

## PROTEÍNAS DE SUPERFÍCIE DE *LEPTOSPIRA INTERROGANS* IDENTIFICADAS POR FERRAMENTAS *IN SILICO*

JÚLIA COUGO DOS SANTOS<sup>1</sup>; ANDRÉ ALEX GRASSMANN<sup>2</sup>; MAURÍCIO TAVARES TAMBORINDEGUY<sup>2</sup>; CAROLINE AMURIM DA SILVA GONÇALVES<sup>2</sup>; FREDERICO SCHMITT KREMER<sup>3</sup>; ALAN JOHN ALEXANDER MCBRIDE<sup>4</sup>

*Laboratório de Pesquisa em Doenças Infecciosas, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas – juliapetrarca@gmail.com*

<sup>2</sup> *Laboratório de Pesquisa em Doenças Infecciosas, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas – grassmann.aa@gmail.com; mauriciotamborindeguy@gmail.com; carolamurim@gmail.com*

<sup>3</sup> *Laboratório de Bioinformática, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas – fred.s.kremer@hotmail.com*

<sup>4</sup> *Laboratório de Pesquisa em Doenças Infecciosas, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas – alan.mcbride@ufpel.edu.br*

### 1. INTRODUÇÃO

A leptospirose é uma doença infecciosa causada por espiroquetas patogênicas pertencentes ao gênero *Leptospira* (MCBRIDE et al., 2005). Os seres humanos são infectados através da exposição direta com animais carreadores dessa bactéria ou através de contato indireto com água ou solo contaminados (SAMIR et al, 2015). Os sinais clínicos variam desde uma doença semelhante à gripe até uma infecção grave com falência de diversos órgãos (ADLER; DE LA PEÑA MOCTEZUMA, 2010). É a zoonose mais difundida no mundo (ADLER et al., 2010). A estimativa é de 873 mil casos de Leptospirose grave e 49 mil óbitos ocorram a cada ano (PICARCEAU et al.; 2014). Atualmente as vacinas disponíveis, com base nas bactérias inteiras inativadas são de baixa eficácia. O foco atual no desenvolvimento de vacina contra Leptospirose é descobrir抗ígenos que induzam proteção a longo prazo, sem efeitos colaterais, contra os mais de 300 sorovares de *Leptospira* spp. patogênica. Novas estratégias para busca de novos e promissores抗ígenos para desenvolvimento de vacinas são necessárias para a prevenção dessa zoonose (Dellagostin et. al 2011).

Dentre os pré-requisitos para desenvolvimento de uma vacina protetora contra doenças bacterianas, destaca-se os抗ígenos secretados, relacionados com algum mecanismo patogênico (como toxinas) e, principalmente os expostos na superfície da bactéria (SETUBAL et al .; 2006). Uma resposta imune montada contra um抗ígeno de superfície deve ser capaz de reconhecer a bactéria viva e elimina-la pelo ataque do sistema imune. Em bactérias didermes como a *Leptospira* spp. as proteínas normalmente expostas na superfície são lipoproteínas ancoradas no folheto externo da membrana externa ou proteínas integrais de membrana externa com estrutura de barril-beta, apresentando uma porção exposta (SETUBAL et al .; 2006). Devido a características estruturais, a predição de localização subcelular de lipoproteínas *in silico* só é possível por análise de homologia (pouco confiável), já proteínas barril beta expostas na superfície possuem uma estrutura secundária extremamente característica, facilmente identificada por softwares específicos (SETUBAL et al .; 2006). Neste trabalho, aplicamos a vacinologia reversa baseada na busca de proteínas barril-beta integrais de membrana, expostas na superfície ou secretadas utilizando uma série de software disponíveis online, a partir do genoma de *L. interrogans*.

### 2. METODOLOGIA

A sequência completa do genoma de *L. interrogans* sorogrupo Icterohaemorrhagiae sorovar Copenhageni cepa Fiocruz L1-130 foi obtida do GenBank (número de acesso AE016823 e AE016824, respectivamente para os cromossomos I e II) e as sequências de proteína correspondentes a este genoma foram obtidas em um arquivo FASTA, utilizando o software Artemis (CARVER et al., 2012). Em seguida foi realizada a análise do proteoma predito a partir da sequencia do genoma, utilizando ferramentas computacionais escolhidas de acordo com as características importantes de um antígeno para o desenvolvimento de vacina contra leptospirose.

Os softwares utilizados foram: Cello v.2.5, Psort-B v.3.0.2 e Gneg-PLoc v.2.0 para predição de localização subcelular; Bomp, MCMBB, TMBETADISC-RBF e Hhomp 1.0 para predição de proteínas com estrutura barril-beta; Phobius, THmm v 2.0, Hmmtop v 2.0 e Mensat2 para proteínas com a estrutura secundária α-hélice transmembrana (TMH); PrediSi, SignalP v.4.1 e SignalCF para predição de proteínas com peptídeo sinal (PS). O resultado dos softwares de localização subcelular foi filtrado de forma a remover da lista proteínas com TMH (proteínas de membrana interna), mantendo proteínas com PS. A partir desta análise, originou-se listas de proteínas: uma de proteínas preditas como proteína de membrana externa (OMP) com a estrutura de barril beta e outra de OMP sem barril beta, além da lista de proteínas secretadas.

Os softwares LipoP v.1.0 e Splip foram utilizados para predição de lipoproteínas, cujo resultado foi filtrado com a lista de proteínas preditas como OMP pelos softwares de predição de localização, tomando por base homologia entre lipoproteínas de OMP de outras bactérias didermes e as preditas para *Leptospira* spp.

Todo o proteoma predito de *L. interrogans* foi analisado em cada um dos softwares, individualmente e a análise foi facilitada por um script escrito em linguagem Python (FRIEDRICH et al., 2014), que teve como objetivo, organizar os diferentes resultados dos softwares e transferi-los para um arquivo CSV e XLS com os dados de cada proteína. Esses dados foram analisados e filtrados entre si quando necessário utilizando Microsoft Excel 2013.

### 3. RESULTADOS E DISCUSSÃO

A ferramenta Artemis encontrou 3.659 proteínas codificadas no genoma. A análise pelo software Cello identificou 132 proteínas extracelulares e 399 de membrana externa; Psort identificou 62 proteínas como de membrana externa e 76 como extracelular; Gneg-PLoc identificou 78 como de membrana externa e 143 como extracelular. A busca por OMPs com barril-beta identificou em cada software: Bomp, 79; MCMBB, 182; TMBETADISC, 1041, e HHOMP, 69. Phobius encontrou 2827 proteínas, THMM 2735, HMMTOP 1894 e Mensat 2487 como proteínas não contendo α-hélice transmembrana (não ancoradas na membrana interna). O software PREDISI identificou 745, SignalP 111, SignalCF 910 como proteínas contendo peptídeo sinal. O LipoP identificou 164 e o SPLIP 175 lipoproteínas.

Combinando os resultados dos softwares de localização celular (Cello, Psort e GengmPloc) obtivemos uma lista de proteínas contidas em todos os três softwares e em pelo menos um deles. Sendo 25 proteínas de membrana externa e 7 extracelulares em todos os programas; 439 proteínas de membrana externa e 278 extracelulares em pelo menos um software. Os resultados dos softwares para identificação de proteínas com TMH (Phobius, THMM, HMMTOP e Mensat) quando combinados apresentaram 2001 proteínas em todos e 3227 em pelo

menos um desses programas, preditas como sem TMH. Foram identificadas 72 proteínas em todos os softwares e 1202 em pelo menos um dos softwares que buscam peptídeo sinal (SignalCF, SignalP e PREDISI). Nos preditores de proteínas com barril-beta (BOMP, MCMBB, HHOMP e TMBETADISC) tivemos 20 proteínas em todos os softwares e 1095 em pelo menos um destes softwares. Ao total, 24 proteínas foram identificadas por todos os softwares como proteína sem TMH e contendo PS, e 849 em pelo menos um dos softwares para estas características. Nenhuma proteína predita como OMP ou proteína extracelular foi predita considerando o resultado em comum entre todos os softwares incluindo a lista de proteínas sem TMH e com PS. Quando consideramos proteínas preditas em pelo menos um software foram preditas 203 proteínas de membrana externa, sem TMH e com PS, assim como, 152 proteínas extracelulares sem TMH e com PS.

Ao total, encontramos 141 proteínas preditas como presentes na membrana externa, sem TMH, com PS e apresentando a estrutura secundária de barril beta, identificadas por pelo menos um software. Desses 141 proteínas 15 são preditas em todos os softwares de estrutura barril beta, 40 pelo HHOMP e 42 pelo Psort como de membrana externa, 19 como contendo peptídeo sinal pelo SignalP e 66 sendo proteínas sem α-hélice. Combinando esses resultados 5 proteínas estão contidas em todas essas listas, sendo elas: LIC10714, anotada como uma proteína de membrana externa; LIC10896, um componente do complexo de transporte da membrana externa TonB; LIC11268, anotada como proteína hipotética; LIC12374 outro componente TonB; e LIC13417, outra proteína hipotética.

Analizando os resultados combinados, não encontramos lipoproteínas que foram preditas como presentes na membrana externa por todos os softwares e 76 em pelo menos um para localização. Apesar da dificuldade para predição correta de lipoproteínas, estes softwares consideram homologia entre essas lipoproteínas de leptospiras e lipoproteínas de Gram-negativas, já validadas como componentes de membrana externa.

#### 4. CONCLUSÕES

A análise do proteoma predito para *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130 possibilitou a identificação de 141 proteínas preditas como proteínas integrais de membrana externa com conformação de barril beta; 152 proteínas secretas e 76 lipoproteínas preditas como presentes na membrana externa. Das 141 OMPs com barril beta, 5 são destaque por terem sido identificadas por softwares confiáveis. Em seguida, faremos uma análise mais detalhada sobre a possível função predita para estas proteínas, modelagem estrutural para buscar porções expostas das mesmas, busca por epítópos imunogênicos e confirmação experimental da localização subcelular destas proteínas. As proteínas expostas na superfície de *Leptospira* spp. serão utilizadas em experimentos para desenvolvimento de vacinas recombinantes contra leptospirose.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

MCBRIDE, A. J.; ATHANAZIO, D. A.; REIS, M. G. e KO, A. I. Leptospirosis. *Curr Opin Infect Dis*, v.18, n.5, p. 376-86. 2005.

PICARDEAU, M.; BERTHERAT, E.; JANCLOES, M.; SKOULLOUDIS, A. N.; DURSKI, K. e HARTSKEERL, R. A. Rapid tests for diagnosis of leptospirosis: current tools and emerging technologies. **Diagn Microbiol Infect Dis**, v.78, n.1, p. 1-8. 2014.

FRIEDRICH, P.; VELLA, M.; GULYÁS A. I.; FREUND T. F.; KÁLI. S. A flexible, interactive software tool for fitting the parameters of neuronal models. **Frontiers in Neuroinformatics**, v.8, n.63.2014

CARVER, T.; HARRIS, S. R.; BERRIMAN, M.; PARKHILL, J. e MCQUILLAN, J. A. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. **Bioinformatics**, v.28, n.4, p. 464-9. 2012.

SAMIR, A.; SOLIMAN, R.; EL-HARIRI, M.; ABDEL-MOEIN, K.; HATEM, M.E. Leptospirosis in animals and human contacts in Egypt: broad range surveillance. **Revista Sociedade Brasileira de Medicina Tropical**, v.48, n.3, p. 272-277.2015.

ADLER, B. e DE LA PENA MOCTEZUMA, A. Leptospira and leptospirosis. **Vet Microbiol**, v.140, n.3-4, p. 287-96. 2010.

DELLAGOSTIN, O. A.; GRASSMANN, A. A.; HARTWIG, D. D.; FELIX, S. R.; DA SILVA, E. F. e MCBRIDE, A. J. Recombinant vaccines against Leptospirosis. **Hum Vaccin**, v.7, n.11, p. 1215-24. 2011

SETUBAL, JC, REIS, M.; MATSUGANA J.; HAAKE D.A. Lipoprotein computational prediction in spirochaetal genomes. **Microbiology**. P.113-121. 2006.