

## SEQUENCIAMENTO DO GENOMA COMPLETO DE UMA NOVA CEPA DE *XANTHOMONAS FUSCANS* SUBSP. *FUSCANS*, OBTIDA DE FEIJÕES INFECTADOS NO RS (Brasil)

FREDERICO SCHMITT KREMER<sup>1</sup>; AMANDA MUNARI GUIMARÃES<sup>2</sup>; ISMAIL  
TEODORO DE SOUZA<sup>3</sup>; ANDREA BITTENCOURT MOURA<sup>4</sup>; LUCIANO DA  
SILVA PINTO<sup>5</sup>

<sup>1</sup> BioPro-Lab, CDTec, UFPel - [fred.s.kremer@gmail.com](mailto:fred.s.kremer@gmail.com)

<sup>2</sup> BioPro-Lab, CDTec, UFPel - [mandimunari@gmail.com](mailto:mandimunari@gmail.com)

<sup>3</sup> FAEM, UFPel - [agrojunior1@yahoo.com.br](mailto:agrojunior1@yahoo.com.br)

<sup>4</sup> FAEM, UFPel - [andreabittencourtmoura@hotmail.com](mailto:andreabittencourtmoura@hotmail.com)

<sup>5</sup> BioPro-Lab, CDTec, UFPel - [ls\\_pinto@hotmail.com](mailto:ls_pinto@hotmail.com)

### 1. INTRODUÇÃO

O gênero *Xanthomonas* abrange mais de 30 espécies de bactérias gram-negativas, sendo muitas destas patógenos de plantas de elevada relevância comercial, como arroz, feijão, soja e *citrus*. As espécies deste gênero são usualmente sub-classificadas em patovares e subespécies de acordo com o hospedeiro que é infectado, doença que é acarretada, ou com base em características fenotípicas específicas (RYAN *et al.*, 2011).

A espécie *X. fuscans* subespécie *fuscans* (*Xff*) é um dos principais patógenos de feijão, agente etiológico da mancha bacteriana foliar (do inglês *common blight of bean*, ou CBB) (AKHAVAN *et al.*, 2013). Esta doença pode levar a perdas de produção de até 40%, e tem um impacto particularmente alto nas américas e África, onde o feijão representa uma das principais fontes de proteínas para alimentação humana (GRAHAM; VANCE, 2003; MORRIS; MONIER, 2003). Deste modo, o devido controle desta bactéria nestas regiões vem se tornando uma problemática cada vez mais relevante.

A patogênese da *Xff* se dá através de um processo progressivo de invasão dos tecidos da planta, que começa pela formação de biofilme na sua superfície (sendo este majoritariamente composto por polissacarídeos como a xantana), passando pela invasão dos estômatos e mesofilo, e que por fim a chega e colonização dos tecidos vasculares (JACQUES *et al.*, 2005; RUH *et al.*, 2017). A capacidade de invadir, infectar e evadir as defesas da planta é garantida por diversos genes de virulência, desde enzimas de degradação, polissacarídeos que forma biofilmes, conjuntos de proteínas efetoras, proteínas adesinas, dentre outros. (DARRASSE *et al.*, 2013) Entretanto, ainda há uma grande limitação de dados genômicos sobre esta linhagem, sobretudo de isolados locais, o que dificulta o entendimento de sua diversidade molecular e patogênica.

Deste modo, no presente trabalho é descrito o sequenciamento do primeiro genoma brasileiro de *X. fuscans* subsp. *fuscans*, tendo sido a cepa, denominada Xap49, obtida de feijões (*Phaseolus vulgaris* L) infectados e sintomáticos no estado do Rio Grande do Sul (Brasil).

## 2. METODOLOGIA

o sequenciamento do genoma completo da cepa Xap49 foi realizado com a plataforma Illumina MiSeq usando-se uma biblioteca paired-end de 500 bp de inserto, sendo este processo executado pela empresa Neopropecta (<https://neopropecta.com/>). A remoção de leituras de baixa qualidade e adaptadores foi realizada com as ferramentas FASTXtoolkit (<http://hannonlab.cshl.edu/>) e Trimmomatic (<http://www.usadellab.org/>).

O processo de montagem *de novo* foi realizado com as ferramentas Velvet (ZERBINO; BIRNEY, 2008), SGA, Ray (BOISVERT; LAVIOLETTE; CORBEIL, 2010), SPAdes (BANKEVICH *et al.*, 2012) e CLC Genome Workbench, sendo os resultados destes programas integrados pelo CISA (LIN; LIAO, 2013) para a geração de uma montagem consenso. A montagem gerada foi alinhada contra o banco de dados do GenBank para a identificação de *contigs* que pertencem a plasmídeos, sendo estas montadas separadamente. As *contigs* identificadas como pertencentes ao cromossomo foram ordenadas com uso da ferramenta CAR com base no genoma da cepa 4834-R (GenBank: NC\_022541.1), e um processo de *gap-closing* foi realizado com a ferramenta IMAGE (TSAI; OTTO; BERRIMAN, 2010). A anotação do genoma foi realizada com a ferramenta Genix (KREMER *et al.*, 2016).

De modo a se analisar as diferenças entre as cepas Xap49 e 4834-R, realizou-se uma análise de variantes a partir dos dados de sequenciamento filtrados após o pré-processamento, sendo usada uma pipeline baseada nos programas Samtools, Bcftools e VCFUtils (LI *et al.*, 2009) para identificação das variantes e SnpEff (REUMERS, 2004) para identificação do efeito de cada variante, sendo os genes preditos como mutados anotados com o banco de dados COG (TATUSOV *et al.*, 2000). Uma análise de genômica estrutural também foi realizada, sendo usada a ferramenta Mauve para identificação de inversões e translocações estruturais no genoma da cepa Xap49.

## 3. RESULTADOS E DISCUSSÃO

Os resultados da montagem e anotação estão sumarizados, respectivamente, nas Tabelas 1 e 2. Além do genoma cromossomal, através da estratégia de montagem empregada foi possível se identificar a presença de 3 plasmídeos, sendo dois destes homólogos dos plasmídeos pIA e pIC que haviam sido previamente descritos na cepa 4834-R, e um terceiro, denominado pIX, que é similar ao plasmídeo pAX22 de *A. xylosoxidans* subsp. *denitrificans*. Neste plasmídeo foi identificado um grande número de genes associados ao processo de conjugação bacteriana, além de muitas proteínas hipotéticas e genes que codificam para enzimas de restrição similares à *notI*. Já no cromossomo foi possível identificar diversos genes associados ao processo de patogênese, incluindo *operons* associados à síntese da goma xantana (associada ao biofilme), como o *operon gumBCFEFGHIJK*, transportadores do tipo ABC e proteínas efetoras da família TALE, que atuam na regulação da expressão gênica do hospedeiro.

**Tabela 1.** Resultado geral da montagem final do genoma da cepa Xap49

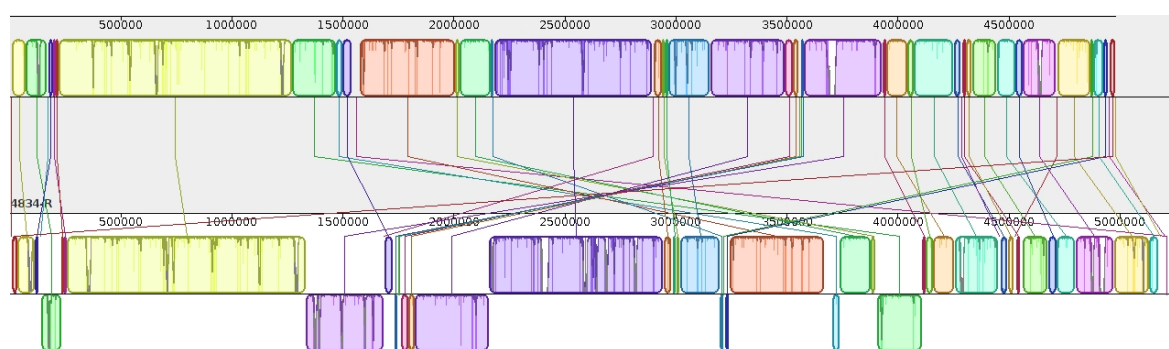
<i>Replicon</i>	Tamanho (bp)	N. Contigs	N50 ( <i>contigs</i> )
Cromossomo	5182257	70	112211
pIA	41822	4	26697
pIC	34064	5	5176
pIX	18734	2	15216

**Tabela 2.** Resultado geral da anotação gerada usando a ferramenta Genix para o genoma da cepa Xap49

<i>Replicon</i>	CDS	tRNA	rRNAs	tmRNAs	Outros*
Cromossomo	4218	61	7	1	49
pIA	49	0	0	0	1
pIC	37	0	0	0	0
pIX	20	0	0	0	0

\*= Outras famílias de elementos não codificantes

Através da análise de variantes foi possível se identificar 299 SNPs entre as duas cepas, e 29 INDELS, indicando que existe uma grande similaridade entre elas. No que diz respeito ao contexto destas mutações, a grande maioria foi predita em regiões intergênicas, *upstream/downstream* ou como tendo efeito silencioso, sendo as mutações *missense* observadas apenas em alguns processos biológicos, como transdução de sinal, mobiloma e metabolismo de íons. Entretanto, observando os resultados da análise estrutural comparativa (Figura 1) foi possível identificar a presença de rearranjos estruturais, indicando a existência de uma certa plasticidade genômica neste grupo de bactérias.



Xap49

**Figura 1.** Resultado da análise de genômica estrutural comparativa gerada pelo programa Mauve para os genomas das cepas Xap49 e 4834-R.

## 4. CONCLUSÕES

A disponibilidade de novos dados genômicos de cepas locais de *X. fuscans* subsp. *fuscans* pode prover dados relevantes para um melhor entendimento das características genéticas destas bactérias, bem como ser usados em análises evolutivas, de patogenicidade, e ser aplicada no desenvolvimento de novas abordagens para controle e rastreamento destas. Os dados do genoma da cepa

Xap49 podem ser acessados no GenBank através dos códigos de acesso CP023294, CP023295, CP023296 e CP023297.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

- AKHAVAN, Alireza *et al.* Bean common bacterial blight: pathogen epiphytic life and effect of irrigation practices. **SpringerPlus** v. 2, n. 1, p. 41, dez. 2013.
- BANKEVICH, Anton *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. **Journal of computational biology: a journal of computational molecular cell biology** v. 19, n. 5, p. 455–77, maio 2012.
- BOISVERT, Sébastien; LAVIOLETTE, François; CORBEIL, Jacques. Ray: Simultaneous Assembly of Reads from a Mix of High-Throughput Sequencing Technologies. , 10 nov. 2010. 2014.10.1089/cmb.2009.0238.
- DARRASSE, Armelle *et al.* Genome sequence of *Xanthomonas fuscans* subsp. *fuscans* strain 4834-R reveals that flagellar motility is not a general feature of xanthomonads. **BMC genomics** v. 14, p. 761, 6 nov. 2013.
- GRAHAM, P. H.; VANCE, Carroll P. Legumes: Importance and Constraints to Greater Use. **PLANT PHYSIOLOGY** v. 131, n. 3, p. 872–877, 1 mar. 2003.
- JACQUES, M-A *et al.* *Xanthomonas axonopodis* pv. *phaseoli* var. *fuscans* is aggregated in stable biofilm population sizes in the phyllosphere of field-grown beans. **Applied and environmental microbiology** v. 71, n. 4, p. 2008–15, abr. 2005.
- KREMER, Frederico Schmitt *et al.* Genix: A New Online Automated Pipeline for Bacterial Genome Annotation. **FEMS Microbiology Letters** v. 11, n. 16, 2016.
- LI, Heng *et al.* The Sequence Alignment/Map format and SAMtools. **Bioinformatics (Oxford, England)** v. 25, n. 16, p. 2078–9, 15 ago. 2009.
- LIN, Shin-Hung; LIAO, Yu-Chieh. CISA: contig integrator for sequence assembly of bacterial genomes. **PloS one** v. 8, n. 3, p. e60843, jan. 2013.
- MORRIS, Cindy E.; MONIER, Jean-Michel. THE ECOLOGICAL SIGNIFICANCE OF BIOFILM FORMATION BY PLANT-ASSOCIATED BACTERIA. **Annual Review of Phytopathology** v. 41, n. 1, p. 429–453, set. 2003.
- REUMERS, J. SNPeff: a database mapping molecular phenotypic effects of human non-synonymous coding SNPs. **Nucleic Acids Research** v. 33, n. Database issue, p. D527–D532, 17 dez. 2004.
- RUH, Mylène *et al.* *Xanthomonas* adaptation to common bean is associated with horizontal transfers of genes encoding TAL effectors. **BMC Genomics** v. 18, n. 1, p. 670, 30 dez. 2017.
- RYAN, Robert P. *et al.* Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions. **Nature Reviews Microbiology** v. 9, n. 5, p. 344–355, 11 maio 2011.
- TATUSOV, R L *et al.* The COG database: a tool for genome-scale analysis of protein functions and evolution. **Nucleic acids research** v. 28, n. 1, p. 33–6, 1 jan. 2000.
- TSAI, Isheng J; OTTO, Thomas D; BERRIMAN, Matthew. Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. **Genome biology** v. 11, n. 4, p. R41, jan. 2010.
- ZERBINO, Daniel R; BIRNEY, Ewan. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. **Genome research** v. 18, n. 5, p. 821–9, maio 2008.