP_080953444.1
479	ASMN341616v1	<i>hopV1</i>	74	4 e-172	77	WP_005777017.1
479	ASMN341616v1	<i>hopW1</i>	39	4 e-79	49	WP_117140040.1
479	ASMN341616v1	<i>hopX1</i>	98	3 e-170	68	WP_080726063.1
479	ASMN341616v1	<i>hopAA1-1</i>	100	0.0	83	WP_117140304.1
479	ASMN341616v1	<i>hopAA1-2</i>	99	0.0	82	WP_117140304.1
479	ASMN341616v1	<i>hopAB1</i>	100	7 e-132	47	WP_117139356.1
479	ASMN341616v1	<i>hopAB2</i>	100	0.0	68	WP_117139356.1
479	ASMN341616v1	<i>hopAB3</i>	100	0.0	63	WP_117139356.1
479	ASMN341616v1	<i>hopAE1</i>	98	0.0	72	WP_117140040.1
479	ASMN341616v1	<i>hopAG1</i>	100	0.0	81	WP_010196193.1
479	ASMN341616v1	<i>hopAH1</i>	100	0.0	83	WP_081023459.1
479	ASMN341616v1	<i>hopAH1</i>	64	1 e-135	69	WP_039847471.1
479	ASMN341616v1	<i>hopAQ1</i>	60	2 e-13	63	WP_024668838.1
479	ASMN341616v1	<i>hopAR1</i>	100	2 e-158	78	WP_117140095.1
479	ASMN341616v1	<i>hopAS1</i>	92	0.0	95	WP_117140549.1
479	ASMN341616v1	<i>hopAY1</i>	71	7 e-12	28	WP_117140095.1
479	ASMN341616v1	<i>hopAZ1</i>	100	3 e-123	73	WP_024663017.1
479	ASMN341616v1	<i>hopBD1</i>	99	0.0	97	WP_055004698.1
479	ASMN341616v1	<i>hopBD2</i>	84	2 e-82	50	WP_055004698.1
479	ASMN341616v1	<i>hopBH1</i>	35	3 e-18	39	WP_117139931.1

Tabela Suplementar S5. Efeitores identificados nas cinco estirpes de *Pseudomonas syringae* pv. *tabaci* usando o *script* customizado em Python que localiza sequências assinaturas presentes nas regiões promotoras (caixa *hrp*).

Estirpe	Número de acesso	Qualificador	Efeitor
440	WP_117140309.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pst</i>	<i>avrE</i>
440	WP_080164413.1	Proteína hipotética	<i>Pst_EF1</i>
440	WP_010196193.1	Proteína com domínio NUDIX	<i>hopAG1</i>
440	WP_081023459.1	Efeitor de tipo III de <i>Pst</i>	<i>hopAH1</i>
440	WP_117140188.1	Proteína <i>helper</i> de tipo III HopAK1	<i>hopAK1</i>
440	WP_044299239.1	Peptidase da família M20	<i>Pst_EF2</i>
440	WP_117140220.1	N-acetiltransferase	<i>Pst_EF3</i>
440	WP_117140221.1	Proteína com domínio de adenilação de amino ácido	<i>Pst_EF4</i>
440	WP_005779785.1	Proteína hipotética	<i>Pst_EF5</i>
440	WP_024663017.1	Proteína com domínio DUF1826	<i>hopAZ1</i>
440	WP_117140040.1	Endonuclease de restrição	<i>hopAE1</i>
440	WP_080953444.1	Efeitor de tipo III de <i>Pst</i>	<i>hopR1</i>
440	WP_117140700.1	Efeitor de tipo III de <i>Pst</i>	<i>avrRpm1</i>
440	WP_117140304.1	Efeitor de tipo III de <i>Pst</i>	<i>hopAA1</i>
440	WP_054994611.1	1-Acil-sn-glicerol-3-fosfato aciltransferase	<i>Pst_EF6</i>
440	WP_002551687.1	Proteína carreadora de acila	<i>Pst_EF7</i>
440	WP_002551686.1	Proteína carreadora de acila	<i>Pst_EF8</i>
440	WP_117140001.1	Proteína hipotética	<i>Pst_EF10</i>
440	WP_080900592.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
440	WP_004664563.1	Piridina nucleotídeo transidrogenase solúvel	<i>Psgc_EF8</i>
440	WP_117140614.1	Enzima dependente de piridoxal fosfato	<i>Pst_EF11</i>
440	WP_117140615.1	N-acetiltransferase	<i>Pst_EF12</i>
440	WP_117139356.1	E3 ubiquitina-proteína ligase	<i>hopAB2</i>
440	WP_005777017.1	Proteína hipotética	<i>hopV1</i>
469	WP_005777017.1	Proteína hipotética	<i>hopV1</i>
469	WP_024663017.1	Proteína com domínio DUF1826	<i>hopAZ1</i>
469	WP_024671645.1	Proteína biosintética de siringolídeo AvrD1	<i>avrD1</i>
469	WP_080268302.1	Proteína da família histidine fosfatase	<i>Psgc_EF13</i>
469	WP_117140304.1	Efeitor de tipo III de <i>Pst</i>	<i>hopAA1</i>
469	WP_117140700.1	Efeitor de tipo III de <i>Pst</i>	<i>avrRpm1</i>
469	WP_117140188.1	Proteína <i>helper</i> de tipo III HopAK1	<i>hopAK1</i>
469	WP_117140040.1	Endonuclease de restrição	<i>hopAE1</i>
469	WP_117140614.1	Enzima dependente de piridoxal fosfato	<i>Pst_EF11</i>

469	WP_117140615.1	N-acetiltransferase	<i>Pst_EF12</i>
469	WP_080953444.1	Efector de tipo III de <i>Pst</i>	<i>hopR1</i>
469	WP_117139356.1	E3 ubiquitina-proteína ligase	<i>hopAB2</i>
469	WP_054994611.1	1-Acil-sn-glicerol-3-fosfato aciltransferase	<i>Pst_EF6</i>
469	WP_002551687.1	Proteína carreadora de acila	<i>Pst_EF7</i>
469	WP_002551686.1	Proteína carreadora de acila	<i>Pst_EF8</i>
469	WP_117140017.1	Histidina quinase sensora	<i>Pst_EF9</i>
469	WP_117140001.1	Proteína hipotética	<i>Pst_EF10</i>
469	WP_080900592.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
469	WP_004664563.1	Piridina nucleotídeo transidrogenase solúvel	<i>Psgc_EF8</i>
469	WP_117140309.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pst</i>	<i>avrE</i>
469	WP_117140688.1	Muramidase	<i>Pst_EF13</i>
469	WP_005779785.1	Proteína hipotética	<i>Pst_EF5</i>
469	WP_024668838.1	Proteína hipotética	<i>avrRps4</i>
478	WP_117140212.1	Proteína efectora AvrPphD de <i>Pst</i>	<i>avrPphD</i>
478	WP_080953444.1	Efector de tipo III de <i>Pst</i>	<i>hopR1</i>
478	WP_024668838.1	Proteína hipotética	<i>avrRps4</i>
478	WP_117140700.1	Efector de tipo III de <i>Pst</i>	<i>avrRpm1</i>
478	WP_117140040.1	Endonuclease de restrição	<i>hopAE1</i>
478	WP_024663017.1	Proteína com domínio DUF1826	<i>hopAZ1</i>
478	WP_117140188.1	Proteína <i>helper</i> de tipo III HopAK1	<i>hopAK1</i>
478	WP_080164413.1	Proteína hipotética	<i>Pst_EF1</i>
478	WP_010196193.1	Proteína com domínio NUDIX	<i>hopAG1</i>
478	WP_081023459.1	Efector de tipo III de <i>Pst</i>	<i>hopAH1</i>
478	WP_117140304.1	Efector de tipo III de <i>Pst</i>	<i>hopAA1</i>
478	WP_117140614.1	Enzima dependente de piridoxal fosfato	<i>Pst_EF11</i>
478	WP_117140615.1	N-acetiltransferase	<i>Pst_EF12</i>
478	WP_024671645.1	Proteína biosintética de siringolídeo AvrD1	<i>avrD1</i>
478	WP_080268302.1	Proteína da família histidine fosfatase	<i>Psgc_EF13</i>
478	WP_117140688.1	Muramidase	<i>Pst_EF13</i>
478	WP_054994611.1	1-Acil-sn-glicerol-3-fosfato aciltransferase	<i>Pst_EF6</i>
478	WP_002551687.1	Proteína carreadora de acila	<i>Pst_EF7</i>
478	WP_002551686.1	Proteína carreadora de acila	<i>Pst_EF8</i>
478	WP_117140017.1	Histidina quinase sensora	<i>Pst_EF9</i>
478	WP_117140001.1	Proteína hipotética	<i>Pst_EF10</i>
478	WP_080900592.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
478	WP_004664563.1	Piridina nucleotídeo transidrogenase solúvel	<i>Psgc_EF8</i>

478	WP_117140309.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pst</i>	<i>avrE</i>
478	WP_117139356.1	E3 ubiquitina-proteína ligase	<i>hopAB2</i>
478	WP_005779785.1	Proteína hipotética	<i>Pst_EF5</i>
479	WP_005779785.1	Proteína hipotética	<i>Pst_EF5</i>
479	WP_044299239.1	Peptidase da família M20	<i>Pst_EF2</i>
479	WP_117140220.1	N-acetiltransferase	<i>Pst_EF3</i>
479	WP_117140221.1	Proteína com domínio de adenilação de amino ácido	<i>Pst_EF4</i>
479	WP_117140688.1	Muramidase	<i>Pst_EF13</i>
479	WP_117139356.1	E3 ubiquitina-proteína ligase	<i>hopAB2</i>
479	WP_024663017.1	Proteína com domínio DUF1826	<i>hopAZ1</i>
479	WP_080953444.1	Efector de tipo III de <i>Pst</i>	<i>hopR1</i>
479	WP_117140040.1	Endonuclease de restrição	<i>hopAE1</i>
479	WP_117140212.1	Proteína efectora AvrPphD de <i>Pst</i>	<i>avrPphD</i>
479	WP_117140309.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pst</i>	<i>avrE</i>
479	WP_117140188.1	Proteína <i>helper</i> de tipo III HopAK1	<i>hopAK1</i>
479	WP_024668838.1	Proteína hipotética	<i>avrRps4</i>
479	WP_117140700.1	Efector de tipo III de <i>Pst</i>	<i>avrRpm1</i>
479	WP_117140001.1	Proteína hipotética	<i>Pst_EF10</i>
479	WP_080900592.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
479	WP_004664563.1	Piridina nucleotídeo transidrogenase solúvel	<i>Psgc_EF8</i>
479	WP_054994611.1	1-Acil-sn-glicerol-3-fosfato aciltransferase	<i>Pst_EF6</i>
479	WP_002551687.1	Proteína carreadora de acila	<i>Pst_EF7</i>
479	WP_002551686.1	Proteína carreadora de acila	<i>Pst_EF8</i>
479	WP_005777017.1	Proteína hipotética	<i>hopV1</i>
479	WP_117140614.1	Enzima dependente de piridoxal fosfato	<i>Pst_EF11</i>
479	WP_117140615.1	N-acetiltransferase	<i>Pst_EF12</i>
479	WP_024671645.1	Proteína biosintética de siringolídeo AvrD1	<i>avrD1</i>
479	WP_080268302.1	Proteína da família histidine fosfatase	<i>Psgc_EF13</i>
479	WP_117140304.1	Efector de tipo III de <i>Pst</i>	<i>hopAA1</i>
480	WP_117140212.1	Proteína efectora AvrPphD de <i>Pst</i>	<i>avrPphD</i>
480	WP_024668838.1	Proteína hipotética	<i>avrRps4</i>
480	WP_117140614.1	Enzima dependente de piridoxal fosfato	<i>Pst_EF11</i>
480	WP_117140615.1	N-acetiltransferase	<i>Pst_EF12</i>
480	WP_117140688.1	Muramidase	<i>Pst_EF13</i>
480	WP_044299239.1	Peptidase da família M20	<i>Pst_EF2</i>
480	WP_117140220.1	N-acetiltransferase	<i>Pst_EF3</i>
480	WP_117140221.1	Proteína com domínio de adenilação de amino ácido	<i>Pst_EF4</i>

480	WP_117140304.1	Efector de tipo III de <i>Pst</i>	<i>hopAA1</i>
480	WP_117140001.1	Proteína hipotética	<i>Pst_EF10</i>
480	WP_080900592.1	FAD:proteína FMN transferase	<i>Psgc_EF7</i>
480	WP_004664563.1	Piridina nucleotídeo transidrogenase solúvel	<i>Psgc_EF8</i>
480	WP_005779785.1	Proteína hipotética	<i>Pst_EF5</i>
480	WP_024671645.1	Proteína biosintética de siringolídeo AvrD1	<i>avrD1</i>
480	WP_080268302.1	Proteína da família histidine fosfatase	<i>Psgc_EF13</i>
480	WP_117140188.1	Proteína <i>helper</i> de tipo III HopAK1	<i>hopAK1</i>
480	WP_080953444.1	Efector de tipo III de <i>Pst</i>	<i>hopR1</i>
480	WP_005777017.1	Proteína hipotética	<i>hopV1</i>
480	WP_117140700.1	Efector de tipo III de <i>Pst</i>	<i>avrRpm1</i>
480	WP_117139356.1	E3 ubiquitina-proteína ligase	<i>hopAB2</i>
480	WP_054994611.1	1-Acil-sn-glicerol-3-fosfato aciltransferase	<i>Pst_EF6</i>
480	WP_002551687.1	Proteína carreadora de acila	<i>Pst_EF7</i>
480	WP_002551686.1	Proteína carreadora de acila	<i>Pst_EF8</i>
480	WP_117140309.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pst</i>	<i>avrE</i>
480	WP_024663017.1	Proteína com domínio DUF1826	<i>hopAZ1</i>

Tabela Suplementar S6. Efeitores de *Pseudomonas cichorii* identificados por Blastp contra o banco de dados de efetores depositados no *Pseudomonas syringae* Genome Resource.

Estirpe	Número de acesso	Efeitor	Cobertura (%)	E-value	Identidade (%)	ID da Proteína
481	ASMN341610v1	<i>avrE1</i>	90	0.0	44	WP_117185806.1
481	ASMN341610v1	<i>hopB1</i>	44	1 e-19	31	WP_117186690.1
481	ASMN341610v1	<i>hopBN1</i>	99	2 e-153	67	WP_117184493.1
482	ASMN341569v1	<i>avrE1</i>	90	0.0	44	WP_117185806.1
482	ASMN341569v1	<i>hopB1</i>	44	1 e-19	31	WP_117186690.1
482	ASMN341569v1	<i>hopBN1</i>	99	2 e-153	67	WP_117184493.1
473	ASMN341577v1	<i>avrE1</i>	87	0.0	44	WP_117165072.1
473	ASMN341577v1	<i>hopA1</i>	99	8 e-123	47	WP_117167454.1
473	ASMN341577v1	<i>hopB1</i>	45	1 e-20	31	WP_117167277.1
474	ASMN341621v1	<i>avrE1</i>	90	0.0	44	WP_117184442.1
474	ASMN341621v1	<i>hopB1</i>	44	1 e-19	31	WP_117181674.1
474	ASMN341621v1	<i>hopBN1</i>	99	2 e-153	67	WP_117184493.1
136	ASMN341675v1	<i>avrE1</i>	90	0.0	44	WP_117184442.1
136	ASMN341675v1	<i>hopB1</i>	44	1 e-19	31	WP_117181674.1
136	ASMN341675v1	<i>hopBN1</i>	99	2 e-153	67	WP_117184493.1

Tabela Suplementar S7. Efeitores identificados nas cinco estirpes de *Pseudomonas cichorii* usando o *script* customizado em Python que localiza sequências assinaturas presentes nas regiões promotoras (caixa *hrp*).

Estirpe	Número de acesso	Qualificador	Efeitor
136	WP_117184385.1	Fator de alongação da transcrição GreB	<i>Pc_EF1</i>
136	WP_025260324.1	Proteína hipotética	<i>Pc_EF2</i>
136	WP_117167452.1	Proteína hipotética	<i>Pc_EF3</i>
136	WP_117184442.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pc</i>	<i>avrE</i>
136	WP_117184441.1	Aspartil beta-hidroxilase	<i>Pc_EF4</i>
136	WP_117184440.1	Proteína hipotética	<i>Pc_EF5</i>
136	WP_117182261.1	Metiltransferase dependente de SAM da classe I	<i>Pc_EF6</i>
136	WP_117182259.1	Proteína com domínio DUF2878	<i>Pc_EF7</i>
136	WP_117166950.1	Proteína do <i>cluster</i> cisteína da família YkgJ	<i>Pc_EF8</i>
473	WP_117165072.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pc</i>	<i>avrE</i>
473	WP_117165073.1	Aspartil beta-hidroxilase	<i>Pc_EF4</i>
473	WP_117165074.1	Proteína com domínio DUF4189	<i>Pc_EF9</i>
473	WP_025260324.1	Proteína hipotética	<i>Pc_EF2</i>
473	WP_117164299.1	Fator de alongação da transcrição GreB	<i>Pc_EF1</i>
473	WP_117167454.1	Proteína hipotética	<i>Pc_EF10</i>
474	WP_117184385.1	Fator de alongação da transcrição GreB	<i>Pc_EF1</i>
474	WP_117184442.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pc</i>	<i>avrE</i>
474	WP_117184441.1	Aspartil beta-hidroxilase	<i>Pc_EF4</i>
474	WP_117184440.1	Proteína hipotética	<i>Pc_EF5</i>
474	WP_117167452.1	Proteína hipotética	<i>Pc_EF3</i>
474	WP_025260324.1	Proteína hipotética	<i>Pc_EF2</i>
474	WP_117182261.1	Metiltransferase dependente de SAM da classe I	<i>Pc_EF6</i>
474	WP_117182259.1	Proteína com domínio DUF2878	<i>Pc_EF7</i>
474	WP_117166950.1	Proteína do <i>cluster</i> cisteína da família YkgJ	<i>Pc_EF8</i>
481	WP_025260324.1	Proteína hipotética	<i>Pc_EF2</i>
481	WP_117185806.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pc</i>	<i>avrE</i>
481	WP_117185807.1	Aspartil beta-hidroxilase	<i>Pc_EF4</i>
481	WP_117184440.1	Proteína hipotética	<i>Pc_EF5</i>
481	WP_117186066.1	Proteína hipotética	<i>Pc_EF3</i>
481	WP_117186810.1	Fator de alongação da transcrição GreB	<i>Pc_EF1</i>
482	WP_117186810.1	Fator de alongação da transcrição GreB	<i>Pc_EF1</i>
482	WP_117185806.1	Proteína da família AvrE de efetores de tipo 3 de <i>Pc</i>	<i>avrE</i>
482	WP_117185807.1	Aspartil beta-hidroxilase	<i>Pc_EF4</i>

482	WP_117184440.1	Proteína hipotética	<i>Pc_EF5</i>
482	WP_025260324.1	Proteína hipotética	<i>Pc_EF2</i>
482	WP_117186066.1	Proteína hipotética	<i>Pc_EF3</i>

Capítulo 3: Efetores de tipo II comuns e específicos de *Pseudomonas* patogênicas ao cafeeiro em Minas Gerais

Francisco Henrique Nunes da Silva Alves¹, Jorge Luis Badel¹

¹ Laboratório de Fitobacteriologia Molecular, Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG, Brasil.

Autor para correspondência: Jorge Luis Badel, e-mail: jorge.badel@ufv.br

RESUMO

Entre as doenças bacterianas sofridas pela cultura do café, as manchas causadas pelas bactérias Gram-negativas, *P. syringae* pv. *garcae* (*Psgc*), *P. syringae* pv. *tabaci* (*Psta*) e *Pseudomonas cichorii* (*Pch*) são de grande importância. Apesar destas espécies/patovares causarem grandes perdas na cultura do cafeeiro, não existem relatos sobre os mecanismos moleculares subjacentes à capacidade de *Psgc*, *Psta* e *Pch* causarem doenças neste hospedeiro. Como consequência, este trabalho teve como objetivo obter informações úteis para entender os mecanismos moleculares específicos e comuns, relacionados com o sistema de secreção do tipo II (T2SS), utilizados pelas *Pseudomonas* patogênicas ao cafeeiro. Para tanto, as sequências das proteínas obtidas do banco de dados (GenBank) foram analisadas para prever a presença de peptídeo sinal e ausência de domínio transmembranar, usando os programas SignalP e TMHMM, respectivamente. Em seguida, as proteínas foram submetidas à ferramenta CD-HIT, para retirada das proteínas redundantes, e na plataforma GOfeat para anotação funcional. Assim, foi identificado a presença de 368 proteínas não redundantes que possuíam sinal de secreção de tipo II e ausência de domínio transmembranar em *Psgc*, 420 em *Psta* e 459 em *Pch*, sendo que somente 240 proteínas estiveram presentes em todas as espécies/patovares analisadas. As proteínas não redundantes identificadas usando termos GO, estiveram relacionadas com atividade enzimática ou transportadora de diferentes substâncias, envolvidas em processos biossintéticos e catabólicos, grande parte localizadas na membrana, fazendo parte do pilus ou do flagelo. Filtragem manual para retirar proteínas associadas à célula bacteriana, demonstrou que algumas proteínas extracelulares

secretadas pelo T2SS possuem atividades enzimáticas que não foram associadas com patogenicidade em estudos prévios. Por fim, os resultados do presente trabalho servem como ponto de partida para trabalhos futuros visando obter conhecimento sobre a importância do T2SS em *Psgc*, *Psta* e *Pch* e para desenvolver novas estratégias que visem prevenir o desenvolvimento das manchas foliares causadas por esses patógenos em plantas de café.

Palavras-chave: *Coffea*, domínio transmembranar, patogenicidade, peptídeo sinal, *tabaci*.

INTRODUÇÃO

O café arábica (*Coffea arábica* L.) e o café robusta (*Coffea canéfora* L.), estão entre as commodities mais vendidas no mundo e são de grande importância para a economia brasileira (DaMatta & Ramalho, 2004). Dentre os patógenos bacterianos que causam doenças em plantas de café, *P. syringae* pv. *garcae* (*Psgc*) (Amaral *et al.*, 1956), *P. syringae* pv. *tabaci* (*Psta*) (Rodrigues Neto *et al.*, 2006) e *Pseudomonas cichorii* (*Pch*) (Swingle 1925) podem causar perdas estimadas de até 70% das mudas em condições de viveiro e de campo (Rodrigues *et al.*, 2013). Apesar disso, não existem trabalhos relacionados aos mecanismos moleculares que permitam compreender como essas bactérias afetam plantas de café (Badel & Zambolim 2019).

Em bactérias fitopatogênicas, existem diversos sistemas de secreção que controlam a exportação de proteínas efetoras para o meio extracelular ou para o interior da célula hospedeira (Chang *et al.*, 2014; Costa *et al.*, 2015). Dentre os sistemas de secreção, o tipo II (T2SS), é amplamente presente em *Proteobacteria*, ocorrendo em muitas, porém não todas, espécies das classes *Alphaproteobacteria*, *Betaproteobacteria*, *Deltaproteobacteria* e *Gammaproteobacteria*, que causam graves doenças em frutos, cereais e vegetais (Cianciotto & White 2017).

Entre as bactérias fitopatogênicas, a importância do T2SS é conhecida expressiva em espécies como *Dickeya dadantii* e *Pectobacterium carotovorum*, que causam podridão mole em inúmeras plantas hospedeiras (Charkowski 2018). A importância do T2SS nas bactérias *Ralstonia solanacearum*, *Xanthomonas*

campestris, *Xanthomonas oryzae*, *Erwinia amylovora* e *Xylella fastidiosa*, que causam outros sintomas, também foi relatada em diferentes estudos (Cianciotto & White 2017).

As proteínas que são secretadas pelo T2SS possuem um sinal na extremidade N-terminal que possibilita serem reconhecidas pelas proteínas que fazem parte do aparato de secreção localizadas na membrana, que por sua vez, as exportam até o espaço periplasmático (Cianciotto & White 2017; Costa *et al.*, 2015). Uma vez no espaço periplasmático, o peptídeo sinal é clivado e a proteína é enovelada para ser secretada para o espaço extracelular (Cianciotto & White 2017; Costa *et al.*, 2015). As proteínas secretadas podem atuar como efetores (toxina, enzima, etc), acarretando na morte das células do hospedeiro, degradação de tecido, supressão da resposta imune e favorecendo para formação de biofilme e assimilação de nutrientes por parte do patógeno (Cianciotto & White 2017; Korotkov *et al.*, 2012).

No entanto, algumas proteínas que possuem sinal de secreção não são secretadas ao meio exterior, devido a presença de domínios transmembranares, permanecendo ancoradas à membrana, onde cumprem suas funções biológicas (Sonnhammer *et al.*, 1998). Essas características fazem possível prever o repertório de proteínas que são secretadas através do T2SS usando algoritmos como o SignalP, que prediz o sinal de secreção (Armenteros *et al.*, 2019), e TMHMM que indica a presença de domínios transmembranares (Sonnhammer *et al.*, 1998). Proteínas com predição positiva no signalP e negativa em TMHMM podem ser considerados efetores de tipo II.

Sendo assim, pretendendo encontrar mecanismos de adaptação comuns entre as espécies/patovares de *Pseudomonas* patogênicas ao cafeeiro, este trabalho visou identificar as proteínas secretadas pelo T2SS e prever os efetores do tipo II (T2SE) presentes em genomas de *Psgc*, *Psta* e *Pch* depositados nas bases de dados. Para essa análise, os programas SignalP e TMHMM foram utilizados para identificar as proteínas com base na presença do sinal de secreção do tipo II e ausência de domínio transmembranar, respectivamente.

MATERIAL E MÉTODOS

Obtenção das sequências das proteínas

As sequências das proteínas preditas nos genomas rascunho de diferentes estirpes de *P. syringae* pv. *garcae*, *P. syringae* pv. *tabaci* e *P. cichorii*, patogênicas ao cafeeiro foi obtido, em formato fasta, no banco de dados (GenBank) do *National Center for Biotechnology Information* (NCBI, <http://www.ncbi.nlm.nih.gov>).

Predição e proteômica comparativa das proteínas secretadas pelo T2SS

Para a predição das proteínas secretadas pelo T2SS de *Psgc*, *Psta* e *Pch*, as proteínas, de cada espécie/patovar, foram analisadas com relação à presença de peptídeo sinal, usando o programa SignalP versão 5.0 (<http://www.cbs.dtu.dk/services/SignalP/index.php>) (Armenteros *et al.*, 2019) e a ausência de domínio transmembranar, usando o TMHMM Server V.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) (Sonnhammer *et al.*, 1998).

Os resultados dos programas SignalP e TMHMM foram então analisados em conjunto usando a ferramenta CD-HIT Suite (http://weizhong-lab.ucsd.edu/cdhit_suite/cgi-bin/index.cgi?cmd=cd-hit) (Huang *et al.* 2010), para retirar as proteínas redundantes, usando um limiar de 50% de identidade, sendo gerado um proteoma do sistema de secreção do tipo II para cada espécie/patovar. Os resultados são apresentados em diagrama de Venn, usando a ferramenta Jvenn (<http://jvenn.toulouse.inra.fr/app/example.html>) (Bardou *et al.*, 2014).

Em seguida, as proteínas pertencentes aos proteomas do T2SS de *Psgc*, *Psta* e *Pch* foram submetidas à plataforma GOfeat (<http://computationalbiology.ufpa.br/gofeat/index/login>), para anotação funcional baseadas em sintenia de sequência e ontologia gênica (Araujo *et al.*, 2018). Em decorrência que muitas proteínas encontradas estão associadas com a membrana, formação do pilus ou flagelo, foi realizada uma filtragem manual dos dados para correta seleção das proteínas secretadas ao meio extracelular e candidatas a efetores do tipo II.

RESULTADOS

Pansecretomas do tipo II

Análise dos secretomas das cinco estirpes de *Psgc*, com base na presença do sinal de secreção do tipo II e ausência de domínio transmembranar, permitiu a identificação de 377 a 383 proteínas. Após a retirada das proteínas redundantes usando CD-HIT, foi determinado que *Psgc* possui em seu pansecretoma do T2SS a presença de 368 proteínas não redundantes, distribuídas entre as estirpes analisadas (Figura 1). Algumas dessas proteínas foram encontradas exclusivamente em algumas estirpes, sendo cinco proteínas encontradas apenas na estirpe 445, e duas nas estirpes 438 e 460 cada.

Em *Psta*, o pansecretoma do T2SS, predito com base na presença do sinal de secreção do tipo II e ausência de domínio transmembranar, analisando as proteínas das cinco estirpes, foi de 428 a 436 proteínas. Após a retirada das proteínas com funções repetitivas pelo CD-HIT, foi definido que *Psta* possui 420 proteínas não redundantes em seu pansecretoma do tipo II, distribuídas entre as cinco estirpes analisadas (Figura 2). Apenas a estirpe 480 apresentou proteínas exclusivas.

Seguindo as mesmas abordagens de predição anteriormente descritas, foi determinado que *Pch* possui 459 proteínas não redundantes, presentes entre as cinco estirpes analisadas (Figura 3). Proteínas exclusivas foram observadas nas estirpes 482 (5), 473 (2) e 136 (4).

***P. cichorii* possui o maior número de proteínas específicas, associadas ao T2SS.**

Quando analisados os pansecretomas do T2SS de *Psgc*, *Psta* e *Pch*, foi possível a identificação do número de proteínas compartilhadas e exclusivas de cada espécie/patovar de *Pseudomonas* patogênica ao cafeeiro. Foi constatado que 42 proteínas não redundantes foram exclusivas de *Psgc*, 68 únicas em *Psta* e 159 presentes apenas em *Pch*. Os resultados também demonstraram que do total de

proteínas não redundantes, somente 240 estiveram presentes em todas as espécies/patovares analisadas (Figura 4).

Identificação das proteínas dos secretomas do T2SS

No processo de identificação, usando a plataforma GOfeat, das proteínas exclusivas e compartilhadas do secretoma do tipo II, entre as três espécies/patovares de *Pseudomonas* analisadas, foi possível a sua classificação em três categorias: função molecular, componente celular e processo biológico. Na análise de ontologia gênica, uma determinada proteína pode ser classificada nas três categorias diferentes, de acordo com a conservação da sequência quando comparada com as proteínas nas bases de dados que foram caracterizadas funcionalmente em estudos prévios.

Assim, das 42 proteínas não redundantes exclusivas de *Psgc*, 20 obtiveram termo GO designados, totalizando 27 termos GO, sendo 13 proteínas (48%) relacionadas com função molecular, 10 (37%) com processo biológico e 4 (15%) com componente celular (Figura 5A). Todos os termos GO dentro da categoria função molecular apresentaram apenas uma proteína, sendo em sua maioria relacionadas com atividade enzimática ou transportadora de diferentes substâncias (Figura 5B). Na categoria processo biológico, foram encontradas proteínas relacionadas com o transporte de diversos tipos de compostos e envolvidas em processos biossintéticos e catabólicos (Figura 5C). Na categoria de componente celular, todos os termos GO apresentaram apenas uma proteína, sendo em sua maioria localizadas na membrana, fazendo parte do pilus ou do flagelo (Figura 5D).

Das 68 proteínas não redundantes, exclusivas de *Psta*, 40 obtiveram termo GO, totalizando 74 termos GO, sendo 30 (41%) proteínas agrupadas em função molecular, 17 (23%) em componente celular e 27 (36%) em processo biológico (Figura 6A). A maioria das proteínas na categoria função molecular são associadas a atividades enzimáticas (Figura 6B). Em processo biológico, em sua grande parte as proteínas são associadas com o transporte de diversos tipos de compostos (Figura 6C). Todas as proteínas na categoria componente celular são localizadas na membrana, fazem parte do pilus ou de complexos transportadores do tipo ABC (Figura 6D).

Oitenta e cinco das 159 proteínas não redundantes de *Pch* obtiveram termo GO designado. No total, 154 termos GO foram denominados, sendo 63 (41%) das proteínas reunidas na categoria função molecular, 40 (26%) em processo biológico e 51 (33%) em componente celular (Figura 7A). Na categoria função molecular, as proteínas são relacionadas com atividades enzimáticas ou transportadoras (Figura 7B). Em processo biológico, as proteínas estão relacionadas com o transporte de diversos tipos de compostos, processos biossintéticos e metabólicos (Figura 7C). Em componente celular, o maior número de proteínas foi identificado como componentes da membrana externa (Figura 7D).

Das 240 proteínas não redundantes, presentes em todas as espécies/patovares analisadas, 128 proteínas obtiveram GO designados, totalizando 234 termos GO. Cento e duas (44%) proteínas foram classificadas na categoria função molecular, 64 (27%) em processo biológico e 68 (29%) em componente celular (Figura 8A). Na categoria função molecular, 31 proteínas foram identificadas com função de ATPase associada ao transporte de substâncias através da membrana e as outras associadas a funções enzimáticas e de transporte de diversas substâncias (Figura 8B). Em processo biológico, as proteínas foram relacionadas ao transporte de diversos tipos de moléculas, além de processos biossintéticos e metabólicos (Figura 8C). Na categoria componente celular, 31 das proteínas foram relacionadas ao complexo transportador do tipo ABC e o resto a componentes da parede celular ou do pilus (Figura 8D).

As proteínas não redundantes das espécies/patovares que obtiveram GO designados e que não foram relacionadas ao pilus, flagelo ou ancoradas nas membranas, foram destacadas e consideradas potenciais efetores do tipo II. Assim, 11 proteínas não redundantes, exclusivas de *Psgc* (Tabela 1), 13 exclusivas de *Psta* (Tabela 2), 28 únicas de *Pch* (Tabela 3) e 44 proteínas comuns entre todas as espécies/patovares (Tabela 4), foram consideradas potenciais efetores do tipo II e em sua grande maioria associadas com atividades enzimáticas.

DISCUSSÃO

Nesse estudo, por meio de métodos de biologia computacional, foi predito as proteínas secretadas pelo sistema de secreção do tipo II de espécies/patovares que

causam mancha foliar em *Coffea* spp. Muitas das proteínas que foram preditas como possuidoras de peptídeo sinal para serem secretadas pelo T2SS e que não possuem domínio transmembrana preditos pelo TMHMM, tiveram funções associadas com a membrana, significando que talvez estas proteínas não tenham um domínio transmembrana clássico ou que o TMHMM não tenha conseguido de fato identificar algumas proteínas que se ancoram nas membranas bacterianas.

Os repertórios de efetores do T2SS de *Psgc*, *Psta* e *Pch* patogênicas ao cafeeiro ainda não foram descritos na literatura. A importância do T2SS nas interações patógeno/hospedeiro já é conhecida (Jha *et al.*, 2005). Proteínas secretadas pelo T2SS estão envolvidas na colonização do hospedeiro e destruição do tecido. A ação das proteínas secretadas pelo T2SS também causa indução de respostas de defesa da planta (Jha *et al.*, 2005).

Proteínas secretadas pelo T2SS são particularmente importantes em bactérias que causam podridão-mole nas suas plantas hospedeiras, tais como espécies de *Pectobacterium* e *Dickeya* (Charkowski 2018). No entanto, também têm sido relatadas tendo função em patógenos vasculares (Charkowski 2018). A maioria dessas proteínas possuem atividade enzimática importante para a degradação da parede celular vegetal, como pectina liase, poligalacturonase e protease.

Quanto ao envolvimento de proteínas secretadas pelo T2SS em bactérias que causam manchas ou crestamentos foliares, enzimas como lipase, poligalacturonase, celulase, pectina liase e outras, têm sido relatadas como fundamentais para degradação do tecido do hospedeiro e para o desenvolvimento das doenças causadas por espécies do gênero *Xanthomonas* (Assis *et al.*, 2017; Esquerré-Tugayé *et al.*, 2000). Em *Xanthomonas campestris* pv. *campestris*, várias enzimas, incluindo amilases, proteases e celulases são secretadas pelo sistema de secreção do tipo II (Hu *et al.*, 1992; Lee *et al.*, 2004), em *Xanthomonas oryzae*, proteínas com atividade de lipase e celulase foram identificadas como efetores do tipo II que induzem a resposta de defesa em plantas de arroz (Assis *et al.*, Jha *et al.*, 2007). Por outro lado, um estudo desenvolvido com *Burkholderia glumae* detectou a presença de 34 proteínas secretadas pelo T2SS, dentre elas lipases, proteases, assim como 16 proteínas que não possuíam sinal característico do sistema de secreção do tipo II e diversas proteínas cujas funções não foram ainda caracterizadas (Goo *et al.*, 2010). Neste estudo, das proteínas identificadas no pansecretoma do tipo II de *Psgc*, *Psta* e *Pch*, foi encontrado um total de 236 proteínas que não apresentaram similaridade

significativa com proteínas anotadas, mas que possuem potencial para novos tipos de efetores com importante contribuição para a patogenicidade bacteriana.

Dentre as atividades enzimáticas previamente relatadas, neste trabalho somente foram encontradas uma poligalacturonase no pansecretoma de *Psta*, e diversas peptidases comuns nos pangenomas das três espécies/patovares analisadas. Diferentes funções enzimáticas e algumas proteínas não caracterizadas foram encontradas. Contudo, a importância desses T2SE, candidatos para a patogenicidade das *Pseudomonas* ao cafeeiro, deve ser demonstrada através de estudos funcionais.

Este estudo foi conduzido pela motivação de entender os mecanismos moleculares comuns que estão associados à adaptação dos patovares de *P. syringae* e da espécie *P. cichorii* em causar doença no cafeeiro. Para tanto, as proteínas comuns secretadas pelo T2SS foram preditas. Encontrou-se um repertório de 44 proteínas que poderiam estar envolvidas no processo de adaptação, a maioria possuindo atividades enzimáticas que não foram descritas como importantes nas interações planta-bactéria. Destaca-se nesse repertório, proteínas envolvidas em manter um balanço redox, como no caso das redutases, oxidases, catalase e glutathione peroxidase. Esses resultados são condizentes com o fato que a primeira linha de defesa da planta que as bactérias devem suplantar é o PTI, durante o qual espécies reativas de oxigênio (ROS) são produzidas (Jones & Dangl 2006; Zipfel & Robatzek 2010). Dentre as atividades comuns aos três patógenos do cafeeiro, destacam-se o número elevado de enzimas que atuam sobre proteínas ou peptídeos, sugerindo que uma das principais funções dos T2SE é a degradação de proteínas para suprimir as respostas de defesa ou para adquirir nutrientes.

A aquisição de nutrientes do apoplasto da planta, implica que as bactérias devem possuir mecanismos para capturá-los e importa-los ao seu interior. Nesse processo, as proteínas transportadoras de membrana cumprem um papel muito importante. Por esse motivo, neste trabalho optou-se por descrever, além dos candidatos a T2SE, as proteínas ancoradas nas membranas. Dentre essas proteínas, destaca-se o número elevado que formam parte de transportadores de tipo ABC, presentes entre os três patógenos do cafeeiro. No entanto, os resultados obtidos neste estudo são ainda muito preliminares, e estudos focados na importância desse tipo de proteínas na adaptação ao cafeeiro devem ser desenvolvidos.

Por fim, os resultados do presente trabalho servem como ponto de partida para trabalhos futuros visando obter conhecimento sobre a importância do T2SS em *Psgc*, *Psta* e *Pch* e para desenvolver novas estratégias que visem prevenir o desenvolvimento das manchas foliares causadas por esses patógenos em plantas de café.

REFERÊNCIAS

Amaral JF, Teixeira C, Pinheiro EDA, 1956. Bactéria causadora da mancha aureolada do cafeeiro. *Arquivos do Instituto Biológico* **23**, 151-155.

Araujo FA, Barh D, Silva A, Guimarães L, Thiago R, Ramos J, 2018. Go Feat: a rapid web-based functional annotation tool for genomic and transcriptomic data. *Scientific Reports* **8**, 8-11.

Armenteros JJA, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, Heijne GV, Nielsen H, 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology* **37**, 420-423.

Assis RAB, Polloni LC, Patané JSL, Thakur S, Felestino EB, Diaz-Caballero J, Digiampietri LA, Goulart LR, Almeida NF, Nascimento R, Dandekar AM, Zaini PA, Setubal JC, Guttman DS, Moreira LM, 2017. Identification and analysis of seven effector protein families with different adaptive and evolutionary histories of the Xanthomonadaceae. *Scientific Reports* **7**, 16133.

Badel JL, Zambolim L, 2019. Coffee bacterial diseases: a plethora of scientific opportunities. *Plant Pathology* **68**, 411- 425.

Bardou P, Mariette J, Escudie F, Djemiel C, Klopp C, 2014. Jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* **15**, 293.

Chang JH, Desveaux D, Creason AL, 2014. The ABCs and 123s of bacterial secretion systems in plant pathogenesis. *Annual Review of Phytopathology* **52**, 317–345.

Charkowski AO, 2018. The changing face of bacterial soft-rot diseases. *Annual Review of Phytopatology* **56**, 269-288.

Cianciotto NP, White RC, 2017. Expanding role of type II secretion in bacterial pathogenesis and beyond. *Infection and Immunity* **85**, 11-16.

Costa TR, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Review of Microbiology* **13**, 343–359.

DaMatta FH, Ramalho JDC, 2004. Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology* **16**, 55-81.

Esquerré-Tugayé MT, Boudart G, Bernard Dumas B, 2000. Cell wall degrading enzymes, inhibitory proteins, and oligosaccharides participate in the molecular dialogue between plants and pathogens. *Plant Physiology and Biochemistry* **38**, 157–163.

Goo E, Kang Y, Kim H, Hwang I, 2010. Proteomic analysis of quorum sensing-dependent proteins in *Burkholderia glumae*. *Journal of Proteome Research* **6**, 3184-3199.

Hu NT, Hung M N, Chiou SJ, Tang F, Chiang DC, Huang HY, Wu CY, 1992. Cloning and characterization of a gene required for the secretion of extracellular enzymes across the outer membrane by *Xanthomonas campestris* pv. *campestris*. *Journal of Bacteriology* **174**, 2679–2687.

Huang Y, Niu B, Gao Y, Fu L, Li W, 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**: 680-682.

Jha G, Rajeshwari R, Sonti RV, 2005. Bacterial type two secretion system secreted proteins: double-edged swords for plant pathogens. *Molecular Plant-Microbe Interactions* **18**, 891–898.

Jha G, Rajeshwari R, Sonti RV, 2007. Functional interplay between two *Xanthomonas oryzae* pv. *oryzae* secretion systems in modulating virulence on rice. *Molecular Plant-Microbe Interactions* **20**, 31–40.

Jones JD, Dangl JL, 2006. The plant immune system. *Nature* **444**, 323-329.

Korotkov KV, Sandkvist M, Hol WG. 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. *Nature Review of Microbiology* **10**, 336–351.

Lee H-M, Chen J-R, Lee H-L, Leu W-M, Chen L-Y, Hu N-T, 2004. Functional dissection of the XpsN (GspC) protein of the *Xanthomonas campestris* pv. *campestris* type II secretion machinery. *Journal of Bacteriology* **186**, 2946-2955.

Rodrigues LMR, Almeida IMG, Patricio FRA, 2013. Mancha Aureolada do Cafeeiro Causada por *Pseudomonas syringae* pv. *garcae*. Campinas, SP, Brasil: Instituto Agronômico de Campinas.

Rodrigues Neto J, Silva CHD, Beriam LOS, 2006. Mancha bacteriana do cafeeiro causada por *Pseudomonas syringae* pv *tabaci*. *Summa Phytopathologica* **32**, 85.

Sonnhammer ELL, Heijne GV, Krogh A, 1998. A hidden markov model for predicting transmembrane hélices in protein sequences. *Proceeding of the International Conference on Intelligent Systems for Molecular Biology* **6**, 175-182.

Zipfel C, Robatzek S, 2010. Pathogen-associated molecular pattern-triggered immunity. *Plant Physiology* **154**, 551-554.

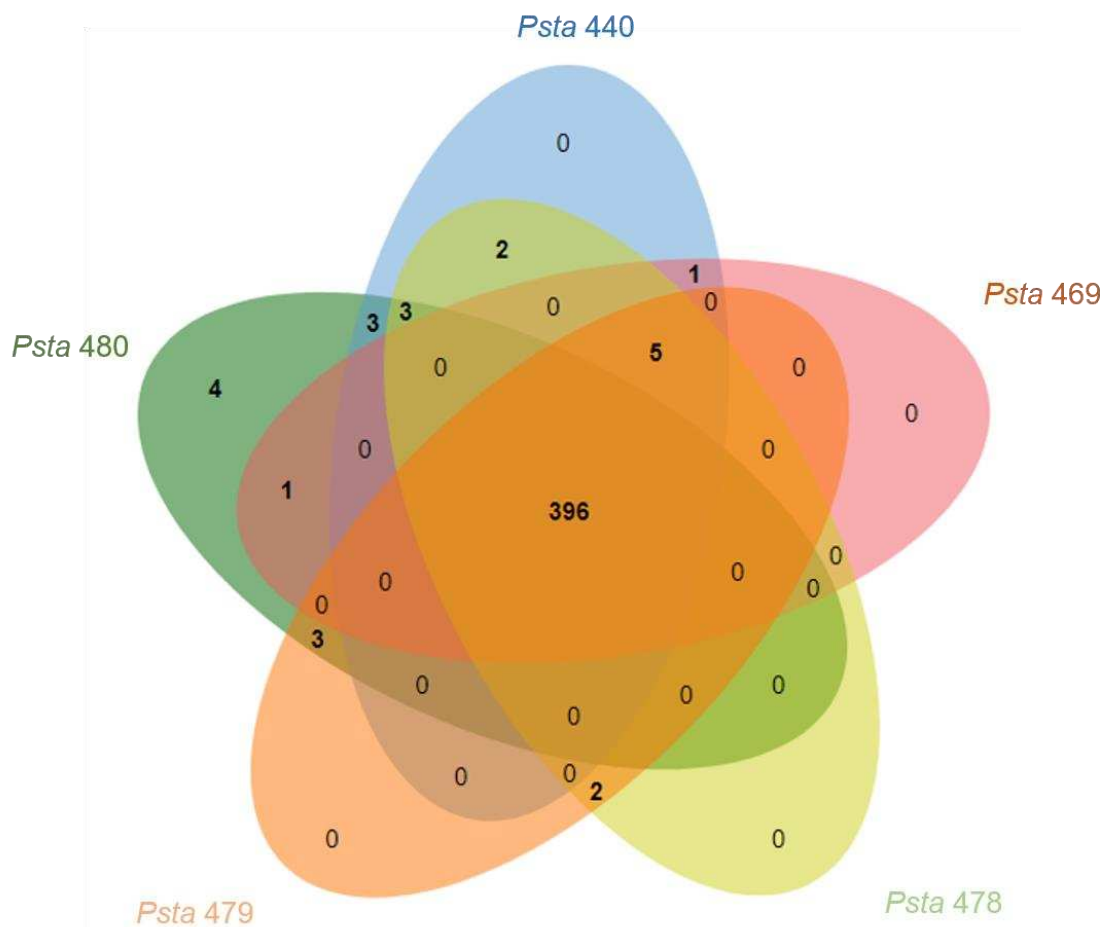


Figura 2. Diagrama de Venn do pansecretoma do T2SS de *Pseudomonas syringae* pv. *tabaci*, gerado com base na presença do sinal de secreção do tipo II e ausência de domínio transmembranar. Os números nas interseções indicam as proteínas que são compartilhadas entre uma ou mais estirpes. O número de proteínas de cada estirpe é agrupado por uma cor.

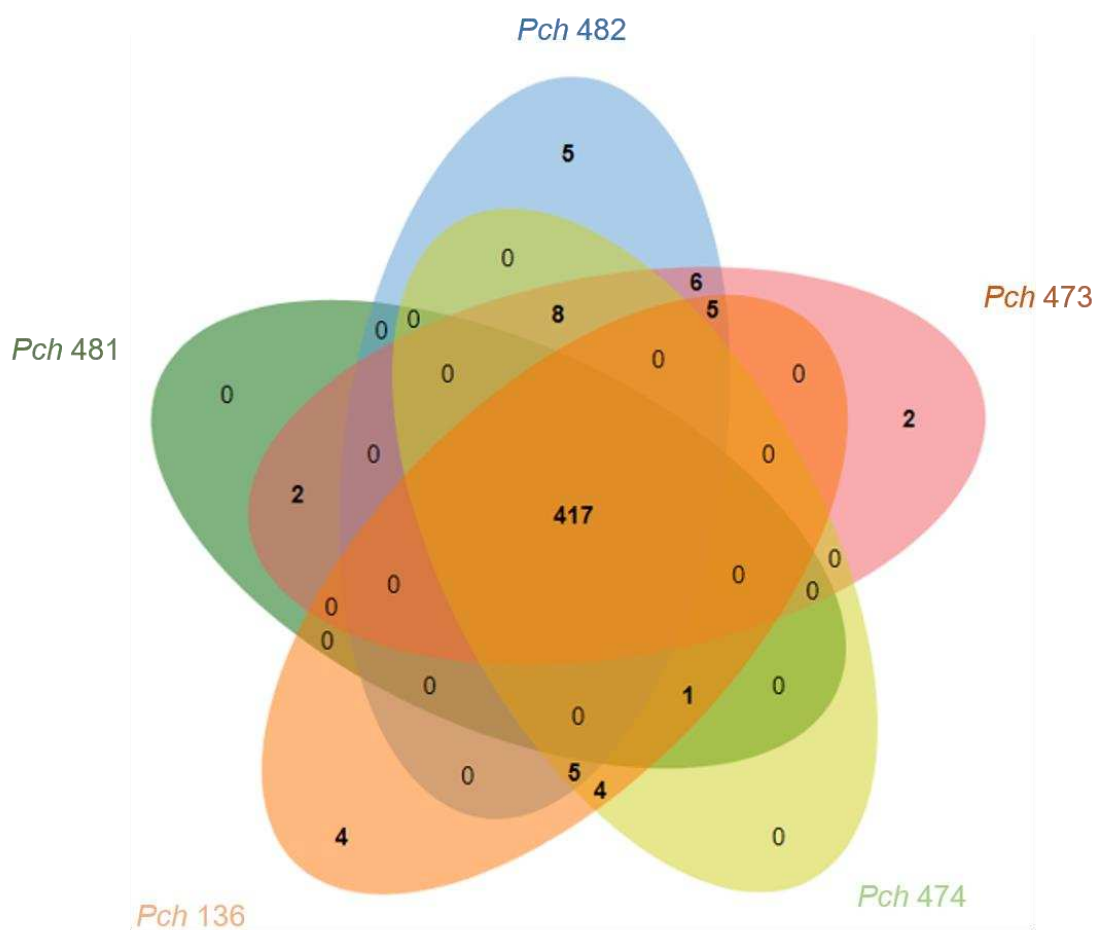


Figura 3. Diagrama de Venn do pansecretoma do T2SS de *Pseudomonas cichorii*, gerado com base na presença do sinal de secreção do tipo II e ausência de domínio transmembranar. Os números nas interseções indicam as proteínas que são compartilhadas entre duas ou mais estirpes. O número de proteínas de cada estirpe é agrupado por uma cor.

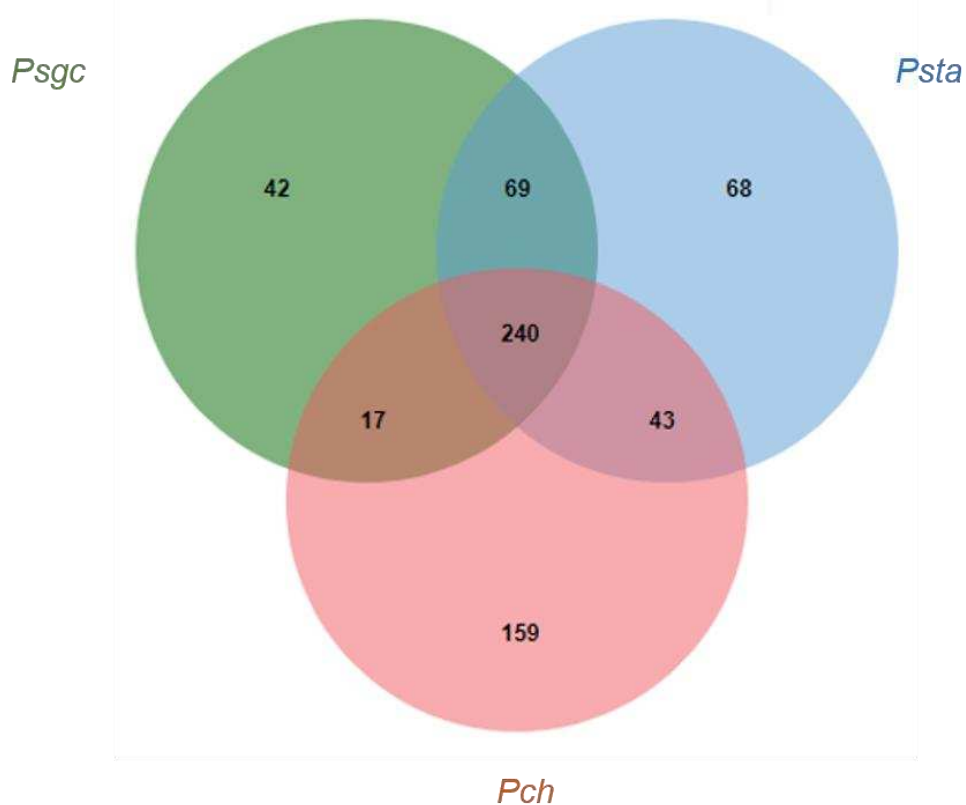


Figura 4. Diagrama de Venn demonstrando as proteínas compartilhadas e exclusivas das espécies/patovares bacterianas patogênicas ao cafeeiro, determinadas pelo secretoma do tipo II. Os números nas interseções indicam as proteínas que são compartilhadas entre duas ou três espécies/patovares. O número de proteínas de cada patógeno é agrupado por uma cor.

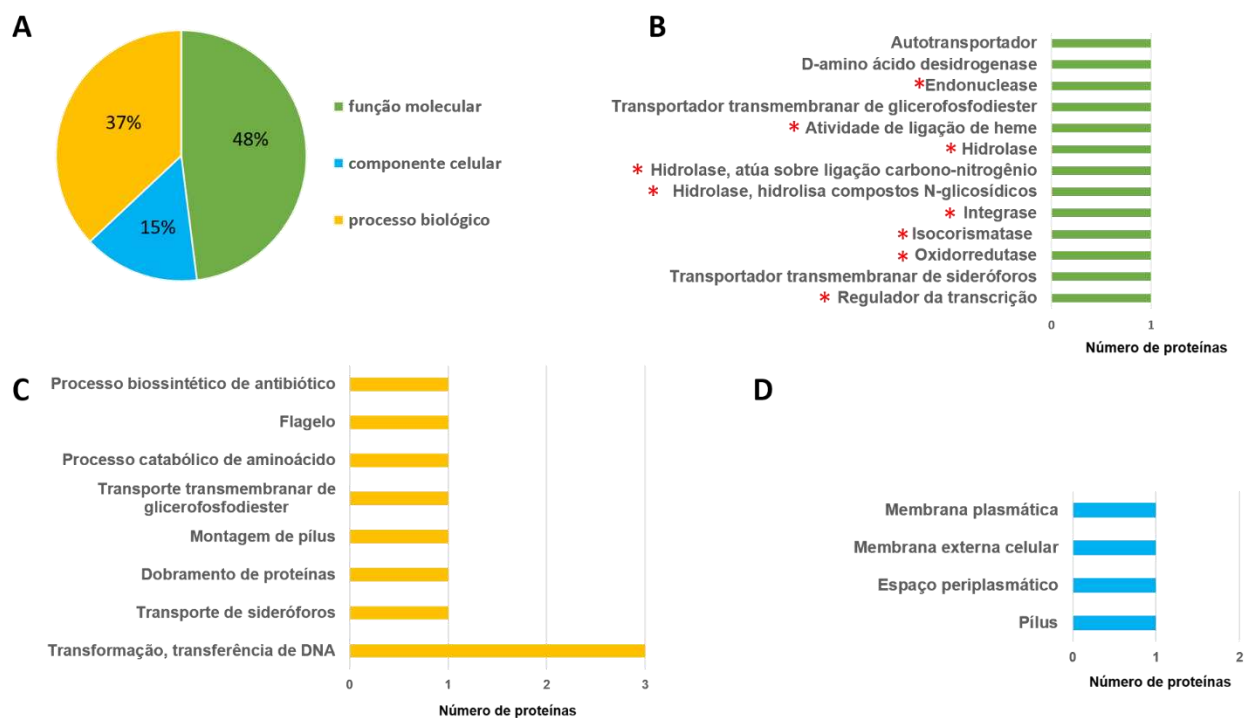


Figura 5. Identificação das proteínas do T2SS, exclusivas de *Pseudomonas syringae* pv. *garcae*, que possuem termo GO designado. (A) Distribuição das proteínas em três categorias com base em seus termos GO. Identificação das proteínas relacionadas a (B) função molecular, (C) processo biológico e (D) componente celular. As proteínas destacadas com um asterisco vermelho nos termos GO foram consideradas candidatas a efetores do tipo II.

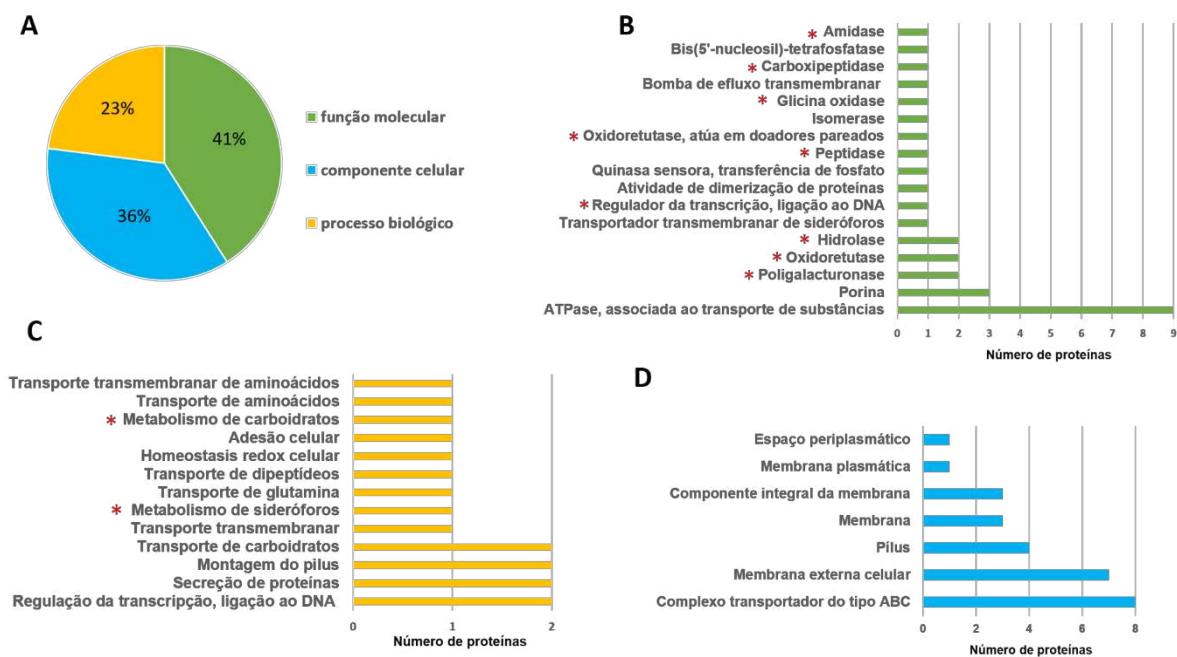


Figura 6. Identificação das proteínas do T2SS, exclusivas de *Pseudomonas syringae* pv. *tabaci*, que possuem termo GO designado. (A) Distribuição das proteínas em três categorias com base em seus termos GO. Identificação das proteínas relacionadas a (B) função molecular, (C) processo biológico e (D) componente celular. As proteínas destacadas com um asterisco vermelho nos termos GO foram consideradas candidatas a efetores do tipo II.

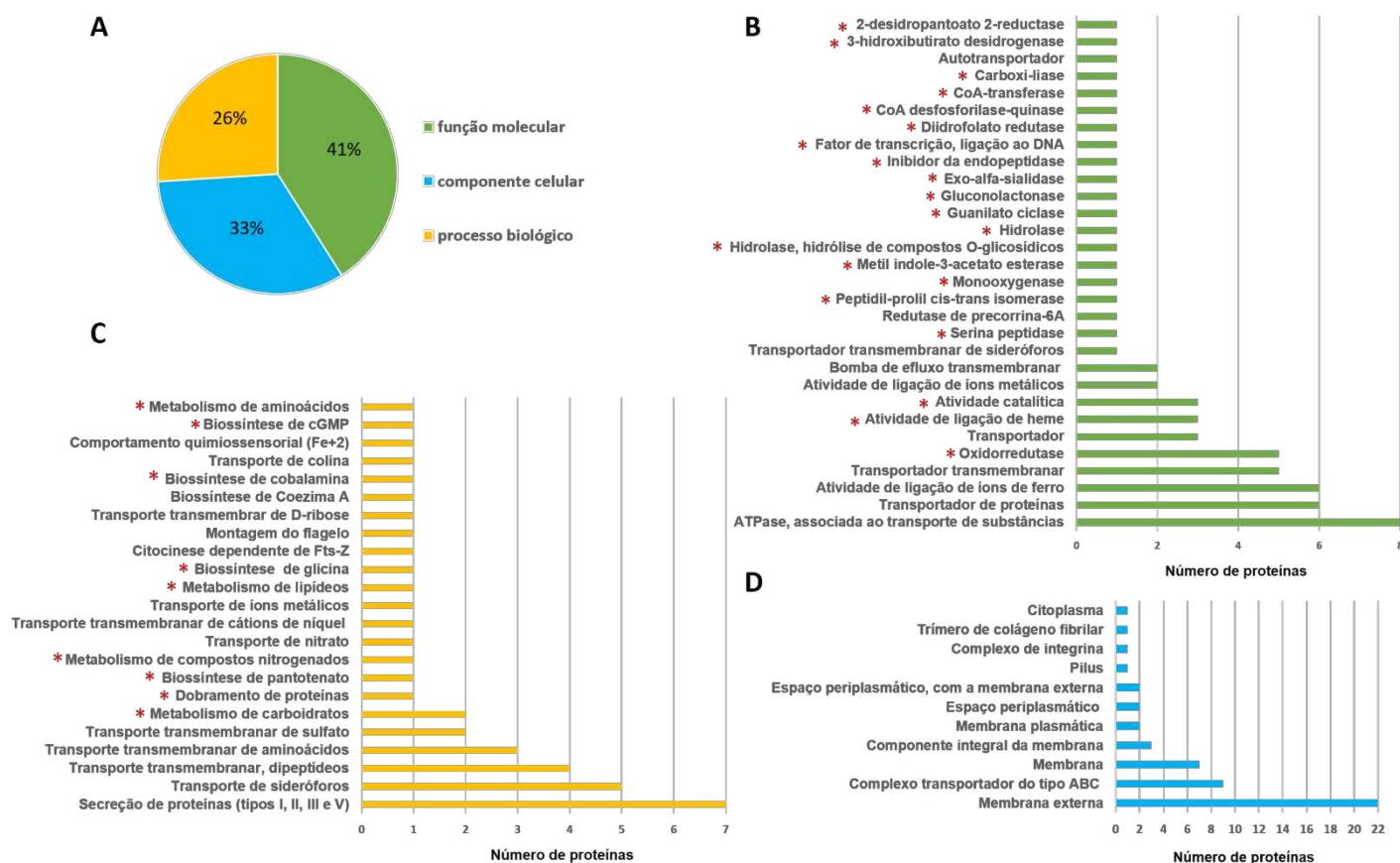


Figura 7. Identificação das proteínas do T2SS, exclusivas de *Pseudomonas cichorii*, que possuem termo GO designado. (A) Distribuição das proteínas em três categorias com base em seus termos GO. Identificação das proteínas relacionadas a (B) função molecular, (C) processo biológico e (D) componente celular. As proteínas destacadas com um asterisco vermelho nos termos GO foram consideradas candidatas a efetores do tipo II.

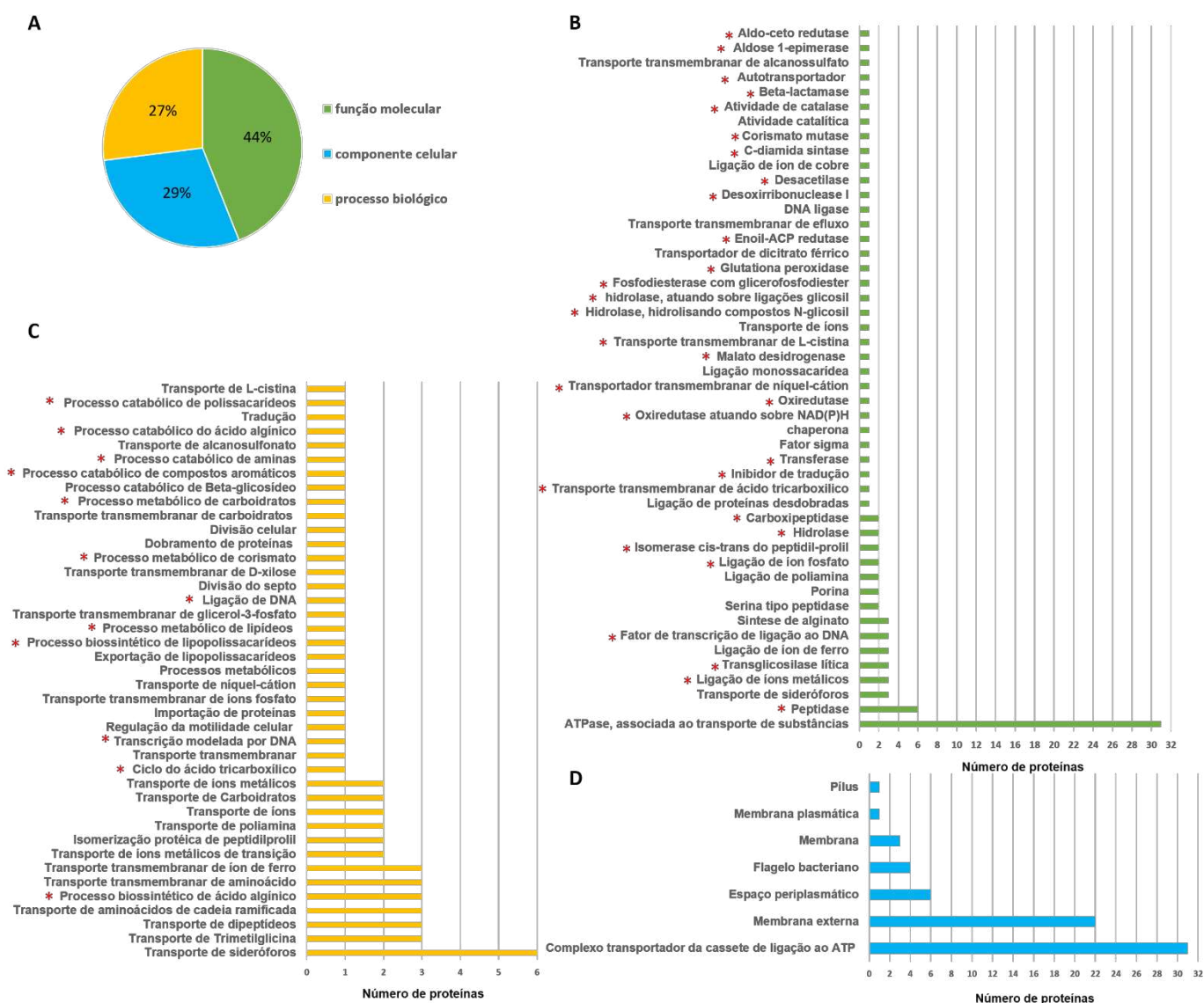


Figura 8. Identificação das proteínas do T2SS presentes em *Pseudomonas syringae* pv. *garcae*, *P. syringae* pv. *tabaci* e *Pseudomonas cichorii* que possuem termo GO designado. (A) Distribuição das proteínas em três categorias com base em seus termos GO. Identificação das proteínas relacionadas a (B) função molecular, (C) processo biológico e (D) componente celular. As proteínas destacadas com asterisco vermelho nos termos GO foram consideradas candidatas a efetores do tipo II.

TABELAS

Tabela 1. Proteínas não redundantes presentes no pansecretoma de *Pseudomonas syringae* pv. *garcae* candidatas a efetores do tipo II, identificadas mediante ontologia gênica.

ID da proteína	Comprimento	Anotação	Cobertura (%)	Ontologia gênica
WP_080267882.1	1107	Alfa/beta hidrolase	100 [369/369]	GO:0016787
WP_053931694.1	1299	Desidrogenase de D-aminoácidos	100 [433/433]	GO:0008718
WP_005889278.1	747	Desidrogenase/reductase de cadeias curtas	100 [249/249]	GO:0016491
WP_055004007.1	549	Endonuclease	100 [183/183]	GO:0004519
WP_081004334.1	303	Integrase	99.0 [101/102]	GO:0008907
WP_055004664.1	738	Isocorismatase hidrolase	100 [246/246]	GO:0008908
WP_081023870.1	603	Nucleosídeo hidrolase de inosina/uridina	100 [201/201]	GO:0016799
WP_053480947.1	2310	Proteína da família penicilina amidase	99.6 [770/773]	GO:0016811
WP_005893236.1	1050	Proteína de ligação a soluto extracelular	100 [350/350]	GO:0005615
WP_024670897.1	804	Proteína de resistencia ao telúrio	100 [268/268]	GO:0009055
WP_005889885.1	990	Regulador transcricional ArgR	100 [330/330]	GO:0140110

Tabela 2. Proteínas não redundantes presentes no pansecretoma de *Pseudomonas syringae* pv. *tabaci* candidatas a efetores do tipo II, identificadas usando ontologia gênica.

ID da proteína	Comprimento	Anotação	Cobertura (%)	Ontologia gênica
WP_117140320.1	1539	Amidase	99.6 [513/515]	GO:0004040
WP_005782793.1	981	D-3-fosfoglicerato desidrogenase	100 [327/327]	GO:0004617
WP_005777878.1	957	GDP-6-desoxi-D-lixo-4-hexulose reductase	100 [319/319]	GO:0016708
WP_117139591.1	621	Glicina oxidase	52.7 [207/393]	GO:0043799
WP_005777246.1	1302	Glucose sorbosona desidrogenase	100 [434/434]	GO:0016491
WP_117139998.1	1236	Glutamato carboxipeptidase	99.5 [412/414]	GO:0004180
WP_005777946.1	1398	Hidroxiacilglutationa hidrolase	100 [466/466]	GO:0016787
WP_117139419.1	1161	N-metiltryptofano oxidase	99.0 [387/391]	GO:0016491
WP_044299239.1	1293	Peptidase da família M20	100 [431/431]	GO:0008233
WP_117139801.1	1593	Poligalacturonase	99.8 [531/532]	GO:0004650
WP_005780218.1	1137	Potencial pectato liase	100 [379/379]	GO:0004650
WP_117139437.1	1005	Proteína não caracterizada	98.0 [335/342]	GO:0016787
WP_054071713.1	1047	Regulador transcricional da família LacI	100 [349/349]	GO:0006355

Tabela 3. Proteínas não redundantes presentes no pansecretoma de *Pseudomonas cichorii* candidatas a efetores do tipo II, identificadas usando ontologia gênica.

ID da proteína	Comprimento	Anotação	Cobertura (%)	Ontologia gênica
WP_117185508.1	1128	Amidohidrolase de N-acil-L-aminoácidos	83.9 [376/448]	GO:0016787
WP_117181951.1	585	2-ceto-3-desoxi-L-fuconato desidrogenase	79.9 [195/244]	GO:0016491
WP_117186280.1	264	Citocromo C	62.9 [88/140]	GO:0009055
WP_117166561.1	993	Decarboxilase dependente de piridoxal	71.7 [331/462]	GO:0016831
WP_117184902.1	528	Defosfo-CoA quinase	85.0 [176/207]	GO:0004140
WP_117184961.1	843	2-desidropantoato 2-reductase	91.8 [281/306]	GO:0008677
WP_117184825.1	678	Desidrogenase dependente de NAD(P) da família das álcool desidrogenases de cadeia curta	88.6 [226/255]	GO:0016491
WP_117181791.1	498	Diidrofolate reductase	97.7 [166/170]	GO:0004146
WP_117184678.1	1464	Glucana 1,4-alfa-maltohexosidase	91.4 [488/534]	GO:0003824
WP_117186579.1	927	Gluconolactonase	92.5 [309/334]	GO:0052689
WP_117185488.1	714	3-hidroxiubutirato desidrogenase	94.1 [238/253]	GO:0003858
WP_117185520.1	327	Monooxigenase, biosintese de antibióticos	84.5 [109/129]	GO:0004497
WP_117180747.1	900	Nitrilase	97.4 [300/308]	GO:0006807
WP_117187101.1	1089	Oxidoreductase	95.5 [363/380]	GO:0016491
WP_117185062.1	2016	Oxidoreductase de ligação a molibdopterina	95.7 [672/702]	GO:0016491
WP_117165294.1	477	Peptidil-prolil cis-trans isomerase	95.2 [159/167]	GO:0003755
WP_117182207.1	1818	Potencial oxidase multi-cobre	88.5 [606/685]	GO:0016491
WP_117185618.1	1407	Potencial proteína da superfamília AP	97.9 [469/479]	GO:0003824
WP_117185081.1	714	Pro-cadeia alfa do colágeno	94.4 [238/252]	GO:0005588
WP_117185590.1	4398	Proteína alfa-helicoidal extracelular	96.2 [1466/1524]	GO:0004866
WP_117185371.1	666	Proteína CbiJ	93.3 [222/238]	GO:0009236
WP_117185315.1	1455	Proteína não caracterizada	97.8 [485/496]	GO:0003824
WP_117186774.1	1122	Proteína não caracterizada	87.8 [374/426]	GO:0004553

WP_117185530.1	1650	Proteína não caracterizada	91.5 [550/601]	GO:000905
WP_117186656.1	3057	Proteína com domínio PAS S- box/diguanylate cyclase (GGDEF)	96.5 [1019/1056]	GO:0004383
WP_117184121.1	981	Proteína com repetições da caixa BNR/Asp	94.5 [327/346]	GO:0004308
WP_117167022.1	927	Regulador transcricional da família LysR	98.1 [309/315]	GO:0003700
WP_117183918.1	765	Subunidade B da 3- cetoadipato:succinil-CoA transferase	98.5 [255/259]	GO:0008410

Tabela 4. Proteínas não redundantes comuns nos pansecretomos de *Pseudomonas syringae* pv. *garcae*, *P. syringae* pv. *tabaci* e *Pseudomonas cichorii*, candidatas a efetores do tipo II, identificadas usando ontologia gênica.

ID da proteína	Comprimento	Anotação	Cobertura (%)	Ontologia gênica
WP_024671446.1	303	Ácido cobirínico ac-diamida sintetase	100 [101/101]	GO:0042242
WP_005893286.1	1035	Aldo/ceto reductase	100 [345/345]	GO:0004033
WP_024670138.1	1140	Aldose 1-epimerase	99.7 [380/381]	GO:0004034
WP_005891336.1	1134	Alginato liase	100 [378/378]	GO:0042122
WP_055004319.1	1560	Amina oxidase	83.9 [520/620]	GO:0016491
WP_024671705.1	2295	Beta-D-glucosídeo glucohidrolase	100 [765/765]	GO:1901805
WP_032618688.1	1530	Catalase	100 [510/510]	GO:0004096
WP_005896221.1	555	Corismato mutase	100 [185/185]	GO:0004106
WP_005888730.1	1158	D-alanil-D-alanina carboxipeptidase	100 [386/386]	GO:0004180
WP_055003771.1	1464	D-alanil-D-alanina carboxipeptidase/D-alanil-D-alanina-endopeptidase	100 [488/488]	GO:0004180
WP_055004369.1	966	Desoxiribonuclease I	100 [322/322]	GO:0004530
WP_005889638.1	786	Dienelactona hidrolase	100 [262/262]	GO:0016787
WP_053931803.1	1680	DNA ligase B	100 [560/560]	GO:0016874
WP_055004332.1	792	Enoil-[proteína-transportadora-de-acila] reductase [NADH]	100 [264/264]	GO:0004318
WP_005888701.1	369	Fator de iniciação da tradução Sui1	100 [123/123]	GO:0003743
WP_024670994.1	1140	Glicerofosforil diester fosfodiesterase	100 [380/380]	GO:0008889
WP_080267853.1	552	Glutationa peroxidase	100 [184/184]	GO:0004602
WP_081023889.1	2739	Glycosil hidrolase da família 3	100 [913/913]	GO:0016798
WP_024671729.1	1104	Hidrolase	100 [368/368]	GO:0016787
WP_053932245.1	1023	Hidrolase com preferencia por inosina-uridina	100 [341/341]	GO:0008152
WP_055003808.1	873	Metalo-beta-lactamase	100 [291/291]	GO:0008800
WP_005886650.1	306	Peptidase	100 [102/102]	GO:0008233
WP_024690433.1	1350	Peptidase da família M16	100 [450/450]	GO:0008233

WP_005896098.1	1326	Peptidase da família M23/M37	100 [442/442]	GO:0008233
WP_055004290.1	1419	Peptidase da família M23/M37	100 [473/473]	GO:0008233
WP_005891650.1	822	Peptidase da família M24/M37	100 [274/274]	GO:0008233
WP_005895120.1	900	Peptidase M23B	100 [300/300]	GO:0008233
WP_005894656.1	672	Peptidil-prolil cis-trans isomerase	100 [224/224]	GO:0003755
WP_024690226.1	561	Peptidil-prolil cis-trans isomerase	100 [187/187]	GO:0003755
WP_005891348.1	1176	Polissacarídeo deacetilase	100 [392/392]	GO:0019213
WP_055002600.1	1335	Proteína da família das proteases carboxi-terminais	100 [445/445]	GO:0008236
WP_024669760.1	1419	Proteína da família das transglicosilases SLT	100 [473/473]	GO:0008933
WP_024669897.1	1281	Proteína de degradação de hidrocarbonetos aromáticos	100 [427/427]	GO:0019439
WP_005888622.1	459	Proteína de ligação de cobre	100 [153/153]	GO:0005507
WP_055002601.1	780	Proteína não caracterizada	100 [260/260]	GO:0003824
WP_005891678.1	561	Proteína não caracterizada	100 [187/187]	GO:0016740
WP_055004270.1	888	Proteína reguladora de ligação a substrato LysR	100 [296/296]	GO:0003700
WP_024671030.1	1926	Proteína SLT (transglicosilase solúvel)	100 [642/642]	GO:0008933
WP_080267841.1	1647	Provável malato:quinona oxidoreductase	100 [549/549]	GO:0008924
WP_005893897.1	888	Regulador transcricional da família LysR	100 [296/296]	GO:0003700
WP_054989297.1	951	Regulador transcricional YelE da família LysR	100 [317/317]	GO:0003700
WP_055002215.1	984	Renalase	100 [328/328]	GO:0050660
WP_002555879.1	165	Rubredoxina	100 [55/55]	GO:0009055
WP_005888726.1	1005	Transglicosilase lítica de mureina B	100 [335/335]	GO:0008933

* As IDs das proteínas correspondem a anotação do genoma de *Pseudomonas syringae* pv. *garcae*