

## Nova abordagem genómica para a investigação da sífilis

### Novel genomic approach to decode syphilis

Miguel Pinto<sup>1</sup>, Vítor Borges<sup>1</sup>, Minia Antelo<sup>1</sup>, Miguel Pinheiro<sup>2</sup>, Alexandra Nunes<sup>1</sup>, Jacinta Azevedo<sup>3</sup>, Maria José Borrego<sup>4</sup>, Joana Mendonça<sup>5</sup>, Dina Carpinteiro<sup>5</sup>, Luís Vieira<sup>5</sup>, João Paulo Gomes<sup>1</sup>

[miguel.pinto@insa.min-saude.pt](mailto:miguel.pinto@insa.min-saude.pt)

(1) Núcleo de Bioinformática. Departamento de Doenças Infecciosas, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal.

(2) Instituto de Biomedicina. Departamento de Ciências Médicas, Universidade de Aveiro, Aveiro, Portugal.

(3) Clínica de Doenças Sexualmente Transmissíveis. Centro de Saúde da Lapa, Agrupamento de Centros de Saúde Lisboa Central, Lisboa, Portugal.

(4) Laboratório de Referência de Infecções Bacterianas Sexualmente Transmissíveis. Departamento de Doenças Infecciosas, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal.

(5) Unidade de Tecnologia e Inovação. Departamento de Genética Humana, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa, Portugal.

### \_Resumo

A sífilis é uma doença sexualmente transmissível causada pela bactéria *Treponema pallidum* e constitui um problema de saúde pública mundial, em parte devido à ausência de uma vacina para prevenção da sua transmissão. A investigação desta doença tem sido atrasada pela incapacidade histórica de cultivar *T. pallidum* *in vitro*, dificultando por exemplo o desenvolvimento de estudos genómicos. De facto, há uma grande lacuna no conhecimento da epidemiologia molecular deste agente patogénico, assim como da base molecular que medeia a patogénese da sífilis. No estudo aqui apresentado, foi possível implementar uma abordagem inovadora para capturar o genoma de *T. pallidum* no contexto de infeção humana, evitando-se, assim, a necessidade da cultura da bactéria em modelo animal. Esta estratégia permitiu estudar, pela primeira vez, como é que este agente patogénico vai alterando o seu genoma para se adaptar e sobreviver como agente infeccioso humano. Nomeadamente permitiu descodificar os principais mecanismos genéticos pelos quais a bactéria *T. pallidum* evade o sistema imunitário e se adapta ao Homem nesta complexa e multifásica doença. A aplicação desta estratégia inovadora de monitorização da interação Homem-bactéria poderá ser importante para o desenvolvimento de novas medidas profiláticas e/ou terapêuticas. Acresce que esta abordagem constitui também um ponto de viragem para o aperfeiçoamento de metodologias de diagnóstico e de epidemiologia molecular, o que permitirá aumentar o conhecimento da distribuição geográfica, das vias de transmissão e das propriedades de virulência deste agente patogénico para bem da saúde pública.

### \_Abstract

*Syphilis, a sexually transmitted disease caused by the bacterium T. pallidum, remains a global problem with an estimated 6 million people infected each year, which is in part attributed to the absence of a vaccine to prevent infection and transmission. Despite its tremendous public health impact, research in syphilis has been considerable hampered due to the historical inability to culture T. pallidum in vitro, which has hampered, for instance, the acquisition of consistent genomic data. As such, there is a strong lack of knowledge on the molecular epidemiology of this important human pathogen as well as on the molecular mechanisms underlying syphilis pathogenesis. In the work presented here, we have bypassed the culture bottleneck by means of a targeted strategy never applied to uncultivable*

*bacterial human pathogens to directly capture whole-genome T. pallidum data in the context of human infection. This strategy allowed, for the first time, to understand how this pathogen shapes its genome towards adaptation and survival during syphilis. While this work demonstrates the exceptional power of monitoring the pathogen adaptability during human infection, it also provides critical data that may guide the development of novel treatments and prophylactic measures, such as a vaccine. In other perspective, it is anticipated that the implemented methodological approach constitutes a disruptive step towards the improvement of the current diagnostics and typing methodologies, which will enhance the knowledge on the geographic distribution of strains, its transmissibility and propensity to cause disease.*

### \_Introdução

A sífilis é uma doença sexualmente transmissível que constitui um problema de saúde pública a nível mundial, que tem sido parcialmente atribuído à ausência de uma vacina para prevenção da sua transmissão (1,2). A bactéria *Treponema pallidum* é o agente etiológico responsável pela sífilis, cujo quadro clínico manifesta-se por lesão no local de contacto (sífilis primária), lesões cutâneas típicas após disseminação na corrente sanguínea (sífilis secundária), bem como quadros clínicos mais preocupantes como sífilis terciária (através, essencialmente, da infeção do sistema nervoso central) ou doença congénita (1). A par da importância da sífilis em saúde pública, a incapacidade histórica de cultivar *T. pallidum* *in vitro* (a propagação laboratorial só é possível com o recurso à inoculação em coelho) tem atrasado a investigação neste campo dificultando, por exemplo, a elaboração de estudos

genómicos. De facto, há uma grande lacuna no conhecimento da epidemiologia molecular deste agente etiológico, patogénico do Homem, particularmente no que respeita à distribuição geográfica e potencial emergência de estirpes *T. pallidum* resistentes aos antibióticos.

Desconhece-se também a base molecular que medeia a patogénese da sífilis e que está na base das diferentes manifestações clínicas que esta doença apresenta. Por exemplo, permanece por esclarecer como é que a bactéria consegue evadir o sistema imunitário após a lesão primária, invadindo e persistindo na corrente sanguínea e também ultrapassar as barreiras hematoencefálica e placentária.

Apesar da sua grande fragilidade *ex vivo*, em 1998 um projeto pioneiro revelou pela primeira vez a sequência completa do genoma de *T. pallidum*, após a propagação laboratorial da bactéria em modelo animal (3). Quase duas décadas depois, o nosso conhecimento a nível genómico desta bactéria aparentou ter estagnado (apenas 5 outros genomas foram publicados neste período), contrastando de forma evidente com outras bactérias patogénicas para as quais se conhecem centenas ou milhares de sequências genómicas. Não obstante, estudos anteriores já mostravam que *T. pallidum* possuía um genoma pequeno (~1,1 Mb), bastante conservado, com dois antigénios principais a serem codificados por genes altamente variáveis, *arp* e *tprK*, devido à presença de motivos genéticos repetitivos variáveis e de regiões variáveis fruto de mecanismos de conversão génica, respetivamente (3-6).

Neste contexto, surgiu a necessidade de encontrar uma alternativa para capturar, sequenciar e estudar o genoma de *T. pallidum*, diretamente de amostras clínicas humanas, evitando a necessidade da cultura em modelo animal. Neste estudo, esse obstáculo foi ultrapassado, permitindo estudar, pela primeira vez, como é que este agente patogénico vai moldando o seu genoma para se adaptar e sobreviver no contexto da sífilis. Assim, foi possível descodificar ao nível genético o “braço-de-ferro” estabelecido entre a bactéria e o ser humano durante a doença, tendo sido identificados quais os mecanismos que a bactéria usa para se adaptar e assim sobreviver ao ataque do sistema imunitário do Homem.

## \_Objetivo

O presente estudo descreve a implementação de uma abordagem inovadora para capturar e sequenciar o genoma de *T. pallidum* no contexto da infeção humana. Os objetivos específicos foram: i) identificação dos principais fatores genéticos associados à capacidade de sobrevivência e persistência de *T. pallidum* durante a infeção; e, ii) caracterização genómica das estirpes de *T. pallidum* circulantes em Portugal.

## \_Material e métodos

Como ponto de partida utilizaram-se amostras de DNA total extraído de amostras clínicas enviadas pela Clínica de Doenças Sexualmente Transmissíveis do Centro de Saúde da Lapa para o Laboratório de Referência de Infeções Bacterianas Sexualmente Transmissíveis do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA). De forma a capturar o genoma completo de *T. pallidum* diretamente de amostras clínicas, foram desenhadas a nível bioinformático e posteriormente sintetizadas cerca de 20000 sondas de RNA complementares ao genoma da bactéria (7). O DNA alvo foi de seguida capturado tirando partido de partículas magnéticas acopladas às sondas desenhadas, sendo subsequentemente sujeito à metodologia de sequenciação do genoma completo e a análise bioinformática detalhada.

## \_Resultados e discussão

A importância da identificação e avaliação dos fatores genéticos envolvidos na adaptação, sobrevivência e persistência de agentes bacterianos patogénicos no contexto de infeção humana é inequívoca. No entanto, no que diz respeito ao agente etiológico causador da sífilis, a impossibilidade de propagá-lo em laboratório tem impedido o desenvolvimento de estudos focados nessa avaliação.

Neste contexto, a abordagem disruptiva descrita neste estudo constitui um importante avanço científico para o conhecimento da sífilis, permitindo analisar não só a diversidade genética entre estirpes de *T. pallidum* que infetam diferentes doentes, mas também a diversidade da população bacteriana que infeta um único doente. Em particular, este estudo

permitiu revelar a existência de uma estirpe de *T. pallidum* epidémica em Portugal, o que não poderia ser detetado através das metodologias moleculares tradicionais usadas até à data. Estudos posteriores encontram perfis genéticos similares em múltiplos países à escala global, apontando que os casos de sífilis atuais são provavelmente causados por uma estirpe de perfil pandémico (8).

Observou-se que o polimorfismo genético entre as estirpes circulantes estava concentrado essencialmente em 13 genes (englobando cerca de 70% do polimorfismo entre estirpes colhidas de doentes diferentes) os quais terão um papel crucial na interação da bactéria com o Homem durante a sífilis. Mais importante, descobriu-se que este agente patogénico modifica o seu genoma durante o processo infeccioso, modificando o *status* “ON/OFF” de algumas proteínas potencialmente associadas à virulência, tais como proteínas membranares e proteínas associadas à motilidade. A geração e manutenção de diversidade genética durante o processo infeccioso (“intra-doente”) parece ser fundamental para que a bactéria *T. pallidum* persista no hospedeiro. Este fenómeno afeta particularmente o maior antigénio da bactéria (TpK), o qual revelou uma enorme e surpreendente diversidade (“inter- e intra-doente”), apresentando perfis genéticos no contexto da infeção humana diferentes daqueles previamente obtidos após propagação laboratorial em modelo animal. Esta observação suporta a hipótese de que o estudo genético deste agente patogénico durante o processo infeccioso no Homem é fundamental para a compreensão da sua capacidade de não ser reconhecido pelo sistema imunitário, invadir e persistir no sangue, e ultrapassar as barreiras hematoencefálica e placentária.

Finalmente, identificou-se que as estirpes circulantes possuem, não só mutações associadas à resistência à azitromicina (antibiótico de segunda linha para o tratamento da sífilis), bem como mutações nunca descritas até à data afetando uma proteína, cujos homólogos noutras bactérias estão associados à resistência aos antibióticos  $\beta$ -lactâmicos. Esta proteína constitui, assim, um alvo importante de vigilância futura de uma potencial diminuição da suscetibilidade à penicilina (o antibiótico de primeira linha para o tratamento da sífilis). Este

estudo ilustra a potencialidade da captura de material genético diretamente de amostras clínicas, acoplada à sequenciação do genoma completo, no estudo de infeções bacterianas. A possibilidade de poder avaliar a população bacteriana que está efetivamente a circular no Homem durante o processo infeccioso abre as portas para um melhor conhecimento dos mecanismos envolvidos na adaptação e persistência dos agentes patogénicos.

### Conclusão

Este estudo permitiu descodificar os principais mecanismos genéticos pelos quais a bactéria *T. pallidum* evade o sistema imunitário e se adapta ao Homem nesta complexa e multifásica doença.

A aplicação desta estratégia inovadora de monitorização da interação Homem-bactéria poderá ser importante para o desenvolvimento de novas medidas terapêuticas e/ou profiláticas, tais como o desenvolvimento de uma vacina, na medida em que se conseguiu identificar os perfis antigénicos dominantes da bactéria no processo infeccioso. Noutra perspetiva, esta abordagem constitui também um ponto de viragem para o aperfeiçoamento de metodologias de diagnóstico e de epidemiologia molecular, o que permitirá aumentar o conhecimento da distribuição geográfica, das vias de transmissão e das propriedades de virulência deste agente patogénico para bem da saúde pública.

O INSA tem um curso um estudo focado em continuar esta importante linha de investigação, o qual foi recentemente financiado pela Sociedade Europeia de Microbiologia Clínica e Doenças Infecciosas ([https://www.escmid.org/profession\\_career/awards\\_grants/research\\_grants/awardees\\_2018/](https://www.escmid.org/profession_career/awards_grants/research_grants/awardees_2018/)).

Dada a relevância do estudo do perfil genético de *T. pallidum* em amostras clínicas associadas aos diferentes locais anatómicos e estadios de sífilis, sublinha-se a importância do envio deste tipo de amostras para o Laboratório de Referência de Infeções Bacterianas Sexualmente Transmissíveis. A equipa do INSA está disponível para encetar as colaborações necessárias nesse sentido.

## Agradecimentos

Este estudo foi financiado pela Fundação para a Ciência e Tecnologia (EXPL/BIA-MIC/0309/2013). Agradece-se à Doutora Cristina Furtado pela revisão do artigo.

Comentário resumido do artigo original publicado: Pinto M, Borges V, Antelo M, Pinheiro M, Nunes A, Azevedo J, Borrego MJ, Mendonça J, Carpinteiro D, Vieira L, Gomes JP. *Genome-scale analysis of the non-cultivable Treponema pallidum reveals extensive within-patient genetic variation*. Nat Microbiol. 2016 Oct 17;2:16190. <https://doi.org/10.1038/nmicrobiol.2016.190>

## Referências bibliográficas:

- (1) Lafond RE, Lukehart SA. Biological basis for syphilis. Clin Microbiol Rev. 2006;19(1):29-49. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1360276/>
- (2) Cullen PA, Cameron CE. Progress towards an effective syphilis vaccine: the past, present and future. Expert Rev Vaccines. 2006;5(1):67-80.
- (3) Fraser CM, Norris SJ, Weinstock GM, et al. Complete genome sequence of Treponema pallidum, the syphilis spirochete. Science. 1998;281(5375):375-88. <http://science.sciencemag.org/content/281/5375/375.long>
- (4) Giacani L, Brandt SL, Puray-Chavez M, et al. Comparative investigation of the genomic regions involved in antigenic variation of the TprK antigen among treponemal species, subspecies, and strains. J Bacteriol. 2012;194(16):4208-25. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3416249/>
- (5) Centurion-Lara A, LaFond RE, Hevner K, et al. Gene conversion: a mechanism for generation of heterogeneity in the tprK gene of Treponema pallidum during infection. Mol Microbiol. 2004;52(6):1579-96. <https://doi.org/10.1111/j.1365-2958.2004.04086.x>
- (6) Pillay A, Liu H, Chen CY, et al. Molecular subtyping of Treponema pallidum subspecies pallidum. Sex Transm Dis. 1998 Sep;25(8):408-14.
- (7) Gnirke A, Melnikov A, Maguire J, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol. 2009; 27(2):182-9. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2663421/>
- (8) Arora N, Schuenemann VJ, Jäger G, et al. Origin of modern syphilis and emergence of a pandemic Treponema pallidum cluster. Nat Microbiol. 2016;2:16245. <https://www.zora.uzh.ch/id/eprint/134132/>