



UNIVERSIDADE
ESTADUAL DE LONDRINA

CÉSAR TOSHIO FACIMOTO

**ANÁLISE DO GENOMA COMPLETO DA CEPA
FRANCISELLA NOATUNENSIS SUBSP. *ORIENTALIS* F1 E
PREDIÇÃO DE CANDIDATOS VACINAIS CONTRA
FRANCISELOSE EM PEIXES DE ÁGUAS QUENTES E
FRIAS**

CÉSAR TOSHIO FACIMOTO

**ANÁLISE DO GENOMA COMPLETO DA CEPA
FRANCISELLA NOATUNENSIS SUBSP. *ORIENTALIS* F1 E
PREDIÇÃO DE CANDIDATOS VACINAIS CONTRA
FRANCISELOSE EM PEIXES DE ÁGUAS QUENTES E
FRIAS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência Animal da Universidade Estadual de Londrina como requisito parcial para a obtenção do título de Mestre.

Orientador: Prof. Dr. Ulisses de Pádua Pereira

Londrina
2019

CÉSAR TOSHIO FACIMOTO

**ANÁLISE DO GENOMA COMPLETO DA CEPA *FRANCISELLA*
NOATUNENSIS SUBSP. *ORIENTALIS* F1 E PREDIÇÃO DE
CANDIDATOS VACINAIS CONTRA FRANCISELOSE EM PEIXES DE
ÁGUAS QUENTES E FRIAS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência Animal da Universidade Estadual de Londrina como requisito parcial para a obtenção do título de Mestre.

BANCA EXAMINADORA

Orientador: Prof. Dr. Ulisses de Pádua Pereira
Universidade Estadual de Londrina – UEL

Prof. Dr. Vasco Ariston de Carvalho Azevedo
Universidade Federal de Minas Gerais – UFMG

Prof. Dr. Siomar de Castro Soares
Universidade Federal do Triângulo Mineiro – UFTM

Prof. Dr. Luís Carlos Guimarães
Universidade Federal do Pará – UFPA

Londrina, 19 de fevereiro de 2019.

AGRADECIMENTOS

Aos amigos pós graduandos Roberta, Felipe, Ed, Raffaella, Suelen e Leonardo que sempre colaboraram e me motivaram no meu amadurecimento pessoal e profissional.

Ao meu orientador Ulisses, por todo o entusiasmo e trabalho duro desempenhado em benefício da ciência, por ter oferecido todo o suporte necessário para o meu crescimento acadêmico e por ter me introduzido e trabalhado ao meu lado no campo da bioinformática.

À equipe do Laboratório de Bacteriologia em Peixes, que proporcionou um ambiente agradável e produtivo para o trabalho.

Aos educadores que participaram de todo o meu processo de formação científica, em especial aos professores Siomar, Nelson Giovanna e Admilton que participaram no desenvolvimento deste trabalho com sugestões e parcerias.

Aos membros da banca examinadora, Prof. Vasco Ariston de Carvalho Azevedo, e Prof. Luís Carlos Guimarães que gentilmente aceitaram participar e colaborar com esta dissertação.

À minha parceira Juliana que esteve ao meu lado nos meus altos e baixos durante a minha jornada pelo mestrado e aos amigos Lucas, John, Felipe e Vinicius pelo companheirismo.

Aos governos que estimularam a educação e a ciência brasileira, por todas as medidas de inclusão social que me permitiram buscar uma educação de qualidade ao longo da minha vida.

Meus sinceros agradecimentos à todos.

FACIMOTO, César Toshio. **Análise do Genoma Completo da Cepa *Francisella noatunensis* subsp. *orientalis* F1 e Predição de Candidatos Vacinais Contra Franciselose em peixes de Águas Quentes e Frias**. 2019. 54 f. Dissertação (Mestrado em Ciência Animal) – Universidade Estadual de Londrina, Londrina, 2019.

RESUMO

A espécie *Francisella noatunensis* subsp. *orientalis*, agente causador da franciselose em tilápias, tem recebido destaque na tilapicultura mundial devido ao seu caráter emergente e por causar perdas na produção, principalmente em países de climas tropicais. Por outro lado, em regiões predominantemente frias, peixes como bacalhau do atlântico, salmão e truta tem sido acometidos pela subespécie *Francisella noatunensis* subsp. *noatunensis*. Este estudo tem por objetivo sequenciar, montar e depositar em bancos de dados públicos o genoma completo de uma cepa de *F. noatunensis* subespécie *orientalis* e predizer *in silico* candidatos proteicos imunogênicos conservados entre as subespécies para o desenvolvimento de uma vacina eficaz contra a franciselose em peixes de águas quentes e frias. A cepa F1 foi isolada de um surto em tilapicultura ocorrido no noroeste do estado de São Paulo em 2015. O DNA foi submetido à sequenciamento de próxima geração, as *reads* resultantes do sequenciamento foram filtradas, montadas em *contigs* e alinhadas com uma cepa referência para criação de um *scaffold*. Os *gaps* resultantes foram preenchidos através da verificação e extensão das bordas dos intervalos em mapeamentos sucessivos com as *reads*. A filogenia foi analisada utilizando a ferramenta Gegenees e a identificação de candidatos vacinais foi realizada utilizando as ferramentas MEDpipe e Vaxign. As proteínas resultantes desta última análise foram checadas quanto à homologia no genoma completo de *Oreochromis niloticus* e *F. noatunensis* subsp. *noatunensis*. A função das proteínas identificadas foram preditas utilizando a ferramenta Interproscan. Através da comparação com 12 genomas completos de *Francisella* spp. disponíveis no GenBank, a análise de filogenia apresentou uma grande similaridade da cepa F1 com as cepas de *F. noatunensis* subsp. *orientalis* isoladas no Brasil, e com uma cepa isolada na Indonésia. Além disso, os grupos de cepas que acometem humanos apresentaram-se distantes das cepas que acometem peixes. A vacinologia reversa e análise das proteínas selecionadas revelaram um total de 5 proteínas (WP_014715657.1, WP_014714512.1, WP_012280666.1, WP_014714564.1 e WP_014714557.1) potencialmente expostas ou secretadas, conservadas entre as duas subespécies de *F. noatunensis*, imunogênicas e sem homologia com proteínas do hospedeiro. As proteínas selecionadas que apresentaram resultado na predição estão associadas à mecanismos de sobrevivência no interior de macrófagos e secreção de hemolisina. Por fim, os resultados obtidos por este estudo servem de base para o desenvolvimento de testes *in vivo* que validem a eficácia dos alvos vacinais e posteriormente sua aplicação no campo.

Palavras-chave: Tilapicultura. Montagem de genoma. Vacina. Vacinologia reversa.

FACIMOTO, César Toshio. **Complete Genome Sequence of *Francisella noatunensis* subsp. *orientalis* F1 Strain Isolated from a Brazilian Tilapia Farm and Prediction of Universal Vaccine Candidates Against Warm and Cold Water Fish Francisellosis**. 2019. 54 p. Dissertation (Master's Degree in Animal Science) – Universidade Estadual de Londrina, Londrina, 2019.

ABSTRACT

Francisella noatunensis subsp. *orientalis* is a pathogen responsible for francisellosis in tilapia. Due to its emergent nature and high mortality profile, this species is a big concern to world aquaculture. In tropical countries, *F. noatunensis* subsp. *orientalis* is responsible for francisellosis in tilapia in winter. On the other hand, in cold countries, *Francisella noatunensis* subsp. *noatunensis* causes disease in atlantic cod and trout. This study aims to sequence, assembly and submit to public databases the complete genome of a *F. noatunensis* subspecies *orientalis* and to predict *in silico* immunogenic conserved proteins among the subspecies that can induce an efficient vaccine against francisellosis in warm and cold waters fish. The F1 strain was isolated from a tilapia farm francisellosis outbreak in Northwest of Sao Paulo State in 2015. The DNA sample was submitted to Next-Generation Sequencing, and resulting reads were filtered, assembled to contigs and aligned to a reference strain to create a scaffold. Remaining gaps were filled through extensions of the gap borders in successive mappings with raw reads. Phylogeny was performed using Gegenees tool and the identification of vaccine targets was ran through MEDpipe and Vaxign tools. Resulting proteins were checked for homology in the complete genomes of *Oreochromis niloticus* and *F. noatunensis* subsp. *orientalis*. Proteins function were predicted using Interproscan database. Twelve complete genomes of *Francisella* spp. available at GenBank were compared to the F1 strain in the phylogeny step. The F1 strain displayed a high similarity with other *F. noatunensis* subsp. *orientalis* isolated from Brazil and with a strain isolated in Indonesia. Beyond, clusters of strains that infect humans were distant from clusters of strains that causes disease in fish. Reverse vaccinology analysis revealed 5 proteins (WP_014715657.1, WP_014714512.1, WP_012280666.1, WP_014714564.1 e WP_014714557.1) potentially surface exposed, conserved in the *F. noatunensis* species, high immunogenic and with no host homology. The proteins with predicted functions were associated with mechanisms of survival within macrophages and hemolysin secretion. Results obtained in this study allow *in vivo* research to validate the efficacy of the vaccine targets found, followed by the development of a subunit vaccine applicable in tilapia farms.

Keywords: Aquaculture. Genome assembly. Vaccine. Reverse vaccinology.

LISTA DE ILUSTRAÇÕES

REVISÃO DE LITERATURA

- Figura 1** – Os maiores produtores de tilápia, em T (2016). 14
- Figura 2** – Fotomicrografia de cepas de *F. noatunensis* subsp. *orientalis* isoladas no Brasil..... 15
- Figura 3** – Rim cranial e baço de tilápia com apresentação de granulomas de franciselose..... 16
- Figura 4** – Ilhas de patogenicidade do gênero *Francisella* (FPI). Representação das ilhas genômicas de *F. novicida*, loci, ORFs e nomenclatura do genes 19
- Figura 5** – Arvore filogenética utilizando o genoma completo de cepas de *Francisella* spp. 23

ARTIGO PARA PUBLICAÇÃO

“COMPLETE GENOME SEQUENCE OF *FRANCISELLA NOATUNENSIS* SUBSP. *ORIENTALIS* F1 STRAIN ISOLATED FROM A BRAZILIAN TILAPIA FARM AND PREDICTION OF UNIVERSAL VACCINE CANDIDATES AGAINST WARM AND COLD WATER FISH FRANCISELLOSIS”.

- Figure 1** – Photomicrograph of the *F. noatunensis* subsp. *orientalis* F1 strain..... 50
- Figure 2** – Heatmap based on whole genome of selected *Francisella* spp. Heatmap of similarity among non-core regions of *Francisella* spp. genomes using Gegenees all in all fragmented comparison with high accuracy 51
- Figure 3** – Ocular globe of Nile tilapia (*Oreochromis niloticus*) infected by *F. noatunensis* subsp. *orientalis* strain F1. A) Presence of marked nodular inflammatory reaction between choroid (arrow) and sclera (*). Hematoxylin and Eosin, 4x. B) Accentuated accumulation of polymorphonuclear and mononuclear inflammatory cells in well vascularized stroma (arrows) with granulomatous formation (*). Hematoxylin and Eosin, 20x. C) Granulomatous area composed of vacuolated macrophages

and epithelioid cells. Hematoxylin and Eosin, 40x. D) Granulomatous portion evidencing the presence of intracytoplasmic cocoid bacteria (arrows). Hematoxylin and Eosin, 100x..... 52

Figure 4 – Phylogenetic tree based on whole genome of *Francisella* spp. strains using EqualAngleRoot and UPGMA methods. A) Phylogenetic tree comparing different species in the *Francisella* genus evidencing the distance of strains. B) A zoomed view of the Fno cluster..... 53

LISTA DE TABELAS

REVISÃO DE LITERATURA

Tabela 1 –	Principais Características de genomas sequenciados de <i>Francisella</i> spp	22
-------------------	--	----

ARTIGO PARA PUBLICAÇÃO

“COMPLETE GENOME SEQUENCE OF *FRANCISELLA NOATUNENSIS* SUBSP. *ORIENTALIS* F1 STRAIN ISOLATED FROM A BRAZILIAN TILAPIA FARM AND PREDICTION OF UNIVERSAL VACCINE CANDIDATES AGAINST WARM AND COLD WATER FISH FRANCISELLOSIS”.

Table 1 –	Classification and general features of <i>F. noatunensis</i> subsp. <i>orientalis</i> strain F1	44
Table 2 –	Genome sequencing project information.....	45
Table 3 –	Genomes used in phylogeny analysis	46
Table 4 –	Genome statistics	47
Table 5 –	Number of genes associated with general COG functional categories.....	48
Table 6 –	Vaccine candidates.....	49

LISTA DE ABREVIATURAS E SIGLAS

anm	Anhydro-N-acetylmuramic acid
BLAST	Basic Local Alignment Search Tool
BLASTP	Protein BLAST
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CDS	Coding Sequence
CHAH	Cysteine Heart Agar with Hemoglobin
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
COG	Cluster of Orthologous Genes
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CYT	Cytoplasmatic
DDBJ	Data Bank of Japan
DNA	Deoxyribonucleic Acid
dot	defect in organelle trafficking
DUF	Domain of Unknown Function
EMBL	European Molecular Biology Laboratory
FPI	<i>Francisella</i> Pathogenicity Island
HE	Hematoxylin and Eosin
HlyD	Haemolysin secretion protein D
IDA	Inferred from Direct Assay
IFN- γ	Interferon-gama
igl	Intracellular growth locus
IL-1b	Interleucina-1-beta
LABBEP	Laboratório de Bacteriologia em Peixes
LGCM	Laboratório de Genética Celular e Molecular
MCE	Mammalian Cell Entry
MED	Maturity Epitope Density
MEM	Membrane
MHC	Major Histocompatibility Complex
MIGS	Minimum Information on Genome Sequence
NAS	Non-traceable Author Statement
NCBI	National Center for Biotechnology Information
NGS	Next-Generation Sequencing
pdp	Pathogenicity determinant protein
TAS	Traceable Author Statement

SUMÁRIO

1	INTRODUÇÃO	12
2	REVISÃO DA LITERATURA	14
2.1	TILAPICULTURA	14
2.2	GÊNERO <i>FRANCISELLA</i> SPP.	15
2.3	FRANCISELOSE	15
2.4	CONTROLE E PREVENÇÃO	16
2.5	ILHAS DE PATOGENICIDADE.....	17
2.6	SAÚDE PÚBLICA	20
2.7	GENÔMICA COMPARATIVA.....	20
2.8	FILOGENÔMICA.....	22
2.9	VACINOLOGIA REVERSA	24
	REFERÊNCIAS	25
3	HIPÓTESE	32
4	OBJETIVOS	33
4.1	OBJETIVO GERAL.....	33
4.2	OBJETIVOS ESPECÍFICOS	33
5	ARTIGO ACEITO PARA PUBLICAÇÃO NO PERIÓDICO	
	STANDARDS IN GENOMICS SCIENCE	34
	ABSTRACT	35
5.1	INTRODUCTION	36
5.2	ORGANISM INFORMATION.....	36
5.2.1	Classification and Features.....	36
5.3	GENOME SEQUENCING INFORMATION	37
5.3.1	Genome Project History.....	37
5.3.2	Growth Conditions and Genomic DNA Preparation.....	37
5.3.3	Genome Sequencing and Assembly	37
5.3.4	Genome Annotation.....	38

5.4	GENOME PROPERTIES.....	38
5.4.1	Extended Insights.....	38
5.5	CONCLUSIONS.....	39
	COMPETING INTERESTS	40
	FUNDING.....	40
	AUTHOR'S CONTRIBUTIONS	40
	REFERENCES	41
6	CONCLUSÃO.....	54

1 INTRODUÇÃO

O Brasil é o 4º maior produtor de tilápia no mundo atualmente, ficando atrás apenas de China, Indonésia e Egito. Além disso, perspectivas sobre o crescimento da aquicultura mundial apontam que América Latina, Caribe, Índia e Sudeste Asiático terão um crescimento superior à 100% no período de 2010 à 2030. (KOBAYASHI et al., 2015; PEIXEIR, 2018). Esse crescimento acelerado estabelece novos desafios à biossegurança na produção de tilápias. Por sua vez, o controle de doenças na piscicultura é um dos maiores entraves na tilapicultura atual devido à perdas econômicas ocasionadas principalmente por estreplococose e franciselose.

A franciselose é uma doença granulomatosa de grande importância na aquicultura mundial devido à alta mortalidade causada no plantel e à ausência de medidas de controle dessa enfermidade. Os surtos associados à este patógeno se manifestam devido à diminuição brusca da temperatura do ambiente, induzindo o hospedeiro à estresse térmico e favorecendo o crescimento da bactéria (LEAL et al., 2014). A transmissão se dá pelo contato direto entre animais, água ou ração contaminada e a bactéria pode persistir por até 2 dias no ambiente na ausência do hospedeiro (SOTO; ABRAMS; REVAN, 2012). O tratamento dessa enfermidade ainda é limitado quanto à disponibilidade e a eficácia de fármacos. Além disso, não há vacina disponível no mercado para a profilaxia da franciselose, embora alguns estudos tenham obtido moderado êxito no seu desenvolvimento (SOTO et al., 2011d; SOTO; FERNANDEZ; HAWKE, 2009).

O gênero *Francisella* spp. possui grande importância na saúde humana e animal, uma vez que as estirpes abrigadas por esse gênero acometem humanos e estão presentes no ambiente. Devido à baixa dose infectante da *Francisella tularensis*, esse patógeno tem sido classificado como arma biológica devido à capacidade de ser utilizado na forma de aerossol (PECHOUS; MCCARTHY; ZAHRT, 2009; OYSTON; SJOSTEDT; TITBALL, 2004). A ocorrência em humanos causa a tularemia, e a infecção se dá por picada de artrópode, inalação, ingestão de tecidos de animais contaminados ou contato com água ou solo contaminado (KILIC et al., 2015).

Os estudos direcionados à espécies que acometem humanos estão em constante desenvolvimento, em consonância com isso, cada vez mais estudos utilizando abordagens *in silico* tem sido produzidos nos últimos anos. Os resultados

in silico colaboram grandemente na investigação da franciselose, no entanto, as particularidades das espécies associadas à esta doença em peixes ainda precisam ser investigados de maneira mais aprofundada.

Desde o estabelecimento de técnicas de sequenciamento de próxima geração (NGS), uma grande quantidade de genomas completos tem sido disponibilizada em bancos de dados públicos, tornando os estudos genômicos mais acessíveis e permitindo que anos de trabalho na bancada fossem reduzidos à meses ou dias no computador. Nesse contexto, mais ferramentas *in silico* tem sido desenvolvidas (WORLD HEALTH ORGANIZATION, 2002). Através do uso de dados de sequenciamento genômico e ferramentas de análise bioinformática, vacinas de subunidades já foram desenvolvidas e tiveram bons resultados experimentais para outras espécies bacterianas utilizando *pipelines* de vacinologia reversa (RAPPUOLI, 2001; GIULIANI et al., 2006; SETTE; RAPPUOLI, 2010).

Embora exista uma boa cobertura de estudos genômicos sobre *F. tularensis*, ainda há escassez de resultados de genômica direcionados à *F. noatunensis*, principalmente na busca de alvos vacinais. Portanto, o objetivo deste trabalho é sequenciar, montar e depositar em bancos de dados o genoma completo da cepa *F. noatunensis* subsp. *orientalis* F1 e predizer *in silico* candidatos proteicos imunogênicos conservados entre as subespécies para o desenvolvimento de vacina eficaz contra a franciselose em peixes de águas quentes e frias.

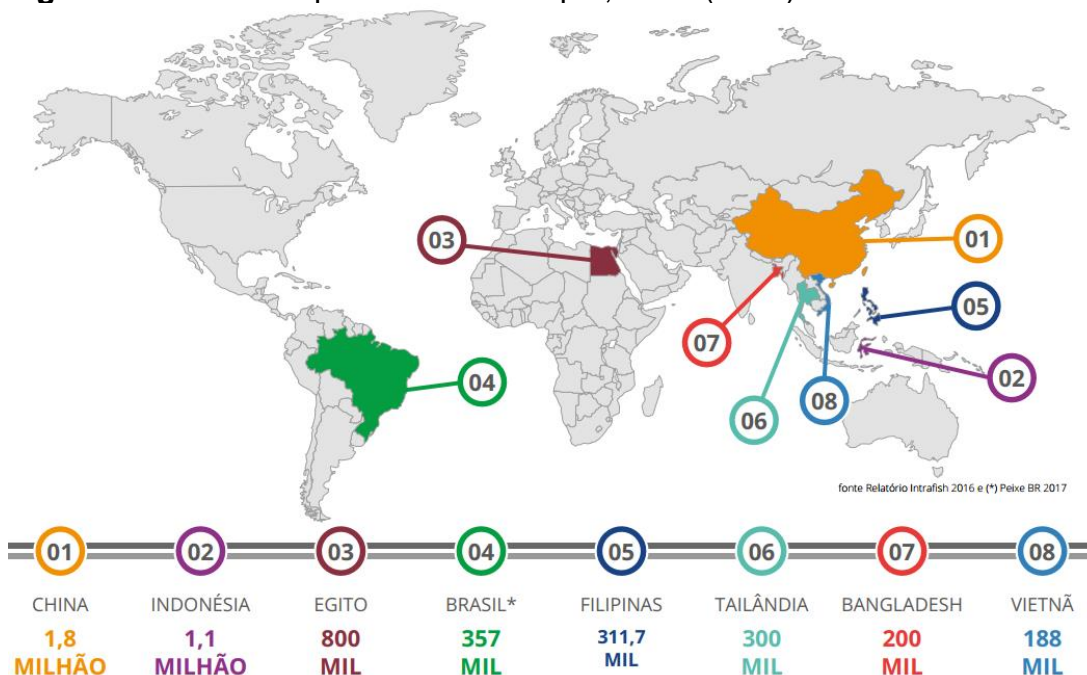
2 REVISÃO DA LITERATURA

2.1 TILAPICULTURA

No ano de 2017 o Brasil produziu 357 mil toneladas de tilápia, ficando atrás apenas de China, Indonésia e Egito (Figura 1). A tilápia é a espécie de maior produção no país, com 51.7% do total da produção de peixes cultivados, e é responsável pela movimentação de U\$1.5 bilhões ao ano (PEIXEBR, 2018). Segundo Kobayashi et al., 2015, avaliando dados do banco mundial, estima-se que América Latina, Caribe, Índia e Sudeste Asiático tenham um crescimento maior que 100% no período de 2010 à 2030, e que a produção de peixes de cultivo ultrapassem peixes de captura até o ano de 2021.

Devido à rusticidade da tilápia e a abundância de recursos, o Brasil tem demonstrado um amplo crescimento na aquicultura liderado pelos estados do Paraná, Rondônia, São Paulo e Mato Grosso principalmente quando considerados os peixes nativos de cultivo. No entanto, esse crescimento acelerado estabelece novos desafios à biossegurança da cadeia de produção de tilápias.

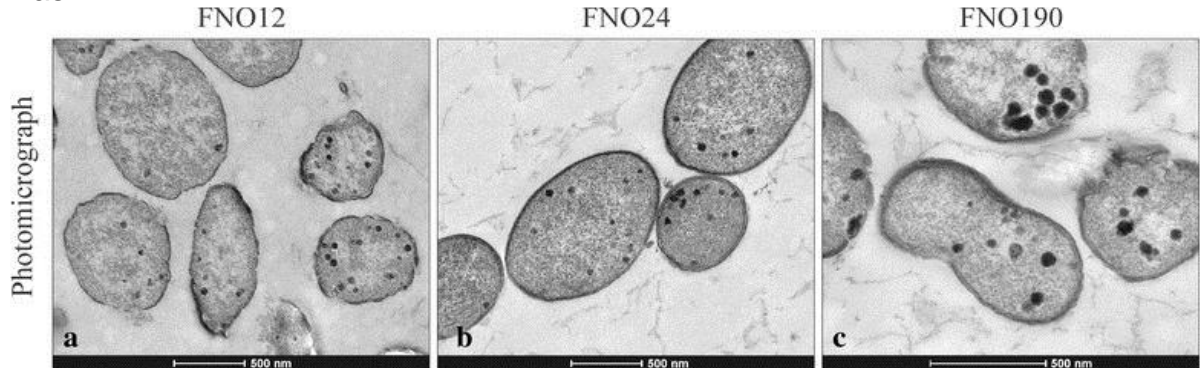
Figura 1 Os maiores produtores de tilápia, em T (2016).



2.2 GÊNERO *FRANCISELLA* SPP.

O gênero *Francisella* spp. é parte da família *Francisellaceae* e abriga microrganismos adaptados à diversos hospedeiros e habitats, podendo causar tularemia em humanos e franciselose em peixes. Bactérias do gênero *Francisella* spp. apresentam morfologia de cocobacilos, Gram negativos, aeróbicos e apresentam crescimento em temperaturas de 20 a 37°C (Figura 2)(OTTEM et al., 2007, 2009; SOTO et al., 2011b; JATOBÁ; KLIPP; HOPPE, 2016; ORTEGA et al., 2016) Embora pouco definida, a taxonomia do gênero *Francisella* spp. pode ser classificada em: *F. tularensis*, que engloba as subespécies *tularensis*, *mediasiatica*, *novicida* e *holarctica*; *F. haliotcida*; *F. hispaniensis*; *F. noatunensis*, que engloba as subespécies *noatunensis* e *orientalis*; *F. persica* e *F. philomiragia* (LARSSON et al., 2005; HUBER et al., 2010; BREVIK et al., 2011; VESTVIK et al., 2013; KINGRY; PETERSEN, 2014; RELICH et al., 2015; LARSON et al., 2016; SEBASTIÃO et al., 2017; TIMOFEEV et al., 2017).

Figura 2 Fotomicrografia de cepas de *F. noatunensis* subsp. *orientalis* isoladas no Brasil.



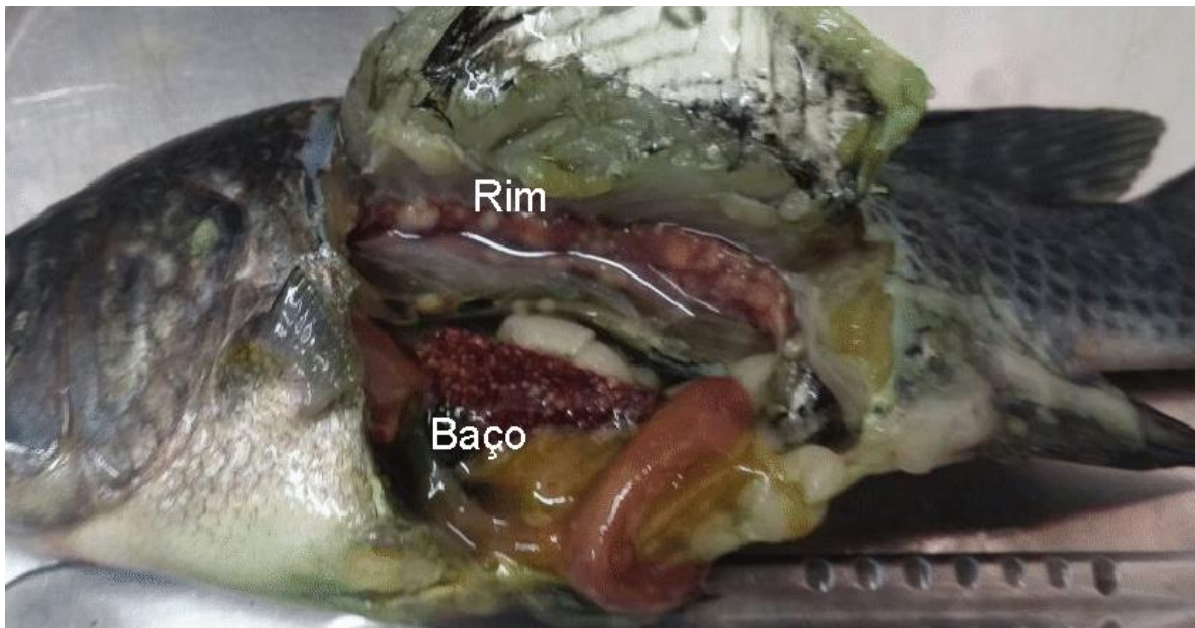
Fonte: Gonçalves, 2016.

2.3 FRANCISELOSE

A ocorrência de franciselose em peixes se dá principalmente devido à espécie *F. noatunensis* subsp. *orientalis* e *F. noatunensis* subsp. *noatunensis*. Os animais acometidos apresentam letargia, melanose, exoftalmia, e natação errática (JEFFERY et al., 2010; SOTO et al., 2011c; ASSIS et al., 2017). Dentre os fatores que predisõem à franciselose, o estresse térmico tem um papel fundamental, pois a

doença se manifesta principalmente em estações frias (LEAL; TAVARES; FIGUEIREDO, 2014). Embora a temperatura seja um fator essencial à ocorrência da doença, tilápias podem desenvolver franciselose a partir de doses infectantes baixas como por exemplo 23 UFC/mL em 10L (SOTO et al., 2014). Casos de *F. noatunensis* ocorrem com maior frequência em peixes pesando menos de 100g (SOTO et al., 2014; ASSIS et al., 2017) e óbitos ocorrem a partir de 2 dias (ASSIS et al., 2017). Em estudos realizados com *zebrafish*, cepas de *Francisella* spp. produziram lesões em fígado, baço, pâncreas e rim, com apresentação de estágios iniciais de granuloma (Figura 3). A resposta do hospedeiro foi avaliada através da expressão gênica, que apresentou aumento da expressão (*upregulation*) de genes associados à produção de interleucina-1-beta (IL-1 β), interferon-gama (IFN- γ) e fator de necrose tumoral (TNF) em peixes *zebrafish* inoculados via intraperitoneal (VOJTECH et al., 2009).

Figura 3 Rim cranial e baço de tilápia com apresentação de granulomas de franciselose



Fonte: Laboratório de Bacteriologia em Peixes/DMVP/CCA, Universidade Estadual de Londrina

2.4 CONTROLE E PREVENÇÃO

Estudos de sensibilidade à antibióticos relatam a resistência de uma cepa de *Francisella* spp. isolada em peixe resistente à trimetoprim, sulfametazol, penicilina, ampicilina, cefuroxima e eritromicina, e susceptibilidade à ceftazidima, tetraciclina, gentamicina e ciprofloxacina (OTTEM et al., 2007). De acordo com Soto et al. 2010, a incorporação da dose de 15mg/Kg de florfenicol na ração fornecida para animais desafiados demonstrou uma taxa de 100% e 86.7 % de sobrevivência nos grupos 1 e 3 dias após infecção, respectivamente. Estudos com a oxitetraciclina no tratamento de peixes relatam que a dose de 30-50mg/Kg de peso vivo incorporados na ração colaboram na redução da mortalidade do plantel (CHERN; CHAO, 1994).

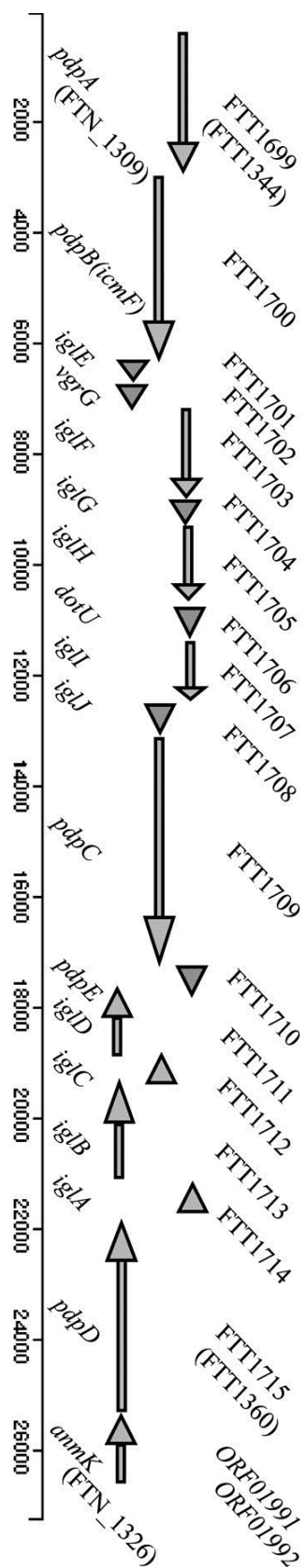
A vacinação em patógenos intracelulares apresentam baixa eficácia quando o patógeno é capaz de evadir os mecanismos de formação do fagolisossomo, o que impede a sua lise e apresentação antigênica via ativação de MHC. Após a fagocitose, as bactérias do gênero *Francisella* spp. são capazes de degradar a membrana fagossômica, liberando o microrganismo no citosol. A replicação da bactéria no interior da célula leva à lise da última, liberando as bactérias ao meio extracelular (FULLER et al., 2008, 2009; PECHOUS; MCCARTHY; ZAHRT, 2009). Embora ainda não disponível, a produção de uma vacina com genes mutantes tem sido considerada. Estudos demonstraram que o gene *iglC* tem papel importante na virulência da bactéria, que apresenta *upregulation* no momento da infecção, portanto uma vacina atenuada utilizando o gene *iglC* inativado tem apresentado 90% de proteção contra *F. noatunensis* subsp. *orientalis* (SOTO; FERNANDEZ; HAWKE, 2009; SOTO et al., 2011a; LAMPE et al., 2017).

2.5 ILHAS DE PATOGENICIDADE

Os genes responsáveis pela sobrevivência da bactéria no interior da célula tem sido analisados através de estudos de genômica comparativa. Por sua vez, verificou-se que existem sítios no genoma de *Francisella* spp. que estão associados à invasão, sobrevivência da bactéria no interior da célula e fuga do fagolisossomo. Essas regiões são conhecidas por *Francisella Pathogenicity Islands* (FPis), e consistem de agrupamentos de genes que foram identificados em estudos de expressão de proteínas de *Francisella* spp. no momento da infecção. Dentre eles,

*IgI*ABCDEFGHIJ, *pdp*ABCE, *dotU*, *vgrG* e *anmK* (Figura 4)(NANO; SCHMERK, 2007; BRÖMS; SJÖSTEDT; LAVANDER, 2010; FARON et al., 2013). Law et al., 2014 observou que a deleção de genes como o *iglC* e *pdpA*, que são componentes do FPI, reduziu a habilidade da bactéria invadir e replicar em células epiteliais, esses mesmos genes quando introduzidos de volta ao genoma, restauraram parcialmente às habilidades de invasão e crescimento intracelular. Resultados semelhantes também foram obtidos por Robertson et al., 2013 que testou a deleção do gene *iglE*.

Figura 4 Ilhas de Patogenicidade do Gênero *Francisella* (FPI). Representação das ilhas genômicas de *F. novicida*, *loci*, ORFs e nomenclatura do genes.



Fonte: Ludu (2008)

2.6 SAÚDE PÚBLICA

Em saúde pública, a espécie *F. tularensis* possui grande importância devido ao seu potencial infeccioso, podendo causar a tularemia em humanos mesmo em baixas contagens de microrganismos (OYSTON; SJOSTEDT; TITBALL, 2004). As ocorrências em humanos geralmente acometem indivíduos de áreas rurais, caçadores, trabalhadores da indústria de couro ou carne. A tularemia também pode ser adquirida pela picada de artrópode, inalação, ingestão de tecidos de animais contaminados como roedores e coelhos, ou contato com água ou solo contaminado (KINGRY; PETERSEN, 2014; KILIC et al., 2015). Na Turquia, Kilic et al., 2015 relata uma alta similaridade entre cepas de *Francisella* spp. obtidas em amostras de água e cepas identificadas de casos clínicos na mesma localidade. Na China, cepas de *Francisella* spp. foram isoladas de sistemas de ar condicionado e as análises comparativas dessas amostras apresentaram 95% de identidade com cepas de *F. philomiragia* e *F. noatunensis* (QU et al., 2013).

De acordo com (SOTO et al., 2009), os fatores determinantes de virulência da *F. noatunensis* e *F. tularensis* são semelhantes, o que sugere a utilização de tilápias como modelo de infecção de *F. tularensis*. No entanto, é pouco provável que humanos venham a ser infectados através do manuseio ou consumo de carne de/ tilápia. Nenhum dos subgrupos de *F. noatunensis* é capaz de sobreviver a temperaturas acima de 30°C, portanto não estão adaptados a humanos. Também não há manifestação de doença em camundongos experimentalmente infectados com a espécie *F. noatunensis* (MIKALSEN et al., 2009).

2.7 GENÔMICA COMPARATIVA

A genômica comparativa objetiva identificar similaridades e diferenças entre os organismos através da comparação do seu material genético. Esta abordagem colabora na identificação de regiões codificantes no genoma, que por sua vez, permite a caracterização da provável função dessas regiões codificantes, plasticidade genômica e estabelecimento de relações filogenéticas entre espécies e cepas (SIVASHANKARI; SHANMUGHAVEL, 2007; OGIER et al., 2010). A busca por genes compartilhados entre humanos e outros organismos possibilita a experimentação de novas terapias em organismos modelo, obtendo resultados

similares. Além disso, os sistemas de produção intensivos tem favorecido a taxa de recombinação de bactérias causadoras de doenças, por sua vez, permitindo uma rápida adaptação a diferentes ambientes e hospedeiros, promovendo o surgimento de cepas zoonóticas (WOODCOCK et al., 2017).

A utilização da genômica comparativa em procariotos permite uma maior definição da diversidade genômica entre espécies correlatas e intra-espécies. Um estudo com 45 genomas de *Francisella* spp. desenvolvido por Sjodin et al., 2012, demonstrou através da análise do core genômico que as espécies *F. tularensis* e subespécies que geralmente infectam mamíferos pertenceriam a um agrupamento diferente da *F. noatunensis*, sugerindo que estas duas espécies teriam evoluído independentemente por adaptação ao hospedeiro, e que seu ancestral comum teve origens em um ambiente marinho. Sridhar et al. 2012, através de comparação das características genômicas de cepas de *Francisella* que acometem mamíferos e do ambiente, observou que cepas não virulentas para humanos possuem apenas uma ilha de patogenicidade, com uma baixa contagem de elementos de inserção (IS)(Tabela 1).

A grande diversidade de genomas bacterianos disponíveis em banco de dados atualmente permite estudos aprofundados na área de filogenômica, pangenômica e plasticidade genômica.

Tabela 1 Principais Características de genomas sequenciados de *Francisella* spp.

	<i>F. noatunensis</i> subsp. <i>orientalis</i> Toba04	<i>F. philomiragia</i> subsp. <i>philomiragia</i> ATCC 25017	<i>F. tularensis</i> subsp. <i>novicida</i> U112	<i>F. tularensis</i> subsp. <i>medisaticca</i> FSC 147	<i>F. tularensis</i> subsp. <i>tularensis</i> SCHU S4	<i>F. tularensis</i> subsp. <i>holarctica</i> OSU18
	Fish parasite	Free-living	Free-living	Mammalian parasite	Mammalian parasite	Mammalian parasite
Genome size(bp)	1,847,202	2,045,775	1,910.03	1,893,886	1,892,775	1,895,727
GC content(%)	32	32	32	32	32	32
ORFs	2289	1966	1781	1750	1852	1932
Protein coding genes	1595	1911	1719	1406	1604	1555
Structural RNAs	39	48	48	47	48	49
IS elements	0	8	29	85	78	116
Pathogenicity Island	1	1	1	2	2	2
Pseudogenes	252	3	14	297	200	328

Fonte: Sridhar,2012

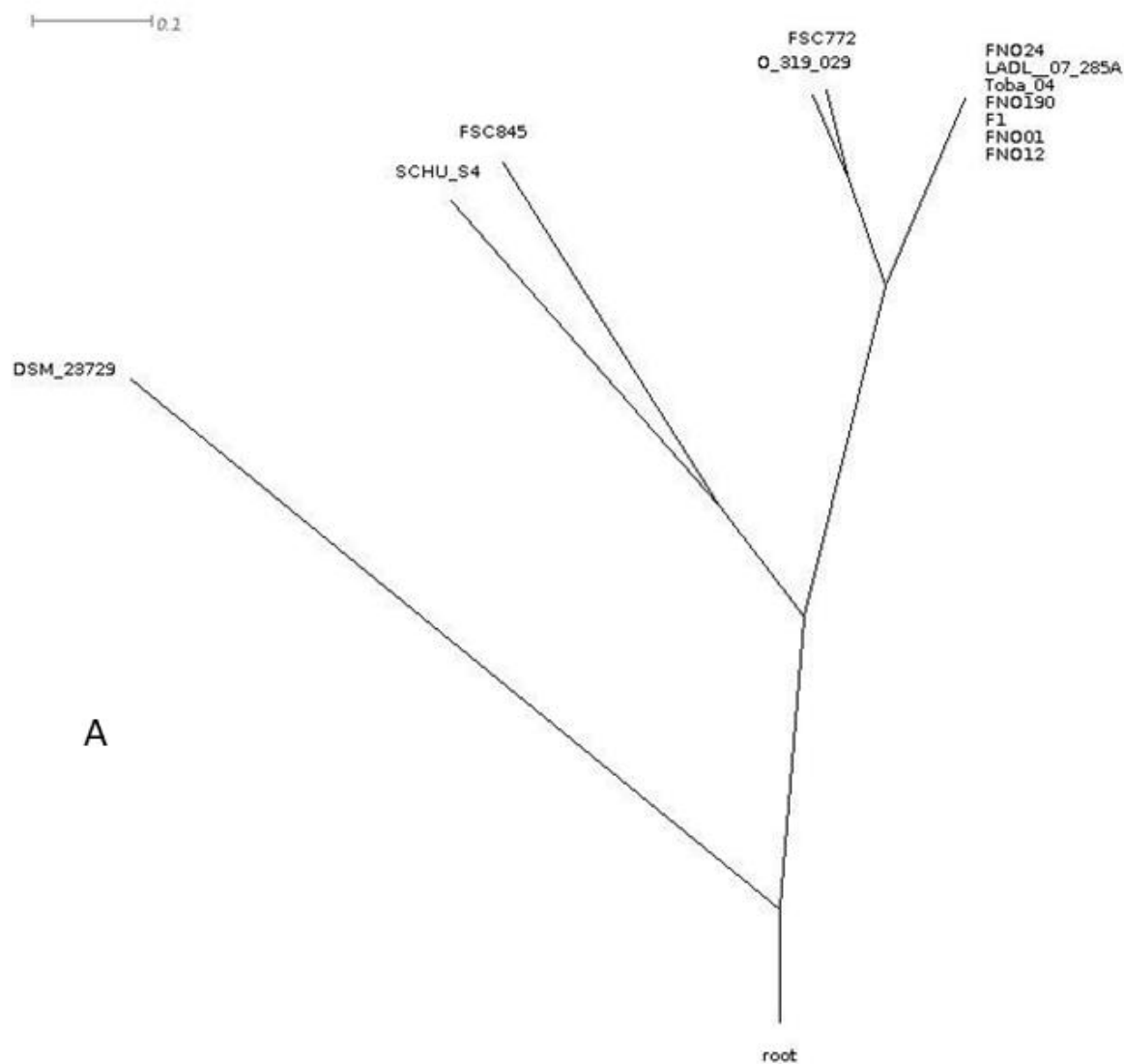
2.8 FILOGENÔMICA

Com o avanço e acessibilidade das técnicas moleculares, os estudos de filogenética foram amplamente aprofundados permitindo avaliar diferenças a nível de nucleotídeos e utilizando genomas completos. O uso de genomas completos permite novas abordagens para a análise de genomas, pois leva em consideração o conteúdo e ordem gênica; ortologia; e, estrutura da fita de DNA (DELSUC; BRINKMANN; PHILIPPE, 2005). Portanto, enquadram-se tais modelos de experimentação como sendo filogenômicos, devido à junção dos estudos de genômica e filogenética (OCAÑA; DÁVILA, 2011; CHAN; RAGAN, 2013).

A reconstrução de árvores filogenômicas e inferência de divergências evolucionárias podem ser realizadas utilizando a ferramenta Gegenees. O Gegenees é um software que quebra os genomas estudados em sequências pequenas

utilizando tamanhos pré-definidos, realiza pesquisas de similaridade utilizando BLAST, identifica genes comumente compartilhados entre os genomas e cria uma matriz de distância baseada na porcentagem de similaridade entre os conteúdos variáveis dos genomas em questão (Figura 5) (ÅGREN et al., 2012). Outra abordagem interessante é a utilização dos genomas para identificação de sítios polimórficos em uma estratégia de tipagem de sequência multilocus de genoma inteiro (whole-genome MultiLocus Sequence Typing ou wgMLST), também conhecida como metodologia gene a gene (MAIDEN et al., 1998; KLUYTMANS-VAN DEN BERGH et al., 2016).

Figura 5 Arvore filogenética utilizando o genoma completo de cepas de *Francisella* spp.



2.9 VACINOLOGIA REVERSA

A vacinologia reversa baseia-se na predição de proteínas que podem produzir imunogenicidade no hospedeiro utilizando a(s) sequência(s) genômica do patógeno (RAPPUOLI, 2000), diferentemente das metodologias tradicionais que se baseiam no cultivo do patógeno e isolamento individual (RAPPUOLI; COVACCI, 2003). A produção de vacinas por métodos convencionais possui algumas limitações, como: patógenos não cultiváveis, bactérias intracelulares que requerem cultivos de alto custo, falhas na inativação dos microorganismos, reversão da virulência (vacinas vivas atenuadas), ou antígenos não expressos ou expressos em pequenas quantidades (MOVAHEDI; HAMPSON, 2008). A primeira vacina desenvolvida utilizando esta metodologia já está licenciada para uso contra o sorogrupo B de *Neisseria meningitidis* (SERRUTO et al., 2012).

Através da análise de genomas de patógenos disponíveis em bancos de dados, a vacinologia reversa permite o levantamento de genes candidatos vacinais pela predição *in silico* para então estes serem clonados e testados *in vitro* e *in vivo*. Esta predição considera proteínas ancoradas na membrana ou secretadas pelo patógeno, devido à interação direta dessas estruturas com o sistema imune do hospedeiro permitindo a adesão, invasão, dano ao tecido hospedeiro, resistência e evasão do sistema imune do hospedeiro (SANTOS et al., 2011). Para essa predição, todos os genes do organismo são considerados independentes de serem expressos ou não. No entanto, antígenos não proteicos (polissacarídeos e glicolipídios) não podem ser estimados (RAPPUOLI, 2001).

Para identificar os candidatos proteicos, busca-se por genes compartilhados em comum a todas espécies (core genoma), pois estes podem induzir uma resposta imune em todas as variações de sorotipo ou de antigênicas das linhagens consideradas. Em seguida, a localização subcelular das proteínas preditas são identificadas (pan-exoproteoma e proteínas de membrana) e análises quanto à probabilidade de adesão e à propriedade de ligação à MHC I e II nos dados resultantes. Por último, pode-se realizar uma busca por fatores de virulência associados à ilhas de patogenicidade, pois estas regiões abrigam candidatos proteicos de maior imunogenicidade (SOARES et al., 2016).

REFERÊNCIAS

- ÅGREN, J.; SUNDSTRÖM, A.; HÅFSTRÖM, T.; SEGERMAN, B. Gegenees: Fragmented Alignment of Multiple Genomes for Determining Phylogenomic Distances and Genetic Signatures Unique for Specified Target Groups. **PLoS ONE**, v. 7, n. 6, p. e39107, 2012.
- ASSIS, G. B. N.; TAVARES, G. C.; PEREIRA, F. L.; FIGUEIREDO, H. C. P.; LEAL, C. A. G. Natural coinfection by *Streptococcus agalactiae* and *Francisella noatunensis* subsp. *orientalis* in farmed Nile tilapia (*Oreochromis niloticus*). **Journal of Fish Diseases**, v. 40, n. 1, p. 51–63, 2017.
- BREVIK, Ø. J.; OTTEM, K. F.; KAMAISHI, T.; WATANABE, K.; NYLUND, A. *Francisella halioticida* sp. nov., a pathogen of farmed giant abalone (*Haliotis gigantea*) in Japan. **Journal of Applied Microbiology**, v. 111, p. 1044–1056, 2011.
- BRÖMS, J. E.; SJÖSTEDT, A.; LAVANDER, M. The role of the *Francisella tularensis* pathogenicity island in type VI secretion, intracellular survival, and modulation of host cell signaling. **Frontiers in Microbiology**, v. 1, n. 136, p. 1–17, 2010.
- CHAN, C. X.; RAGAN, M. A. Next-generation phylogenomics. **Biology Direct**, v. 8, n. 1, p. 3, 2013.
- CHERN, R. S.; CHAO, C. B. Outbreaks of a Disease Caused by *Rickettsia*-like in Cultured Tilapias in Taiwan. **Fish Pathology**, v. 29, n. 2, p. 61-71, 1994.
- DELSUC, F.; BRINKMANN, H.; PHILIPPE, H. Phylogenomics and the reconstruction of the tree of life. **Nature Reviews Genetics**, v. 6, n. 5, p. 361–375, 2005.
- FARON, M.; FLETCHER, J. R.; RASMUSSEN, J. A.; LONG, M. E.; ALLEN, L. A. H.; JONES, B. D. The *Francisella tularensis* migr, trme, and cpha genes contribute to *F. tularensis* pathogenicity island gene regulation and intracellular growth by modulation of the stress alarmone ppGpp. **Infection and Immunity**, v. 81, n. 8, p. 2800–2811, 2013.
- FIELD, D.; GARRITY, G.; GRAY, T.; MORRISON, N.; SELENGUT, J.; STERK, P.; TATUSOVA, T.; THOMSON, N.; ALLEN, M. J.; ANGIUOLI, S. V.; ASHBURNER, M.; AXELROD, N.; OLIVER GLÖCKNER, F.; GOLDSTEIN, P.; GURALNICK, R.; HAFT, D.; HANCOCK, D.; HERMJAKOB, H.; HERTZ-FOWLER, C.; HUGENHOLTZ, P.; JOINT, I.; KAGAN, L.; KANE, M.; KENNEDY, J.; BALDAUF, S.; BALLARD, S.; COCHRANE, G.; COLE, J.; DAWYNDT, P.; VOS, P. De; GILNA, P.; GLÖCKNER, F. O.; GOLDSTEIN, P.; GURALNICK, R.; HAFT, D.; KAGAN, L.; KANE, M.; KENNEDY, J.; KOWALCHUK, G.; NELSON, K.; PARKHILL, J.; PROCTOR, L.; WHITE, O.; SPIERS, A.; STEVENS, R.; SWIFT, P.; TAYLOR, C.; TETT, A.; TURNER, S.; USSERY, D.; VAUGHAN, B.; WARD, N. The minimum information about a genome sequences (MIGS) specification. **Nature Biotechnology**, v. 26, n. 5, p. 541–547, 2008.

FULLER, J. R.; CRAVEN, R. R.; HALL, J. D.; KJEK, T. M.; TAFT-BENZ, S.; KAWULA, T. H. RipA, a cytoplasmic membrane protein conserved among *Francisella* species, is required for intracellular survival. **Infection and Immunity**, v. 76, n. 11, p. 4934–4943, 2008.

FULLER, J. R.; KJEK, T. M.; TAFT-BENZ, S.; KAWULA, T. H. Environmental and intracellular regulation of *Francisella tularensis* ripA. **BMC Microbiology**, v. 9, n. 1, p. 216, 2009.

GIULIANI, M. M.; ADU-BOBIE, J.; COMANDUCCI, M.; ARICO, B.; SAVINO, S.; SANTINI, L.; BRUNELLI, B.; BAMBINI, S.; BIOLCHI, A.; CAPECCHI, B.; CARTOCCI, E.; CIUCCHI, L.; DI MARCELLO, F.; FERLICCA, F.; GALLI, B.; LUZZI, E.; MASIGNANI, V.; SERRUTO, D.; VEGGI, D.; CONTORNI, M.; MORANDI, M.; BARTALESI, A.; CINOTTI, V.; MANNUCCI, D.; TITTA, F.; OVIDI, E.; WELSCH, J. A.; GRANOFF, D.; RAPPUOLI, R.; PIZZA, M. A universal vaccine for serogroup B meningococcus. **Proceedings of the National Academy of Sciences**, v. 103, n. 29, p. 10834–10839, 2006.

GONÇALVES, L. A.; SOARES, S. C.; PEREIRA, F. L.; DORELLA, F. A.; DE CARVALHO, A. F.; ALMEIDA, G. M. F. Complete genome sequences of *Francisella noatunensis* subsp. *orientalis* strains FNO12, FNO24 and FNO190: A fish pathogen with genomic clonal behavior. **Standards in Genomic Sciences**, v. 11, p. 30, 2016.

HUBER, B.; ESCUDERO, R.; BUSSE, H.-J.; SEIBOLD, E.; SCHOLZ, H. C.; ANDA, P.; KAMPFER, P.; SPLETTSTOESSER, W. D. Description of *Francisella hispaniensis* sp. nov., isolated from human blood, reclassification of *Francisella novicida* (Larson et al. 1955) Olsufiev et al. 1959 as *Francisella tularensis* subsp. *novicida* comb. nov. and emended description of the genus *Francisella*. **International Journal of Systematic and Evolutionary Microbiology**, v. 60, n. 8, p. 1887–1896, 2010.

JATOBÁ, A.; KLIPP, S. P.; HOPPE, R. Primeiro relato de *Francisella noatunensis* subespécie *orientalis* no sul do Brazil – Relato de caso. **Acta Veterinaria Brasilica**, v. 10, n. 2, p. 172, 2016.

JEFFERY, K.; STONE, D.; FEIST, S.; VERNER-JEFFREYS, D. An outbreak of disease caused by *Francisella* sp. in Nile tilapia *Oreochromis niloticus* at a recirculation fish farm in the UK. **Diseases of Aquatic Organisms**, v. 91, n. 2, p. 161–165, 2010.

KILIC, S.; BIRDSELL, D. N.; KARAGÖZ, A.; ÇELEBI, B.; BAKKALOĞLU, Z.; ARIKAN, M.; SAHL, J. W.; MITCHELL, C.; RIVERA, A.; MALTINSKY, S.; KEIM, P.; DURMAZ, R.; WAGNER, D. M. Water as source of *Francisella tularensis* infection in humans, Turkey. **Emerging Infectious Diseases**, v. 21, n. 12, p. 2213–2216, 2015.

KINGRY, L. C.; PETERSEN, J. M. Comparative review of *Francisella tularensis* and *Francisella novicida*. **Frontiers in Cellular and Infection Microbiology**, v. 4, p. 1–12, 2014.

KLUYTMANS-VAN DEN BERGH, M. F. Q.; ROSSEN, J. W. A.; BRUJNING-VERHAGEN, P. C. J.; BONTEN, M. J. M.; FRIEDRICH, A. W.; VANDENBROUCKE-GRAULS, C. M. J. E.; WILLEMS, R. J. L.; KLUYTMANS, J. A. J. W. Whole-Genome Multilocus Sequence Typing of Extended-Spectrum-Beta-Lactamase-Producing *Enterobacteriaceae*. **Journal of Clinical Microbiology**, v. 54, n. 12, p. 2919–2927, 2016.

KOBAYASHI, M.; MSANGI, S.; BATKA, M.; VANNUCCINI, S.; DEY, M. M.; ANDERSON, J. L. Fish to 2030: The Role and Opportunity for Aquaculture. **Aquaculture Economics and Management**, v. 19, n. 3, p. 282–300, 2015.

LAMPE, E. O.; TANDBERG, J. I.; RISHOVD, A. L.; WINTHER-LARSEN, H. C. *Francisella noatunensis* ssp. *noatunensis* igIC deletion mutant protects adult zebrafish challenged with acute mortality dose of wild-type strain. **Diseases of Aquatic Organisms**, v. 123, n. 2, p. 123–140, 2017.

LARSON, M. A.; NALBANTOGLU, U.; SAYOOD, K.; ZENTZ, E. B.; CER, R. Z.; IWEN, P. C.; FRANCESCONI, S. C.; BISHOP-LILLY, K. A.; MOKASHI, V. P.; SJÖSTEDT, A.; HINRICHS, S. H. Transfer of *Wolbachia persica* to the genus *Francisella* and reclassification as *Francisella persica* comb. nov. with emended description of the family *Francisellaceae*. **International Journal of Systematic and Evolutionary Microbiology**, v. 66, p. 1200–1205, 2016.

LARSSON, P.; OYSTON, P. C. F.; CHAIN, P.; CHU, M. C.; DUFFIELD, M.; FUXELIUS, H.-H.; GARCIA, E.; HÄLLTORP, G.; JOHANSSON, D.; ISHERWOOD, K. E.; KARP, P. D.; LARSSON, E.; LIU, Y.; MICHELL, S.; PRIOR, J.; PRIOR, R.; MALFATTI, S.; SJÖSTEDT, A.; SVENSSON, K.; THOMPSON, N.; VERGEZ, L.; WAGG, J. K.; WREN, B. W.; LINDLER, L. E.; ANDERSSON, S. G. E.; FORSMAN, M.; TITBALL, R. W. The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. **Nature Genetics**, v. 37, n. 2, p. 153–159, 2005.

LAW, H. T.; SRIRAM, A.; FEVANG, C.; NIX, E. B.; NANO, F. E.; GUTTMAN, J. A. IgIC and PdpA are important for promoting *Francisella* invasion and intracellular growth in epithelial cells. **PLoS ONE**, v. 9, n. 8, p. 1–12, 2014.

LEAL, C. A. G.; TAVARES, G. C.; FIGUEIREDO, H. C. P. Outbreaks and genetic diversity of *Francisella noatunensis* subsp. *orientalis* isolated from farm-raised Nile tilapia (*Oreochromis niloticus*) in Brazil. **Genetics and Molecular Research**, v. 13, n. 3, p. 5704–5712, 2014.

LUDU, J. S.; DE BRUIN, O. M.; DUPLANTIS, B. N.; SCHMERK, C. L.; CHOU, A. Y.; ELKINS, K. L., NANO, F. E. The *Francisella* pathogenicity island protein PdpD is required for full virulence and associates with homologues of the type VI secretion system. **Journal of Bacteriology**, v. 190, n. 13, p. 4584–4595, 2008.

MAIDEN, M. C.; BYGRAVES, J. A.; FEIL, E.; MORELLI, G.; RUSSELL, J. E.; URWIN, R.; ZHANG, Q.; ZHOU, J.; ZURTH, K.; CAUGANT, D. A.; FEAVERS, I. M.; ACHTMAN, M.; SPRATT, B. G.; FERLICCA, F.; GALLI, B.; LUZZI, E.; MASIGNANI, V.; SERRUTO, D.; VEGGI, D.; CONTORNI, M.; MORANDI, M.; BARTALESI, A.; CINOTTI, V.; MANNUCCI, D.; TITTA, F.; OVIDI, E.; WELSCH, J. A.; GRANOFF, D.; RAPPUOLI, R.; PIZZA, M. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. **Proceedings of the National Academy of Sciences of the United States of America**, v. 95, n. 6, p. 3140–5, 1998.

MIKALSEN, J.; OLSEN, A. B.; RUDRA, H.; MOLDAL, T.; LUND, H.; DJ??NNE, B.; BERGH, O.; COLQUHOUN, D. J. Virulence and pathogenicity of *Francisella philomiragia* subsp. *noatunensis* for Atlantic cod, *Gadus morhua*, and laboratory mice. **Journal of Fish Diseases**, v. 32, n. 4, p. 377–381, 2009.

MOVAHEDI, A. (Reza); HAMPSON, D. J. New ways to identify novel bacterial antigens for vaccine development. **Veterinary Microbiology**, v. 131, n. 1–2, p. 1–13, 2008.

NANO, F. E.; SCHMERK, C. The Francisella Pathogenicity Island. **Annals of the New York Academy of Sciences**, v. 1105, n. 1, p. 122–137, 2007.

OCAÑA, K. A. C. S.; DÁVILA, A. M. R. Phylogenomics-Based Reconstruction of Protozoan Species Tree. **Evolutionary Bioinformatics**, v. 7, p. 107–121, 2011.

OGIER, J.-C.; CALTEAU, A.; FORST, S.; GOODRICH-BLAIR, H.; ROCHE, D.; ROUY, Z.; SUEN, G.; ZUMBIHL, R.; GIVAUDAN, A.; TAILLIEZ, P.; MEDIGUE, C.; GAUDRIULT, S. Units of plasticity in bacterial genomes: new insight from the comparative genomics of two bacteria interacting with invertebrates, *Photobacterium* and *Xenorhabdus*. **BMC Genomics**, v. 11, n. 1, p. 568, 2010.

ORTEGA, C.; MANCERA, G.; ENRÍQUEZ, R.; VARGAS, A.; MARTÍNEZ, S.; FAJARDO, R.; AVENDAÑO-HERRERA, R.; NAVARRETE, M. J.; ROMERO, A. First identification of *Francisella noatunensis* subsp. *orientalis* causing mortality in mexican tilapia *Oreochromis* spp. **Diseases of Aquatic Organisms**, v. 120, n. 3, p. 205–215, 2016.

OTTEM, K. F.; NYLUND, A.; KARLSBAKK, E.; FRIIS-MØLLER, A.; KAMAISHI, T. Elevation of *Francisella philomiragia* subsp. *noatunensis* Mikalsen et al. (2007) to *Francisella noatunensis* comb. nov. [syn. *Francisella piscicida* Ottem et al. (2008) syn. nov.] and characterization of *Francisella noatunensis* subsp. *orientalis* subsp. nov., two important fish pathogens. **Journal of Applied Microbiology**, v. 106, n. 4, p. 1231–1243, 2009.

OTTEM, K. F.; NYLUND, A.; KARLSBAKK, E.; FRIIS-MØLLER, A.; KROSSØY, B. Characterization of *Francisella* sp., GM2212, the first *Francisella* isolate from marine fish, Atlantic cod (*Gadus morhua*). **Archives of Microbiology**, v. 187, n. 5, p. 343–350, 2007.

OYSTON, P. C. F.; SJOSTEDT, A.; TITBALL, R. W. Tularaemia: Bioterrorism defence renews interest in *Francisella tularensis*. **Nature Reviews Microbiology**, v. 2, n. 12, p. 967–978, 2004.

PECHOUS, R. D.; MCCARTHY, T. R.; ZAHRT, T. C. Working toward the future: insights into *Francisella tularensis* pathogenesis and vaccine development. **Microbiology and molecular biology reviews : MMBR**, v. 73, n. 4, p. 684–711, 2009.

PEIXEBR. **Anuário PeixeBR da Piscicultura 2018**. p. 12-30, 2018.

PEREIRA, U. P.; MIAN, G. F.; OLIVEIRA, I. C. M.; BENCHETRIT, L. C.; COSTA, G. M.; FIGUEIREDO, H. C. P. Genotyping of *Streptococcus agalactiae* strains isolated from fish, human and cattle and their virulence potential in Nile tilapia. **Veterinary Microbiology**, v. 140, n. 1–2, p. 186–192, 2010.

QU, P. H.; CHEN, S. Y.; SCHOLZ, H. C.; BUSSE, H. J.; GU, Q.; KÄMPFER, P.; FOSTER, J. T.; GLAESER, S. P.; CHEN, C.; YANG, Z. C. *Francisella guangzhouensis* sp. nov., isolated from air-conditioning systems. **International Journal of Systematic and Evolutionary Microbiology**, v. 63, n. 10, p. 3628–3635, 2013.

RAPPUOLI, R. Reverse vaccinology. **Current opinion in microbiology**, v. 3, n. 5, p. 445–50, 2000.

RAPPUOLI, R. Reverse vaccinology, a genome-based approach to vaccine development. **Vaccine**, v. 19, n. 17–19, p. 2688–91, 2001.

RAPPUOLI, R.; COVACCI, A. Reverse Vaccinology and Genomics. **Science**, v. 302, n. 5645, p. 602–602, 2003.

RELICH, R. F.; HUMPHRIES, R. M.; MATTISON, H. R.; MILES, J. E.; SIMPSON, E. R.; CORBETT, I. J.; SCHMITT, B. H.; MAY, M. *Francisella philomiragia* Bacteremia in a Patient with Acute Respiratory Insufficiency and Acute-on-Chronic Kidney Disease. **Journal of clinical microbiology**, v. 53, n. 12, p. 3947–50, 2015.

ROBERTSON, G. T.; CHILD, R.; INGLE, C.; CELLI, J.; NORGARD, M. V. IgE is an outer membrane-associated lipoprotein essential for intracellular survival and murine virulence of type A *Francisella tularensis*. **Infection and Immunity**, v. 81, n. 11, p. 4026–4040, 2013.

SANTOS, A.; ALI, A.; BARBOSA, E.; SILVA, A.; MIYOSHI, A.; BARH, D.; AZEVEDO, V. The reverse vaccinology - a contextual overview. **IIOAB Journal**, v. 2, n. 4, p. 8–15, 2011.

SEBASTIÃO, F. A. . A. b; PILARSKI, F. .; KEARNEY, M. T. . T.; SOTO, E. . Molecular detection of *Francisella noatunensis* subsp. *orientalis* in cultured Nile tilapia (*Oreochromis niloticus*) in three Brazilian states. **Journal of Fish Diseases**, v. 40, n. 11, p. 1731–1735, 2017.

SERRUTO, D.; BOTTOMLEY, M. J.; RAM, S.; GIULIANI, M. M.; RAPPUOLI, R. The new multicomponent vaccine against meningococcal serogroup B, 4CMenB: Immunological, functional and structural characterization of the antigens. **Vaccine**, v. 30, p. 87–97, 2012.

SETTE, A.; RAPPUOLI, R. Reverse vaccinology: developing vaccines in the era of genomics. **Immunity**, v. 33, n. 4, p. 530–541, 2010.

SIVASHANKARI, S.; SHANMUGHAVAL, P. Comparative genomics - a perspective. **Bioinformatics**, v. 1, n. 9, p. 376–378, 2007.

SJÖDIN, A.; SVENSSON, K.; OHRMAN, C.; AHLINDER, J.; LINDGREN, P.; DUODU, S.; JOHANSSON, A.; COLQUHOUN, D. J.; LARSSON, P.; FORSMAN, M. Genome characterisation of the genus *Francisella* reveals insight into similar evolutionary paths in pathogens of mammals and fish. **BMC genomics**, v. 13, n. 1, p. 268, 2012.

SOARES, S.C.; GEYIK, H.; RAMOS, R.T.J; SÁ, P.H.C.G; BARBOSA, E.G.V.; BAUMBACH, J.; FIGUEIREDO, H.C.P.; MIYOSHI, A TAUCH, A. ; SILVA, A. AZEVEDO, C.GIPSy: Genomic island prediction software. **Journal of Biotechnology**, v. 232, p. 2-16, 2016.

SOTO, E.; BROWN, N.; GARDENFORS, Z. O.; YOUNT, S.; REVAN, F.; FRANCIS, S.; KEARNEY, M. T.; CAMUS, A. Effect of size and temperature at vaccination on immunization and protection conferred by a live attenuated *Francisella noatunensis* immersion vaccine in red hybrid tilapia. **Fish and Shellfish Immunology**, v. 41, n. 2, p. 593–599, 2014.

SOTO, E.; WILES, J.; ELZER, P.; MACALUSO, K.; HAWKE, J. P. Attenuated *Francisella asiatica* iglC mutant induces protective immunity to francisellosis in tilapia. **Vaccine**, v. 29, n. 3, p. 593–598, 2011a.

SOTO, E.; ABRAMS, S.; REVAN, F. Effects of temperature and salt concentration on *Francisella noatunensis* subsp. *orientalis* infections in Nile tilapia *Oreochromis niloticus*. **Diseases of Aquatic Organisms**, v. 101, n. 3, p. 217–223, 2012.

SOTO, E.; BAUMGARTNER, W.; WILES, J.; HAWKE, J. P. *Francisella asiatica* as the causative agent of piscine francisellosis in cultured tilapia (*Oreochromis* sp.) in the United States. **Journal of Veterinary Diagnostic Investigation**, v. 23, n. 4, p. 821–825, 2011c.

SOTO, E.; ENDRIS, R. G.; HAWKE, J. P. *In vitro* and *in vivo* efficacy of florfenicol for treatment of *Francisella asiatica* infection in tilapia. **Antimicrobial Agents and Chemotherapy**, v. 54, n. 11, p. 4664–4670, 2010.

SOTO, E.; FERNANDEZ, D.; HAWKE, J. P. Attenuation of the fish pathogen *Francisella* sp. by mutation of the iglC* gene. **Journal of Aquatic Animal Health**, v. 21, n. 3, p. 140–149, 2009.

SOTO, E.; HAWKE, J. P.; FERNANDEZ, D.; A MORALES, J. *Francisella* sp., an emerging pathogen of tilapia, *Oreochromis niloticus*, in Costa Rica. **Journal of Fish Diseases**, v. 32, n. 8, p. 713–722, 2009.

SOTO, E.; WILES, J.; ELZER, P.; MACALUSO, K.; HAWKE, J. P. Attenuated *Francisella asiatica* igIC mutant induces protective immunity to francisellosis in tilapia. **Vaccine**, v. 29, n. 3, p. 593–598, 2011d.

SRIDHAR, S.; SHARMA, A.; KONGSHAUG, H.; NILSEN, F.; JONASSEN, F. Whole genome sequencing of the fish pathogen *Francisella noatunensis* subsp. *orientalis* Toba04 gives novel insights into *Francisella* evolution and pathogenicity. **BMC Genomics**, v. 13, n. 1, p. 598, 2012.

TIMOFEEV, V.; BAKHTEEVA, I.; TITAREVA, G.; KOPYLOV, P.; CHRISTIANY, D.; MOKRIEVICH, A.; DYATLOV, I.; VERGNAUD, G. Russian isolates enlarge the known geographic diversity of *Francisella tularensis* subsp. *mediasiatica*. **PLoS ONE**, v. 12, n. 9, p. 1–20, 2017.

VESTVIK, N.; RØNNESETH, A.; KALGRAFF, C. A. K.; WINTHER-LARSEN, H. C.; WERGELAND, H. I.; HAUGLAND, G. T. *Francisella noatunensis* subsp. *noatunensis* replicates within Atlantic cod (*Gadus morhua*) leucocytes and inhibits respiratory burst activity. **Fish & shellfish immunology**, v. 35, n. 3, p. 725–33, 2013.

VOJTECH, L. N.; SANDERS, G. E.; CONWAY, C.; OSTLAND, V.; HANSEN, J. D. Host immune response and acute disease in a zebrafish model of *Francisella* pathogenesis. **Infection and Immunity**, v. 77, n. 2, p. 914–925, 2009.

WOODCOCK, D. J.; KRUSCHE, P.; STRACHAN, N. J. C.; FORBES, K. J.; COHAN, F. M.; MÉRIC, G.; SHEPPARD, S. K. Genomic plasticity and rapid host switching can promote the evolution of generalism: a case study in the zoonotic pathogen *Campylobacter*. **Scientific Reports**, v. 7, n. 1, p. 9650, 2017.

WORLD HEALTH ORGANIZATION. Genomics and World Health. **World Health Organization**, v. 96, n. 6, p. 254, 2002.

3 HIPÓTESE

É possível identificar alvos vacinais conservados entre as duas subespécies de *F. noatunensis* para estudos de vacinologia em peixes de águas quentes e frias.

4 OBJETIVOS

4.1 OBJETIVO GERAL

Sequenciar, montar e depositar em bancos de dados públicos o genoma completo de uma cepa de *F. noatunensis* subsp. *orientalis* e predizer *in silico* candidatos proteicos imunogênicos conservados entre as cepas *F. noatunensis* subsp. *orientalis* e *F. noatunensis* subsp. *noatunensis* para o desenvolvimento de vacina eficaz contra a franciselose em peixes de águas quentes e frias.

4.2 OBJETIVOS ESPECÍFICOS

- Montar o genoma completo da cepa F1 utilizando dados de sequenciamento de nova geração;
- Identificar a localização subcelular de proteínas presentes na cepa F1;
- Investigar os mecanismos que envolvem as potenciais proteínas alvos de vacina;
- Testar *in silico* a imunogenicidade das proteínas alvo obtidas pela vacinologia reversa.

5 ARTIGO ACEITO PARA PUBLICAÇÃO NO PERIÓDICO STANDARDS IN GENOMICS SCIENCE

Complete genome sequence of *Francisella noatunensis* subsp. *orientalis* F1 strain isolated from a Brazilian tilapia farm and prediction of universal vaccine candidates against warm and cold water fish francisellosis

Authors

César T Facimoto¹, Roberta T Chideroli¹, Giovana W Di Santis¹, Daniela D Gonçalves², Anderson O Do Carmo³, Evanguedes Kalapothakis³, Alberto F de Oliveira Junior³, Vasco Azevedo³, Leandro de Jesus Benevides^{3,5}, Siomar Soares Castro⁴, Ulisses P Pereira^{*1}.

Institutional Affiliation

1 Laboratory of Bacteriology in Fish (LABBEP) – Department of Preventive Veterinary Medicine, State University of Londrina, Londrina, PR, Brazil;

2 Laboratory of Preventive Veterinary Medicine and Public Health – Universidade Paranaense, Umuarama, PR, Brazil;

3 Department of General Biology, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil;

4 Department of Immunology, Microbiology and Parasitology, Institute of Biological Sciences and Natural Sciences, Federal University of Triângulo Mineiro, Uberaba, MG, Brazil;

5 Commensal and Probiotics-Host Interactions Laboratory, Micalis Institute, Institut National de la Recherche Agronomique, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France.

Corresponding author

Ulisses de Padua Pereira, DVM, Ph.D.; Tel.: +55 43 33714765; fax: +55 43 33714485. E-mail address:

upaduapereira@uel.br

Abstract

Francisella noatunensis subsp. *orientalis* is a Gram negative bacteria that causes granulomatous infections in fish from low water temperatures and has been responsible for large number of deaths in tilapia farms. *F. noatunensis* has been reported in many countries in the last decade. However, studies about the phylogenomic relationships between isolates are necessary due to the emerging nature of the disease. This study aims to present the complete genome of the *F. noatunensis* F1 strain isolated in 2015 from a tilapia farm in the São Paulo state, Brazil. The genome is a circular chromosome with a final length of 1,854,333 bp, containing 32.26% G+C DNA content, 1,440 protein-coding genes and 369 pseudogenes. In addition, a prediction of conserved vaccine targets was performed between the two subspecies of *F. noatunensis* that causes the disease in tropical and cold water fish species. Five proteins conserved between *F. noatunensis* subsp. *orientalis* and *Francisella noatunensis* subsp. *noatunensis* were predicted to be good vaccine candidates for the development of a recombinant vaccine against francisellosis. Further, this genome provides additional valuable data to better understand the agent's evolution and epidemiology in Brazil.

Keywords

Infection diseases, vaccine, fish pathogen, genome sequencing, phylogeny.

Abbreviations

CHAH: Cysteine Heart Agar with Hemoglobin; HE: Hematoxylin and eosin; LGCM: Laboratory of Cellular and Molecular Genetics; LABBEP: Laboratory of Bacteriology in Fish; MIGS: Minimum Information on Genome Sequence; rRNA: ribosomal RNA; tRNA: transporter RNA; UPGMA: unweighted pair-group method with arithmetic mean; DDBJ: Data Bank of Japan; EMBL: European Molecular Biology Laboratory; NCBI: National Center for Biotechnology Information; CYT: cytoplasmatic; SEC: secreted; MEM: membrane; PSE: potentially surface exposed

5.1 Introduction

World aquaculture has increasingly intensified and diversified over the years, and this, resulted in the recognition of “new” infectious agents, previously considered unknown diseases [1]. In recent years, bacteria from the *Francisella* genus, *Francisella noatunensis* subsp. *noatunensis*, and *Francisella noatunensis* subsp. *orientalis*, have been associated with bacteriosis in fish species from cold and warm waters, respectively [2–4]. *Francisella* spp. infections are characterized as an acute (with few clinical signs and high mortality), a subacute, or a chronic (non-specific clinical signs with varying degrees of mortality) syndrome [1]. In Brazil, francisellosis losses in Nile Tilapia (*Oreochromis niloticus*) cultivated in cage tanks are reported since 2012 and the mortality rates are up to 60%. The most common clinical signs associated with infections are: anorexia, melanosis, exophthalmia, erratic swimming, skin ulcers (primarily at the base of the fins) and gill pallor [5].

This pathogen affects mainly fry, fingerling, and young adult fish, exhibiting multifocal granulomas in internal organs, mostly in liver, spleen, and kidney [2,6]. Cases of *F. noatunensis* subsp. *orientalis* in farm-raised tilapia have been reported in Taiwan [7], Costa Rica [2], Indonesia[8], UK [9], USA [10], and Brazil [11]. Research on *Francisella* spp. phylogenomic relationships and diversity are in constant development, at the same pace whole genomes are made public available. Due to its importance to aquaculture, *F. noatunensis* gained more attention in recent years, however, only few whole genomes are deposited in public databases for study. In this paper, we describe the whole genome features and vaccine candidates for the F1 strain isolated from a fish farm outbreak in the state of São Paulo, Brazil in 2015.

5.2 Organism Information

5.2.1 Classification and features

Bacteria from genus *Francisella*, are part of the *Francisellaceae* family and *Thiotrichales* order. Bacteriology assays characterize them as strictly aerobic, intracellular facultative, nonmotile, Gram negative coccobacilli of 0.5–1.5 μm (Table 1 and Figure 1). In addition, these bacteria are considered fastidious microorganisms that require Cysteine Heart Agar with Hemoglobin (CHAH) for growth, not growing on MacConkey agar. They are adapted to temperatures ranging from 6 to 28 °C, with ideal growth between 17 and 22°C [12]. *Francisella* spp. is typically catalase positive and cytochrome oxidase negative. Biochemical characterization showed negative for: nitrate reduction, urease, raffinose fermentation, esculin, adonitol, arabinose, rhamnose, mannitol, cellobiose, melibiose, glucuronate, malonate, para-phenylalanine deaminase, 5-ketogluconate, palatinose, galacturonate, tetrathionate reductase, raffinose, C14 lipase, valine arylamidase, cystine arylamidase, trypsin, a-chymotrypsin, agalactosidase, b-glucuronidase, a-mannosidase, a-fucosidase, and acid production from lactose. Also, resulted positive for the following biochemical tests: b lactamase, alkaline phosphatase, acid phosphatase, C4 esterase, C8 esterase, lipase, naphthol-AS-BI-phosphohydrolase, and acid production from maltose [13].

Non-specific clinical signs of this bacteriosis in fish include loss of appetite, lethargy, erratic swimming, and exophthalmia. More specific signs of the disease are observed internally such as: presence of a large number of white nodules on the gills, spleen, kidney, gonads, liver and heart.

In the present study, a histopathologic exam of the ocular globe, cranial kidney and spleen was performed using hematoxylin and eosin staining (HE). Microscopically, we observed a large accumulation of inflammatory cells promoting an increase in nodular volume between the choroid and the sclera (Figure 3A), with areas of well vascularized stroma, marked monophyletic and polymorphonuclear infiltrate (Figure 3B),

alternating with some that are clearly granulomatous (Figure 3C), composed of epithelioid cells and macrophages containing intracytoplasmic coccoid bacteria (Figure 3D). The same lesion was also present in other histological sections, extending discreetly to the retrobulbar periocular tissues. Granulomatous lesions were also observed on the spleen and cranial kidney.

This genus also contains species that infect humans, such as *Francisella tularensis* (the agent responsible for tularemia), and *Francisella philomiragia*, both associated with chronic and necrotizing granulomas. Although the available genomes of species from the genus *Francisella* have relatively similar features, their genomes are highly rearranged [14].

Twelve whole genome strains in the *Francisella* genus were submitted to Geneious V2.2.1 [15] analysis with high accuracy, generating a heatmap. A phylogenetic tree was drawn through SplitsTree4 set to EqualAngleRoot split and UPGMA methods in order to estimate the phylogenetic distance of our strain among other species in the genus.

5.3 Genome sequencing information

5.3.1 Genome project history

The F1 strain was selected for sequencing due to the emerging nature of this pathogen in global aquaculture. It was also selected due to its high mortality profile in aquatic environments [3,4].

The genome project is housed in the Genomes OnLine Database (GOLD) [16] project ID Gp0177338 and complete genome sequence and annotation data were uploaded in the DDBJ/EMBL/GenBank under the accession number CP018051. Sequencing and partial assembly were performed at the Laboratory of Cellular and Molecular Genetics (LGCM), Minas Gerais, Brazil. Final assembly, reverse vaccinology and comparative genomics were performed at the Laboratory of Bacteriology in Fish (LABBEP), Paraná, Brazil. Table 2 presents a summary of the information collected during this project and its association with MIGS version 2.0 compliance [17].

5.3.2 Growth conditions and genomic DNA preparation

The F1 strain was obtained from the Northwest region of the state of São Paulo in 2015. Swabs of kidney tissue were sampled aseptically, streaked onto cysteine heart agar supplemented with 1 % bovine hemoglobin (HIMEDIA, India) and incubated at 28 °C for 4–7 days [13]. Genomic DNA was obtained through phenol-chloroform-isoamyl alcohol extraction protocol [18].

5.3.3 Genome sequencing and assembly

Genome sequencing of the F1 strain was performed with the MiSEQ platform (Illumina®, USA). Reads were uploaded in FASTQ format to CLC Genomics Workbench 8 (Qiagen, USA) software, for trimming and assembly steps.

Reads with average phred scores below 30 and presenting any ambiguities were discarded. The last 10 nucleotides of each 3' end and reads smaller than 50 bp were also discarded. Overall, 2,524,734 reads were processed, resulting in 2,492,893 reads after trimming.

The F1 strain assembly *ab initio* resulted in 12 contigs obtained, with a N50 value of 304,012. The maximum length was 478,741 bp and the minimum 1,782 bp. A total of 11 gaps were filled by iterative reads mapping against gap margin of scaffold to complete the whole genome sequence.

In order to build the scaffold draft of the genome, reads were processed through CONTIGuator 2.7 software [19], utilizing the FNO01 (CP012153) strain as a reference. Remaining gaps were processed through CLC Genomics Workbench 8.02. During

the step, the reads were mapped multiple times lining up the reads with gap terminations all the way through the genome in pursue of overlapping regions that would merge contigs together. All the raw sequencing data was mapped and the absence of contamination with other genomes was confirmed by the coverage and the low number of unmapped reads (less than 0,1%).

5.3.4 Genome annotation

Automatic structural and functional annotation (to predict genes, rRNAs and tRNAs) of the genome was performed with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [20].

5.4 Genome Properties

The complete genome of F1 strain comprises a single circular chromosome 1,854,333 bp in length, with 32.26% G+C content, 7 rRNA operons, 37 tRNA genes and 369 pseudogenes (Table 3). The following data was annotated for the F1 strain: 1,235 genes with functional prediction and 357 CDSs classified as hypothetical proteins by COG database [21]. The distribution of genes into the COG functional categories is presented in Table 4.

5.4.1 Extended insights

Phylogeny of the F1 strain was compared with other strains among genus *Francisella* downloaded from GenBank. Accession numbers and hosts are displayed at Table 3. This approach was performed in order to evaluate evolutionary relationship of *Francisella* spp. completed genomes deposited at GenBank, and also, to identify F1 vaccine candidates with reverse vaccinology methodology.

Phylogenetic analysis was performed using Gegenees V2.2.1 [15] with high accuracy, generating a heatmap (Figure 2) displaying the similarity in non-conserved regions among whole genome *Francisella* strains. The heatmap presented all *F. noatunensis subsp. orientalis* strains forming a cluster with high similarity, while the only *F. noatunensis subsp. noatunensis* (GenBank accession number CP022207) strain evaluated, displayed a relatively low similarity with the *F. noatunensis subsp. orientalis* cluster, similar with another study using the *Francisella* genus [22]. Also, *F. philomiragia* O_319_029 strain, presented a higher similarity with the *F. noatunensis subsp. noatunensis* strain, which indicate a close relationship between these species [22]. The results obtained from Gegenees were exported to SplitsTree4 V4.14.6 [23] aiming to create a phylogenetic tree using EqualAngleRoot and UPGMA methods. The phylogenetic tree (Figure 4A) give an overview of the phylogenetic distances within *Francisella* genus. In agreement with another study with whole genome analysis [24], species that infect fish, are grouped in a different cluster from species that infect mammals, suggesting an independent evolutionary path through host adaptation of these clusters. Noteworthy, *F. haliotica* DSM_23729, a Giant Abalone strain, was closer to the phylogenetic tree root, reinforcing the hypothesis that *Francisella* spp. species originated from a marine habitat [24]. Interestingly, *F. haliotica* DSM_23729, showed a closer ancestral relationship with human strains instead of marine species. An enlarged view of the *F. noatunensis* cluster (Figure 4B) presented a high similarity between strains isolated in Latin America and Toba_04 strain isolated from Indonesia. Although, the cluster branch displays Toba_04 slightly out of the group, which may allow us to speculate a possible path of dissemination/introduction of *F. noatunensis subsp. orientalis* in the American continent.

In order to find vaccine candidates, a reverse vaccinology approach was performed using the complete genome of F1 strain. All protein sequences of F1 genome were submitted to MEDpipe [25] analysis, classifying them as cytoplasmic (CYT), secreted (SEC), membrane (MEM) or potentially surface exposed (PSE) proteins. In total, 188 proteins were

assigned as MEM, 1053 CYT, 49 SEC and 158 PSE. Only SEC and PSE proteins are used in reverse vaccinology studies, due to the close contact with host immunity [26]. A MED (maturity epitope density) score analysis was performed to evaluate the immunogenicity of the protein candidates. One hundred of the highest MED scores proteins candidates were checked for protein conservation and presence of premature stop codons. Using BLASTP, protein sequences were compared with other proteins belonging to *F. noatunensis subsp. orientalis* and *F. noatunensis subsp. noatunensis* aiming to find conserved proteins intra and inter-subspecies. Protein candidates with more than 90% identity and 90% similarity were considered conserved. Premature stop codons were checked in the amino acid FASTA file of the F1 strain and, when present, these proteins were disposed from analysis. Only 8 SEC and 40 PSE proteins attended these criteria. The highest and lowest MED scores picked were 55.27 and 11.05 respectively. Selected protein sequences were included into a FASTA amino acid file and submitted to the webtool Vaxign [27], in order to verify adhesivity score of interactions between selected proteins and host. Among the 48 proteins submitted, only 5 had binding this score higher than 0.51 and have no similarity with human, mice and pig. In this study, proteins were also checked for similarity with *Oreochromis* spp. proteins, to avoid interactions with the host immunity, and none of them presented considerable coverage or identity with host proteins (more than 50%). The highest adhesivity scores were displayed by proteins classified as SEC (WP_014715657, locus_tag BMT43_RS09195 and WP_014714512, locus_tag BMT43_RS02215) and the remaining proteins were classified as PSE (Table 6). Only WP_014714564.1 (locus_tag BMT43_RS06520) and WP_014714557.1 (locus_tag BMT43_RS06560) were assigned to a protein function domain through Interproscan analysis (<https://www.ebi.ac.uk/interpro/search/>), which classified these proteins as MCE family protein and HlyD family secretion protein respectively. The MCE family proteins are related to survival in macrophages, and are reported in live vaccines of attenuated mutant strain, yielding in reduction of virulence in guinea pigs to *Mycobacterium tuberculosis* [28]. HlyD, alongside with other membrane proteins, form a transport apparatus for hemolysin secretion. Further studies argue that this system is suitable to development of vaccine heterologous antigens [29]

5.5 Conclusions

In this study, we present the complete genome of a *F. noatunensis subsp. orientalis* F1 strain isolated from an outbreak in the state of Sao Paulo, Brazil. Our results resemble other strains isolated in Brazil, due to similar count of RNA genes and CDS, mainly the high number of pseudogenes, which suggests genome decay in this species. The phylogenetic approach displayed a high similarity among species in the *F. noatunensis* cluster. Interestingly, Toba_04 strain is slightly distant from *F. noatunensis subsp. orientalis* American isolates, suggesting that the Indonesian strain was not directly “introduced” in the American continent. Noteworthy, *F. noatunensis subsp. noatunensis* strain was relatively out of this cluster, closer to a *F. philomiragia* isolate. The vaccinology study, demonstrated five potential proteins that may induce immune response to *F. noatunensis* both subspecies without affecting the host, suggesting the potential for vaccine development against francisellosis in fish (in tropical and cold regions).

Competing interests

The authors declare that they have no competing interests

Funding

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We would also like to acknowledge the support provided by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Authors' contributions

CTF, RTC and UPP drafted the manuscript. LAA and UPP ran bacteriology assays and isolation of the strain. GWDS performed histopathology exams and analyses of results. LJB and SSC helped in the development of phylogenetic tree and vaccinology. DDG, LJB, SSC and UPP provided insights on the manuscript. ALC, EK, AJ and VA worked on library preparation, sequencing and assembly. UPP, RTC and CTF annotated the genome. All authors read and approved the final manuscript.

References

1. Colquhoun DJ, Duodu S. *Francisella* infections in farmed and wild aquatic organisms. Vet Res [Internet]. BioMed Central Ltd. 2011; 42–47.
2. Soto E, Hawke JP, Fernandez D, Morales J. *Francisella* sp., an emerging pathogen of tilapia, *Oreochromis niloticus* (L.), in Costa Rica. J Fish Dis. 2009; 32:713–22.
3. Vestvik N, Rønneseth A, Kalgraff CAK, Winther-Larsen HC, Wergeland HI, Haugland GT. *Francisella noatunensis* subsp. *noatunensis* replicates within Atlantic cod (*Gadus morhua* L.) leucocytes and inhibits respiratory burst activity. Fish Shellfish Immunol. 2013; 35:725–33.
4. Nguyen VV, Dong HT, Senapin S, Pirarat N, Rodkhum C. *Francisella noatunensis* subsp. *orientalis*, an emerging bacterial pathogen affecting cultured red tilapia (*Oreochromis* sp.) in Thailand. Aquac Res. 2016; 47:3697–702.
5. Leal CAG, Tavares GC, Figueiredo HCP. Outbreaks and genetic diversity of *Francisella noatunensis* subsp. *orientalis* isolated from farm-raised Nile tilapia (*Oreochromis niloticus*) in Brazil. Genet Mol Res. 2014; 13:5704–12.
6. Soto E, Illanes O, Hilchie D, Morales JA, Sunyakumthorn P, Hawke JP. Molecular and immunohistochemical diagnosis of *Francisella noatunensis* subsp. *orientalis* from formalin-fixed, paraffin-embedded tissues. J Vet Diagnostic Investig. 2012; 24:840–5.
7. Chen SC, Tung MC, Chen SP, Tsai JF, Wang PC, Chen RS. Systematic granulomas caused by a rickettsia-like organism in Nile tilapia, *Oreochromis niloticus* (L.), from southern Taiwan. J Fish Dis. Wiley/Blackwell (10.1111). 1994; 17:591–9.
8. Ottem KF, Nylund A, Isaksen TE, Karlsbakk E, Bergh Ø. Occurrence of *Francisella piscicida* in farmed and wild Atlantic cod, *Gadus morhua* L., in Norway. J Fish Dis. 2008; 31:525–34.
9. Jeffery K, Stone D, Feist S, Verner-Jeffreys D. An outbreak of disease caused by *Francisella* sp. in Nile tilapia *Oreochromis niloticus* at a recirculation fish farm in the UK. Dis Aquat Organ. 2010; 91:161–5.
10. Soto E, Baumgartner W, Wiles J, Hawke JP. *Francisella asiatica* as the causative agent of piscine francisellosis in cultured tilapia (*Oreochromis* sp.) in the United States. J Vet Diagnostic Investig. 2011; 23:821–5.
11. Jatobá A., Klipp SP., Hoppe R. Primeiro relato de *Francisella noatunensis* subespécie *orientalis* no sul do Brasil - relato de caso. Acta Vet Bras. 2016; 10:172–6.
12. Roberts RJ. Fish Pathology. In: Roberts RJ. The Bacteriology of Teleosts. 4th ed. Edinburgh: W. B. Saunders, 2012; p. 339-382.
13. Birkbeck TH, Feist SW, Verner-Jeffreys DW. *Francisella* infections in fish and shellfish. J Fish Dis. 2011; 34:173–87.
14. Rohmer L, Fong C, Abmayr S, Wasnick M, Larson Freeman T, Radey M. Comparison of *Francisella tularensis* genomes reveals evolutionary events associated with the emergence of human pathogenic strains. Genome Biol. 2007; 8:102.
15. Agren J, Sundström A, Håfström T, Segerman B. Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. PLoS One. Public Library of Science; 2012; 7(6):e39107.
16. Liolios K, Chen I-MA, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res. Oxford University Press; 2010; 38:D346-54.

17. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P. The minimum information about a genome sequences (MIGS) specification. *Nat Biotechnol.* 2008; 26:541–7.
18. Bollet C, Gevaudan MJ, de Lamballerie X, Zandotti C, de Micco P. A simple method for the isolation of chromosomal DNA from Gram positive or acid-fast bacteria. *Nucleic Acids Res.* 1991; 19:1955.
19. Galardini M, Biondi EG, Bazzicalupo M, Mengoni A. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol Med.* 2011; 6:11.
20. Tatusov T, Dicuccio M, Badretdin A, Ciuffo S, Li W. Prokaryotic Genome Annotation Pipeline. 2002; 1–9.
21. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 2000; 28:33–6.
22. Gonçalves LA, Soares SC, Pereira FL, Dorella FA, de Carvalho AF, Almeida GMF. Complete genome sequences of *Francisella noatunensis* subsp. *orientalis* strains FNO12, FNO24 and FNO190: A fish pathogen with genomic clonal behavior. *Stand Genomic Sci.* 2016; 11:30.
23. Huson DH. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics.* 1998; 14:68–73.
24. Sjödin A, Svensson K, Ohrman C, Ahlinder J, Lindgren P, Duodu S. Genome characterisation of the genus *Francisella* reveals insight into similar evolutionary paths in pathogens of mammals and fish. *BMC Genomics.* 2012; 13:268.
25. Santos AR, Pereira VB, Barbosa E, Baumbach J, Pauling J, Röttger R. Mature Epitope Density—a strategy for target selection based on immunoinformatics and exported prokaryotic proteins. *BMC Genomics.* BioMed Central; 2013; 14 (Suppl 6):S4.
26. Rappuoli R. Reverse vaccinology, a genome-based approach to vaccine development. *Vaccine.* 2001; 19:2688–91.
27. He Y, Xiang Z, Mobley HLT. Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol.* Hindawi Limited; 2010:297505.
28. Obregón-Henao A, Shanley C, Bianco MV, Cataldi AA, Basaraba RJ, Orme IM. Vaccination of guinea pigs using mce operon mutants of *Mycobacterium tuberculosis*. *Vaccine.* NIH Public Access; 2011; 29(26):4302–7.
29. Gentschev I, Dietrich G, Goebel W. The *E. coli* α -hemolysin secretion system and its use in vaccine development. *Trends Microbiol.* Elsevier Current Trends; 2002; 10:39–45.
30. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA.* 1990; 87:4576–9.
31. Garrity GM, Bell JA, Lilburn TG. Phylum XIV. Proteobacteria phyl. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, Volume 2. 2nd edition, Part B. New York: Springer; 2005. p. 1.
32. Garrity GM, Bell JA, Lilburn TG. Class III. Gammaproteobacteria class. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, Volume 2. 2nd edition, Part B. New York: Springer; 2005. p. 1.
33. Garrity GM, Bell JA, Lilburn TG. Order V. Thiotrichales ord. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, Volume 2. 2nd edition, Part B. New York: Springer; 2005. p. 1.
34. Sjöstedt AB. Family III. Francisellaceae fam. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. *Bergey's Manual of Systematic Bacteriology*, Volume 2. 2nd edition, Part B. New York: Springer; 2005; p. 199–200.
35. Skerman V, McGowan V, Sneath P. Approved Lists of Bacterial Names (Amended). *Approv. List. Bact. Names.* ASM Press; 1989.
36. Ottem KF, Nylund A, Karlsbakk E, Friis-Møller A, Kamaishi T. Elevation of *Francisella*

philomiragia subsp. *noatunensis* Mikalsen *et al.* (2007) to *Francisella noatunensis* comb. nov. [syn. *Francisella piscicida* Ottem *et al.* (2008) syn. nov.] and characterization of *Francisella noatunensis* subsp. *orientalis* subsp. nov., two important fish pathogens. *J Appl Microbiol.* 2009; 106:1231–43.

37. Soto E, Abrams S, Revan F. Effects of temperature and salt concentration on *Francisella noatunensis* subsp. *orientalis* infections in Nile tilapia *Oreochromis niloticus*. *Dis Aquat Organ.* 2012; 101:217–23.

Table 1. Classification and general features of Fno strain F1 (FIELD et al., 2008)

MIGS ID	Property	Term	Evidence code
	Classification	Domain <i>Bacteria</i>	TAS [30]
		Phylum <i>Proteobacteria</i>	TAS [31]
		Class <i>Gammaproteobacteria</i>	TAS [32]
		Order <i>Thiotrichales</i>	TAS [33]
		Family <i>Francisellaceae</i>	TAS [34]
		Genus <i>Francisella</i>	TAS [35]
		Species <i>Francisella noatunensis subsp. orientalis</i>	TAS [36]
		Strain F1	IDA
	Gram stain	Gram-negative	TAS [36]
	Cell shape	Pleomorphic	TAS [36]
	Motility	non-motile	TAS [36]
	Sporulation	Not reported	NAS
	Temperature range	Mesophilic (15–34 °C)	TAS [36]
	Optimum temperature	<25 °C	TAS [36]
	pH range;	Not reported	NAS
	Optimum Carbon source	Not reported	NAS
MIGS-6	Habitat	Host (Nile tilapia)	NAS
MIGS-6.3	Salinity	Survive different salinities (0 to 35 %)	TAS [37]
MIGS-22	Oxygen requirement	Strictly aerobic	TAS [36]
MIGS-15	Biotic relationship	Intracellular facultative pathogen	TAS [13]
MIGS-14	Pathogenicity	Pathogenic for fish	TAS [13]
MIGS-4	Geographic location	Brazil, State of São Paulo,	NAS
MIGS-5	Sample collection time	2015	NAS
MIGS-4.1	Latitude	not reported	NAS
MIGS-4.2	Longitude	not reported	NAS
MIGS-4.3	Depth	not reported	NAS
MIGS-4.4	Altitude	not reported	NAS

^aEvidence codes – IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project.

Table 2. Genome sequencing project information.

MIGS ID	Property	Term
MIGS-31	Finishing quality	Finished
MIGS-28	Libraries used	One paired-and library (mean size 300 bp, DNA insert size of ~200bp)
MIGS-29	Sequencing platforms	Illumina MiSeq
MIGS-31.2	Fold coverage	146.0x
MIGS-30	Assemblers	CLC genomics workbench v8.02
MIGS-32	Gene calling method	NCBI Prokaryotic Genome Annotation Pipeline
	Locus Tag	BMT43
	Genbank ID	CP018051.1
	Genbank Date of Release	November 22, 2016
	GOLD ID	Gp0177338
	BIOPROJECT	PRJNA352676
MIGS 13	Source Material Identifier	F1
	Project relevance	fish pathogen

Table 3. Genomes used in phylogeny analysis

Species	Strain	Accession	Size (Mb)	GC%	Genes	Protein	Host	Country
<i>F. noatunensis</i> subsp. <i>noatunensis</i>	FSC772	CP022207.1	1.93382	32.7	1891	1754	Atlantic Salmon	Chile
<i>F. noatunensis</i> subsp. <i>orientalis</i>	F1	CP018051.1	1.85433	32.3	1889	1448	Tilapia	Brazil
<i>F. noatunensis</i> subsp. <i>orientalis</i>	FNO01	CP012153.2	1.86244	32.3	1900	1452	Tilapia	Brazil
<i>F. noatunensis</i> subsp. <i>orientalis</i>	FNO12	CP011921.2	1.86221	32.3	1899	1446	Tilapia	Brazil
<i>F. noatunensis</i> subsp. <i>orientalis</i>	FNO190	CP011923.2	1.86221	32.3	1900	1451	Tilapia	Brazil
<i>F. noatunensis</i> subsp. <i>orientalis</i>	FNO24	CP011922.2	1.86232	32.3	1899	1446	Tilapia	Brazil
<i>F. noatunensis</i> subsp. <i>orientalis</i>	LADL--07-285A	CP006875.1	1.85899	32.3	1896	1422	Tilapia	Costa Rica
<i>F. noatunensis</i> subsp. <i>orientalis</i>	Toba 04	CP003402.1	1.8472	32.2	1886	1418	Tilapia	Indonesia
<i>F. halioticida</i>	DSM_23729	CP022132.1	2.19743	31.2	2351	2096	Giant Abalone	Japan
<i>F. tularensis</i> subsp. <i>tularensis</i>	Schu_S4	NC_006570	1.89278	32.3	1834	1556	Human	Human
<i>F. persica</i>	FSC845	CP012505.1	1.51668	31.4	1485	1067	Tick	Egypt
<i>F. philomiragia</i>	O_319_029	CP009343.1	2.04493	32.6	1986	1876	Water	United States

Table 4. Genome statistics

Attribute	Value	% of Total^a
Genome size (bp)	1,854,333	100
DNA coding region (bp)	1,329,162	71.67
DNA G+C content (bp)	601,242	32.26
DNA scaffolds	1	100
Total genes	1,857	100
Protein-coding genes	1,440	77.54
RNA genes	48	2.58
Pseudo genes	369	19.87
Genes with function prediction	1,235	66.5
Genes assigned to COGs	1,331	71.67
Genes assigned Pfam domains	1,572	84.65
Genes with signal peptides	96	5.17
Genes with transmembrane hélices	585	31.5
CRISPR repeats	0	0

^aThe total is based on either size of genome in base pairs or the genes in the annotate genome

Table 5. Number of genes associated with general COG functional categories.

Code	Value	%age	Description
J	143	9.93	Translation, ribosomal structure and biogenesis
A	1	0.07	RNA processing and modification
K	52	3.61	Transcription
L	71	4.93	Replication, recombination and repair
B	0	0.0	Chromatin structure and dynamics
D	18	1.25	Cell cycle control, Cell division, chromosome partitioning
V	14	0.97	Defense mechanisms
T	17	1.18	Signal transduction mechanisms
M	115	7.98	Cell wall/membrane biogenesis
N	04	0.27	Cell motility
U	27	1.87	Intracellular trafficking and secretion
O	63	4.37	Posttranslational modification, protein turnover, chaperones
C	85	5.90	Energy production and conversion
G	75	5.20	Carbohydrate transport and metabolism
E	111	7.70	Amino acid transport and metabolism
F	63	4.37	Nucleotide transport and metabolism
H	80	5.55	Coenzyme transport and metabolism
I	65	4.51	Lipid transport and metabolism
P	43	2.98	Inorganic ion transport and metabolism
Q	39	2.70	Secondary metabolites biosynthesis, transport and catabolism
R	116	8.05	General function prediction only
S	96	6.66	Function unknown
-	526	36.52	Not in COGs

^aThe percentage is based on the total number of protein coding genes in the annotated genome

^bThe total does not correspond to the final quantity of CDSs for each genome because some genes are associated with more than one COG functional category

Table 6. Vaccine candidates

Protein Accession	Protein Note	Locus_tag	Adhesin Probability	Location
WP_014715657	DUF3568	BMT43_RS09195	0.701	SECRETED
WP_014714512	DUF3573	BMT43_RS02215	0.664	SECRETED
WP_012280666	MULTISPECIES: hypothetical protein	BMT43_RS07125	0.577	PSE
WP_014714564	MCE family protein	BMT43_RS06520	0.565	PSE
WP_014714557	HlyD family secretion protein	BMT43_RS06560	0.551	PSE

^aDUF – Domain of Unknown Function

Figure legends

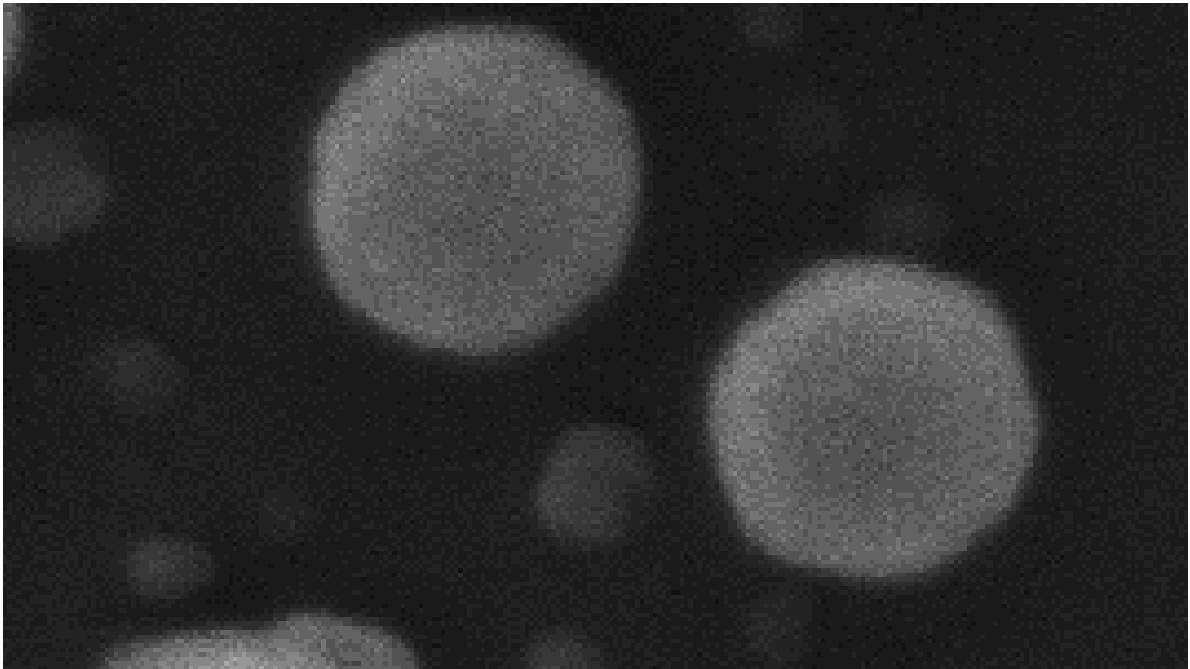


Figure 1. Photomicrograph of the *F. noatunensis* subsp. *orientalis* F1 strain.

Organism	1	2	3	4	5	6	7	8	9	10	11	12
1: <i>F. halioticida</i> DSM 23729	100.0	3.95	3.31	4.46	4.47	4.27	4.27	4.27	4.27	4.27	4.27	4.27
2: <i>F. tularensis</i> subsp <i>tularensis</i> SCHU S4	4.6	100.0	25.68	14.04	13.23	11.2	11.21	11.21	11.21	11.21	11.21	11.2
3: <i>F. persica</i> FSC845	4.55	31.74	100.0	6.54	6.4	6.05	6.05	6.05	6.05	6.05	6.05	6.04
4: <i>F. philomiragia</i> O_319_029	4.82	12.54	5.06	100.0	80.08	59.87	59.86	59.88	59.88	59.88	59.87	59.8
5: <i>F. noatunensis</i> subsp <i>noatunensis</i> FSC772	5.03	12.95	5.23	84.47	100.0	63.41	63.4	63.4	63.41	63.41	63.41	63.35
6: <i>F. noatunensis</i> subsp <i>orientalis</i> FNO24	4.98	11.36	5.09	65.41	65.78	100.0	99.96	99.96	99.98	99.84	99.85	99.86
7: <i>F. noatunensis</i> subsp <i>orientalis</i> FNO190	4.93	11.46	5.13	65.52	65.76	99.96	100.0	99.96	99.98	99.85	99.86	99.86
8: <i>F. noatunensis</i> subsp <i>orientalis</i> FNO01	4.94	11.38	5.12	65.36	65.73	99.96	99.95	100.0	99.97	99.84	99.85	99.86
9: <i>F. noatunensis</i> subsp <i>orientalis</i> FNO12	4.91	11.39	5.1	65.39	65.78	99.98	99.98	99.97	100.0	99.87	99.87	99.88
10: <i>F. noatunensis</i> subsp <i>orientalis</i> LADL-07-285A	4.86	11.33	5.07	65.5	65.83	99.96	99.95	99.95	99.97	100.0	99.96	99.87
11: <i>F. noatunensis</i> subsp <i>orientalis</i> F1	4.69	11.23	4.93	65.42	65.72	99.97	99.97	99.97	99.98	99.97	100.0	99.89
12: <i>F. noatunensis</i> subsp <i>orientalis</i> Toba04	4.48	10.96	4.68	65.3	65.67	99.89	99.89	99.88	99.91	99.79	99.79	100.0

Figure 2. Heatmap based on whole genome of selected *Francisella* spp.
Heatmap of similarity among non-core regions of *Francisella* spp. genomes using Gegenees
all in all fragmented comparison with high accuracy.

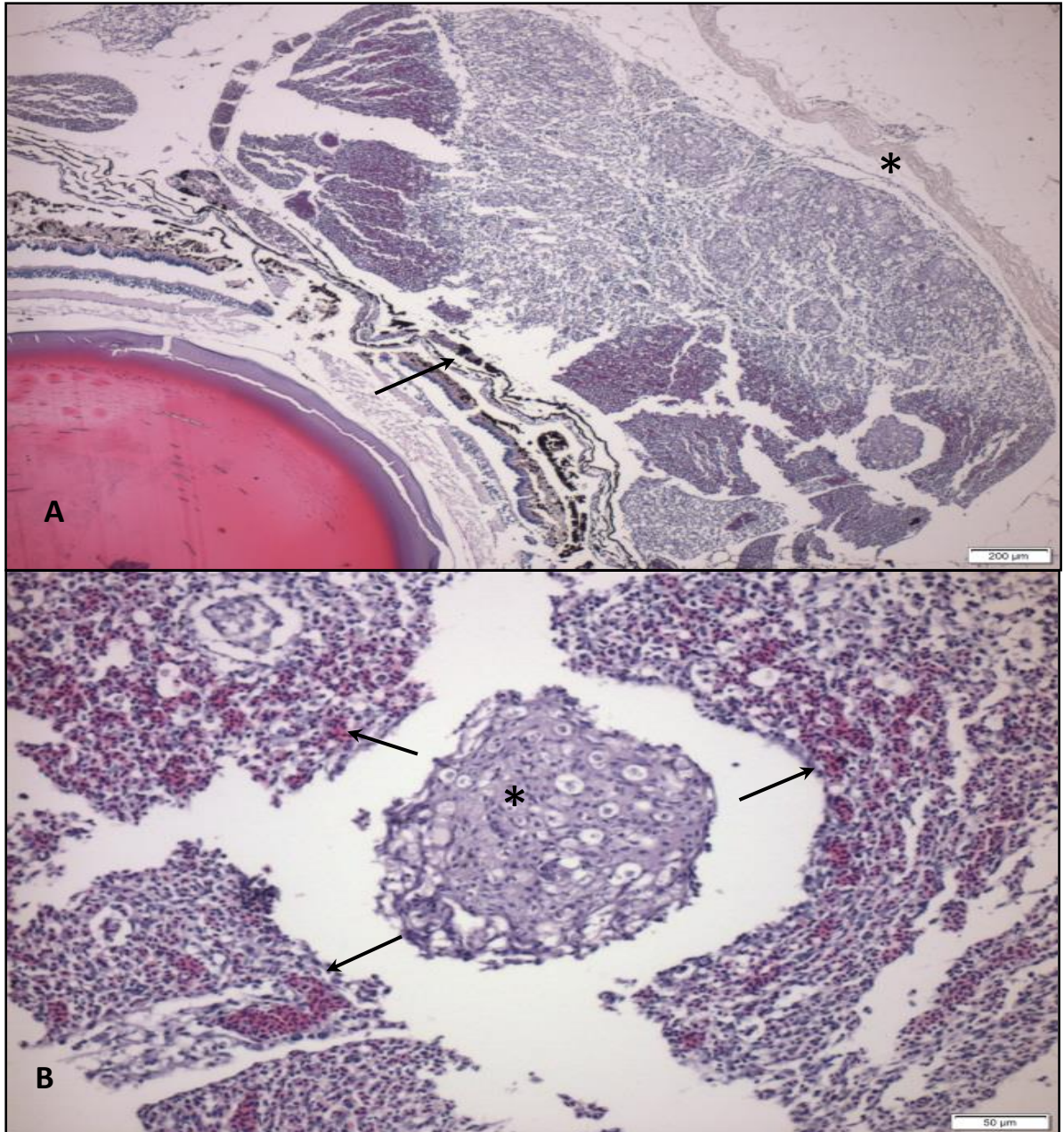


Figure 3: Ocular globe of Nile tilapia (*Oreochromis niloticus*) infected by *F. noatunensis* subsp. *orientalis* strain F1.

A) Presence of marked nodular inflammatory reaction between choroid (arrow) and sclera (*). Hematoxylin and Eosin, 4x. B) Accentuated accumulation of polymorphonuclear and mononuclear inflammatory cells in well vascularized stroma (arrows) with granulomatous formation (*). Hematoxylin and Eosin, 20x. C) Granulomatous area composed of vacuolated macrophages and epithelioid cells. Hematoxylin and Eosin, 40x. D) Granulomatous portion evidencing the presence of intracytoplasmic cocoid bacteria (arrows). Hematoxylin and Eosin, 100x.

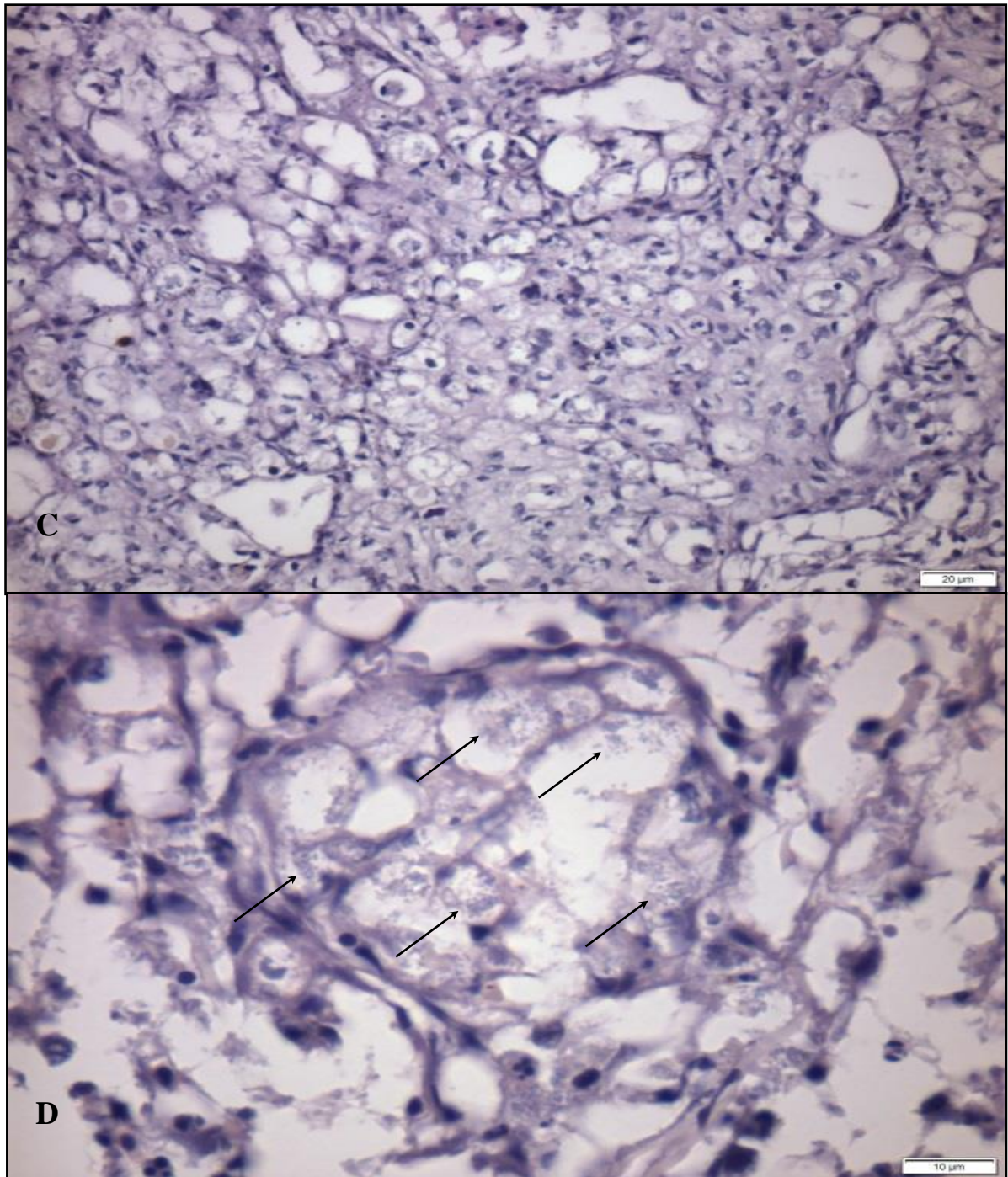
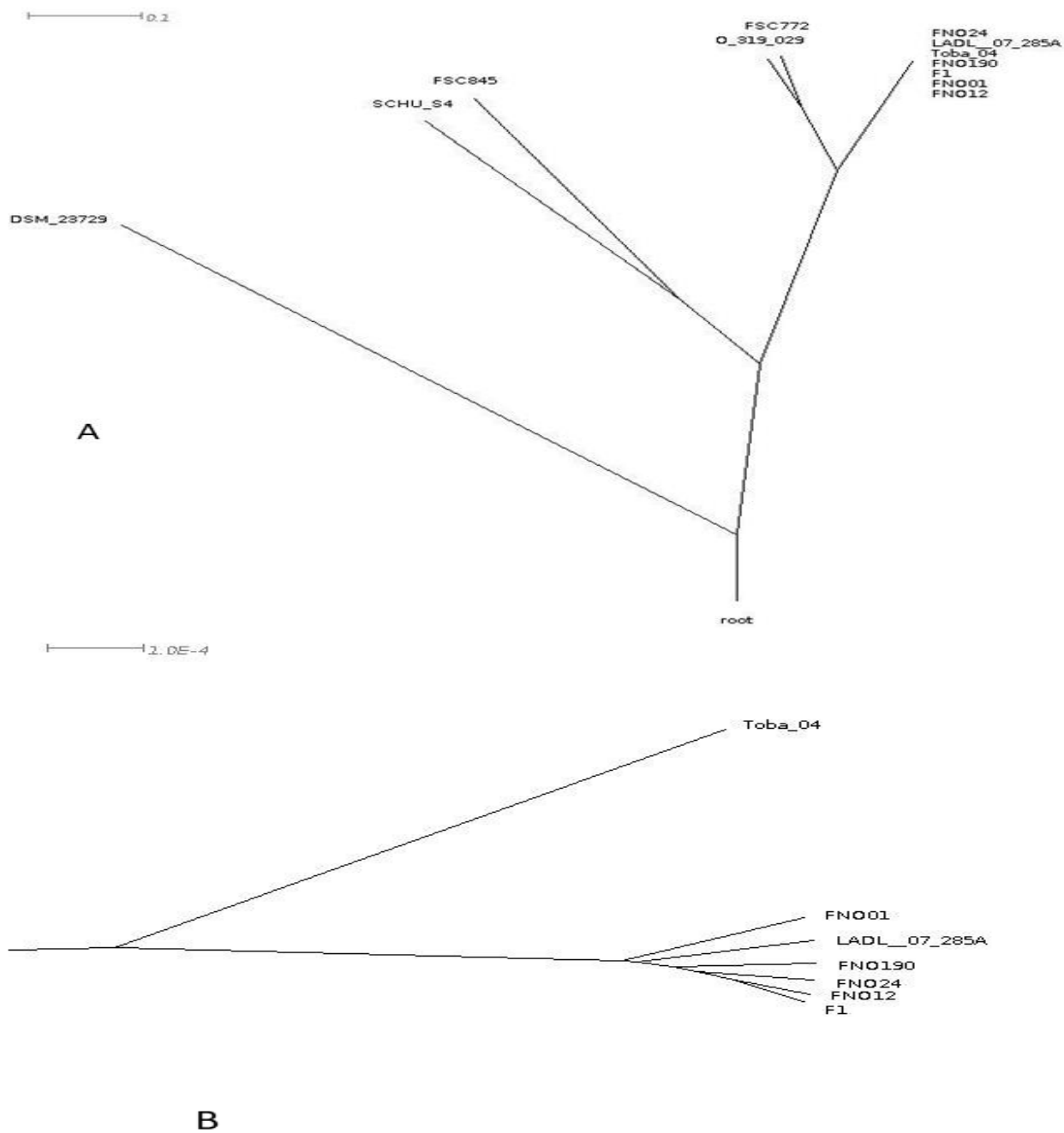


Figure 4: Phylogenetic tree based on whole genome of *Francisella* spp. strains using EqualAngleRoot and UPGMA methods. A) Phylogenetic tree comparing different species in the *Francisella* genus evidencing the distance of strains. B) A zoomed view of the Fno cluster.



6 CONCLUSÃO

- Foi possível concluir a montagem do genoma completo da cepa *F. noatunensis* subsp. *orientalis* (Genbank ID: CP018051.1).
- O genoma completo da cepa *F. noatunensis* subsp. *orientalis* F1 apresentou alta similaridade com cepas isoladas na América Latina. Uma menor similaridade foi identificada com uma cepa proveniente da Indonésia.
- As proteínas WP_014715657, WP_014714512, WP_012280666, WP_014714564 e WP_014714557 identificadas e analisadas *in silico* são bons candidatos à vacina de subunidade contra as subespécies *F. noatunensis* subsp. *orientalis* e *F. noatunensis* subsp. *noatunensis*.