

Genômica comparativa e evolutiva de *Leptospira*

DOI: 10.3395/reciis.v1i2.Sup.103pt



Natalia Rego

Unidad de Bioinformática,
Institut Pasteur de Montevideo,
Montevideo, Uruguai
nrego@pasteur.edu.uy



Hugo Naya

Unidad de Bioinformática,
Institut Pasteur de Montevideo,
Montevideo, Uruguai
naya@pasteur.edu.uy

Guillermo Lamolle

Unidad de Bioinformática, Institut Pasteur de Montevideo / Sección Biomatemática, Facultad de Ciencias, Universidad de la República (UdelaR), Montevideo, Uruguay
glamolle@pasteur.edu.uy

Fernando Álvarez-Valin

Sección Biomatemática, Facultad de Ciencias, Universidad de la República (UdelaR), Montevideo, Uruguay
falvarez@fcien.edu.uy

Resumo

A disponibilidade de seqüências genômicas completas tem possibilitado o estudo de mudanças no conteúdo gênico e nucleotídico durante a evolução de linhagens de bactérias patogênicas e simbióticas. Aqui nós apresentamos resultados de genômica comparativa entre quatro cepas pertencentes a duas espécies, *Leptospira interrogans* (*L. interrogans* Lai str. 56601 e *L. interrogans* Copenhageni str. Fiocruz L1-130) e *L. borgpetersenii* (*L. borgpetersenii* sorovar Hardjo-bovis str. JB197 e *L. borgpetersenii* sorovar Hardjo-bovis str. L550). As duas cepas da última espécie reduziram sua gama de hospedeiros estando restritas a um ciclo de transmissão hospedeiro-a-hospedeiro enquanto o anterior pode sobreviver por longos períodos em água doce. Uma análise composicional e substitucional detalhada em 2416 quartetos de ortólogos indica que esta mudança nas condições do ambiente está acompanhada de mudanças no conteúdo GC genômico afetando todo o genoma. Além disso, esta análise revela que a divergência entre espécies é surpreendentemente vasta, com uma razão de distâncias sinônimos / aminoácidos igual a 7.17, uma proporção muito maior do que em outros grupos. Nós mostramos que esta razão elevada é uma indicação de aceleração nas taxas de substituição que afetaram especialmente as posições sinônimas estando conectada à mudança na composição de bases já descrita. Nós postulamos que este processo ocorreu especificamente no ramo que leva a *L. borgpetersenii*.

Palavras-chave

Leptospira, redução de genoma, conteúdo GC, taxas de substituição, gama de hospedeiros

Introdução

A ordem Spirochaetales consiste de várias espécies de bactérias alongadas e finas cuja origem monofilética é evidente a partir de atributos estruturais múltiplos que são únicos deste grupo (particularmente o cilindro protoplasmático helicoidal envolvido por um revestimento externo). A reconstrução filogenética molecular baseada no 16S rRNA sustenta estes dados ultra-estruturais, apesar de clados de espiroquetas se ramificarem mais profundamente do que clados em outras ordens de bactéria (PASTER et al., 1991, p.6104-6108). A gama ecológica de espiroquetas é muito ampla, estendendo-se desde organismos de vida livre a formas parasitas e de espécies aeróbicas a anaeróbicas.

Durante sua história evolutiva, as espiroquetas sofreram variações muito significativas na composição e tamanho do genoma. Uma destas mudanças está descrita na Figura 1, onde a variabilidade do conteúdo GC

é representada em uma perspectiva filogenética. Esta figura mostra que a variabilidade não é restrita apenas às comparações entre gêneros, mas mais surpreendentemente, há várias espécies pertencentes a um mesmo gênero, e portanto supostamente relacionadas evolutivamente, cujos conteúdos GC genômicos são notadamente diferentes. Como um exemplo extremo, nós podemos mencionar o gênero *Treponema*, onde *Treponema pallidum* (agente causador da sífilis) e *Treponema denticola* (*espiroqueta de placa periodontal*) exibem conteúdos GC genômicos de 53% e 38% respectivamente. Outro exemplo surpreendente é fornecido pelas espécies do gênero *Leptospira*, cujas ligações filogenéticas são ainda mais próximas, e ainda assim elas exibem diferenças muito significantes no conteúdo GC genômico. No momento, a heterogeneidade na composição de bases das espiroquetas tem sido discutida como um atributo diagnóstico para propósitos taxonômicos (OLSEN et al., 2000, p.45-46) mas ainda não foi avaliada em um contexto evolutivo.

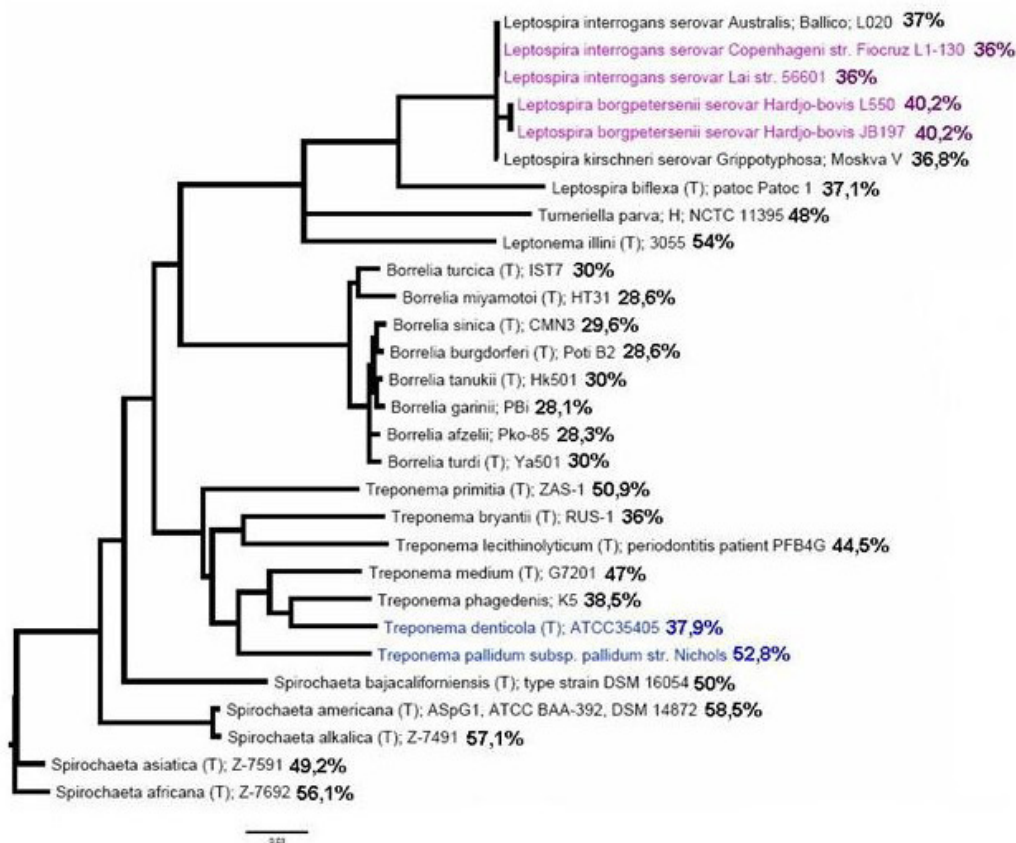


Figura 1 – Relações filogenéticas entre 29 espiroquetas inferidas de seqüências 16SrRNA obtidas no Projeto Banco Dados Ribossômico RDP-II (The Ribosomal Database Project): <http://rdp.cme.msu.edu>. A árvore foi construída com Weighbor, o método implementado neste local. O conteúdo GC genômico para cada taxa é mostrado. O comprimento dos ramos é proporcional ao número de substituições. As espécies de interesse estão destacadas com cores diferentes. A escala de referência (que aparece abaixo) representa 0.03 substituições de nucleotídeos por posição.

Não apenas na espécie de *Treponema* mencionada acima, mas também para o par de espécies *L. interrogans* – *L. borgpetersenii*, a diferença no tamanho de seus genomas tem sido explicada como um resultado da expansão genômica em uma espécie (por duplicação e transferência de gene horizontal) e da redução genômica na outra (e.g. perda de gene mediado por IS). Ambos os processos, redução de genoma e transferência de gene horizontal, estão relacionados a estilos de vida diferentes encontrados no gênero *Leptospira*. Isto foi proposto com base nas duas observações seguintes: primeiro, a redução genômica confronta a diminuição da gama de hospedeiros, e segundo, o processo de perda de gene não é aleatório, mas afeta principalmente genes cuja presença permitiria percepção ambiental e maior capacidade metabólica de transporte e utilização (BULACH et al., 2006, p.14562-14564; SESHADRI et al., 2004, p.5648-5649). Diferentemente de outras reduções de genoma microbiais, onde há um enriquecimento do conteúdo AT do genoma associado ao aumento da especificidade de hospedeiro (e.g. *Buchnera*, *Mycobacterium*; MORAN, 2002, p.585), estas espiroquetas mostram a tendência oposta, a qual indica que o conteúdo GC genômico aumenta em associação a alta especificidade de hospedeiros.

Neste trabalho, nós analisamos vários aspectos evolutivos no gênero *Leptospira* com uma ênfase especial aos processos que envolvem mudanças no conteúdo GC genômico e taxas de substituição de nucleotídeos, tentando caracterizar em detalhes estes processos evolutivos e esclarecer as forças evolutivas que as dirigiram. Por esta razão, nos concentramos em quatro cepas cujos genomas foram completamente determinados: duas cepas pertencentes ao *L. interrogans* (*L. interrogans* Lai str. 56601 e *L. interrogans* Copenhageni str. Fiocruz L1-130) e duas cepas do *L. borgpetersenii* sorovar Hardjo-bovis (str. JB197 e str. L550). Na média, os últimos são 17% menores em tamanho de genoma e 5% maiores em conteúdo GC.

Ambas as espécies têm sido extensivamente estudadas porque elas são os agentes causadores da maioria dos casos de leptospirose, uma das zoonoses mais difundidas, responsável por mais de meio milhão de casos humanos por ano no mundo todo. Apesar das duas produzirem sintomas clínicos da doença semelhantes, elas são epidemio-

logicamente muito diferentes. Enquanto o *L. interrogans* obtém solo úmido ou água doce na urina do hospedeiro e é capaz de sobreviver mais de 200 dias em ambientes aquáticos até infectar um novo hospedeiro mamífero, o *L. borgpetersenii* não tolera a privação de nutrientes e é limitado a um ciclo de transmissão de hospedeiro-a-hospedeiro (BULACH et al., 2006, p.14562-14563).

Materiais e Métodos

As seqüências genômicas completas das espécies aqui analisadas foram obtidas de bancos de dados públicos. Ortólogos disponíveis nas quatro espécies foram identificados utilizando Blastp através dos melhores *hits* recíprocos (ALTSCHUL et al., 1990). Desta forma, foi possível identificar 2419 grupos de genes ortólogos que estão presentes nas quatro espécies. Cada um destes grupos foi precisamente alinhado ao nível de aminoácido utilizando MUSCLE (EDGAR, 2004) e então traduzido novamente para a seqüência conhecida de DNA.

Distâncias sinônimas e não-sinônimas foram calculadas utilizando o método de NEI E GOJOBORI (1985) com as modificações sugeridas por ZHANG et al. (1998) que corrige desvios de transição/transversão que afetam principalmente a forma de contar o número de posições sinônimas e não-sinônimas na terceira posição do códon de duetos (códon duplamente degenerados). Neste trabalho, a correção foi feita de acordo com a proporção transição/transversão observada na terceira posição do códon de quartetos.

Comparação de padrões composicionais em genes ortólogos

Uma análise substitucional e composicional detalhada de 2416 quartetos de ortólogos foi realizada e é apresentada na Tabela 1. Deve-se enfatizar que todos os cálculos foram efetuados após a eliminação de intervalos de alinhamentos, garantindo, desta forma, que os segmentos comparados para cada gene fossem estritamente homólogos. Isto garante que qualquer diferença que pudesse ser detectada entre os gêneros em comparação não seja atribuída a inserções ou exclusões diferenciais ocorrendo no gênero em questão, mas exclusivamente ao processo de substituição de nucleotídeos.

Tabela 1 – Padrões composicionais em genes ortólogos

Sorovar - Espécie	GC ₁	GC ₂	GC ₃
<i>L. borgpetersenii</i> Hardjo-bovis str. JB197	0.4772	0.3520	0.3933
<i>L. borgpetersenii</i> Hardjo-bovis str. L550	0.4772	0.3521	0.3933
<i>L. interrogans</i> Lai str. 56601	0.4600	0.3453	0.2813
<i>L. interrogans</i> Copenhageni str. Fiocruz L1-130	0.4601	0.3453	0.2812
T-teste para comparações pareadas	37.5, P << 10 ⁻¹²	18.2, P << 10 ⁻¹²	112, P=0

A primeira observação evidente que pode ser retirada da Tabela 1 é a de que ambas as cepas do *L. interrogans* e ambas as cepas do *L. borgpetersenii* são muito semelhantes, ainda assim, eles são consistentemente diferentes quando a comparação é feita entre cepas pertencentes a espécies diferentes. De fato, as cepas do *L. borgpetersenii* são mais ricas em GC do que as cepas do *L. interrogans*, sendo estas diferenças consistentes e estatisticamente muito significantes (t-teste para comparações pareadas) para as posições dos três códons. Contudo, a diferenciação é especialmente manifestada na posição do terceiro códon, onde a média de diferença é superior a 10%.

Torna-se necessário verificar que este comportamento não é uma diferenciação média simples e ilegítima, causada por alguns genes com conteúdo GC diferentes adquiridos talvez por transferência de genes horizontal, mas que afetou o genoma como um todo e a maioria de

seus genes constituintes. Por esta razão, nós desenhamos o nível de GC de um sorovar contra o outro para cada gene e cada posição de códon. (Figura 2). A linha diagonal foi desenhada nestas figuras para indicar o lugar ao redor de onde os pontos (genes) deveriam simetricamente cair se a distribuição não fosse tendenciosa. Em outras palavras, se o conteúdo GC de um determinado gene é o mesmo em ambas as espécies, então o ponto correspondente àquele gene deveria cair exatamente na linha. Entretanto, como é evidente pelas figuras, a vasta maioria de pontos está localizada acima da linha diagonal, indicando então que para estes genes o nível de GC é superior no *L. borgpetersenii*. De acordo com os resultados apresentados da Tabela 1, os desvios acima da linha diagonal são mais marcantes no GC₁ do que no GC₂, mas eles são particularmente fortes para a posição do terceiro códon.

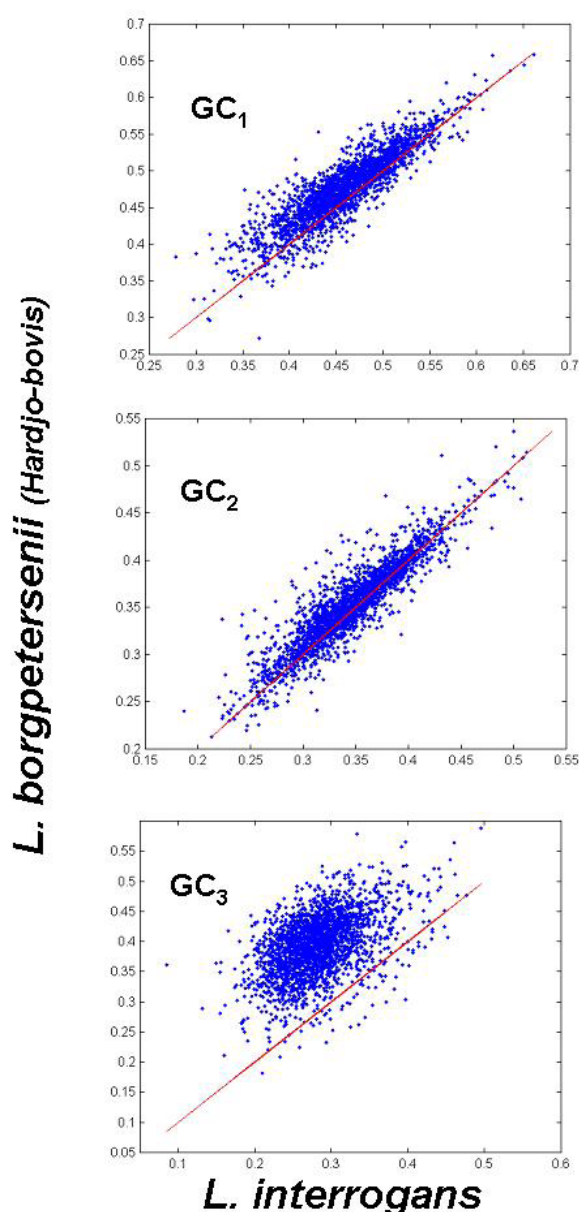


Figura 2 – Gráfico de dispersão entre níveis de GC em *Leptospira* para cada gene e para cada posição.

Tais diferenças marcantes na composição de GC no gênero *Leptospira* tornam-se particularmente enigmáticas porque as cepas do *L. borgpetersenii*, que perderam a habilidade ancestral de sobreviver fora do hospedeiro, são notavelmente mais ricas em GC do que as cepas do *L. interrogans* relacionadas, em contradição ao comportamento esperado para o conteúdo GC em reduções genômicas associadas ao parasitismo (ROCHA e DANCHIN, 2002, p.291-294). Por outro lado, a diminuição na extensão do nicho correspondente à redução do genoma em um grupo de *Leptospira* poderia levar a um aumento na importância de fatores relativamente menores por outro lado, tais

como a temperatura de crescimento ideal (MUSTO *et al.*, 2004).

Taxas evolutivas no *Leptospira*: aceleração de taxas silenciosas

Nós estimamos as distâncias sinônimas e dos aminoácidos para os 2419 grupos de ortólogos. Os resultados destas análises estão apresentados na Tabela 2 e Figura 3. Esta última mostra a distribuição nas frequências de comparações intergrupais, para ambas as distâncias sinônimas e de aminoácidos, painéis a e b respectivamente.

Tabela 2 – Distâncias de nucleotídeo no *Leptospira*

Sorovar-Espécie	1	2	3	4
Taxas sinônimas				
1- <i>L. interrogans</i> Copenhageni str. Fiocruz L1-130	0			
2- <i>L. interrogans</i> Lai str. 56601	0.0038	0		
3- <i>L. borgpetersenii</i> Hardjo-bovis str. L550	1.2266	1.2273	0	
4- <i>L. borgpetersenii</i> Hardjo-bovis str. JB197	1.2257	1.2253	0.0115	0
Taxas de aminoácido				
1- <i>L. interrogans</i> Copenhageni str. Fiocruz L1-130	0			
2- <i>L. Interrogans</i> Lai str. 56601	0.0021	0		
3- <i>L. borgpetersenii</i> Hardjo-bovis str. L550	0.1687	0.1687	0	
4- <i>L. borgpetersenii</i> Hardjo-bovis str. JB197	0.1688	0.1687	0.003	0
Pequenas subunidades de genes rRNA				
1- <i>L. interrogans</i> Copenhageni str. Fiocruz L1-130	0			
2- <i>L. interrogans</i> Lai str. 56601	0.00072	0		
3- <i>L. borgpetersenii</i> Hardjo-bovis str. L550	0.007889	0.007889	0	
4- <i>L. borgpetersenii</i> Hardjo-bovis str. JB197	0.007883	0.007883	0	0

Estes resultados são dignos de alguns comentários. Em primeiro lugar, as distâncias entre cada subgrupo (sorovar) são muito pequenas, ambas no nível sinônimo e de aminoácidos. Por outro lado, as distâncias intergrupais são relativamente grandes, em uma média de cerca de 0.17 mudanças de aminoácidos por códon (para mudanças de aminoácidos), enquanto a divergência silenciosa chega ao número de 1.22 mudanças por sítio sinônimo (em média); que está no alcance de saturação. Tal extensão de divergência de nucleotídeos silenciosos é alarmante se levarmos em conta o fato de que estamos comparando espécies pertencentes ao mesmo gênero.

Finalmente e o mais relevante, a razão das distâncias sinônimas/aminoácidos é de 7.17 (1.22/0.17) um número alto e notável.

A fim de comparar a quantidade de divergências (assim como sua proporção) observada no *Leptospira*, torna-se necessário ter um grupo de referência para o qual as informações de número de divergências e taxas de substituição estejam disponíveis. Estas distâncias foram então comparadas à quantidade de divergências entre ordens de mamíferos diferentes. Vale lembrar que as ordens de mamíferos divergiram umas das outras há 100 milhões de anos atrás (Mya) (SPRINGER *et al.*, 2003). De

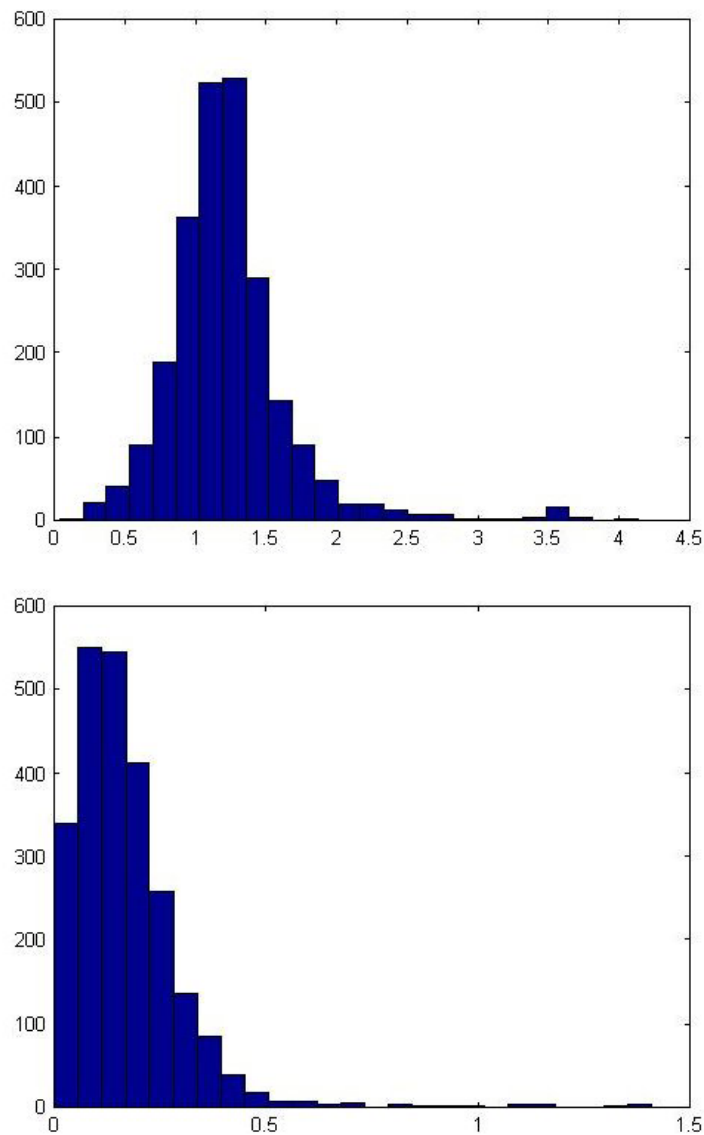


Figura 3 – Distribuição de distâncias (a) sinônimas e (b) de aminoácidos (correção de Poisson) entre genes ortólogos de serovares de *Leptospira interrogans* e *Leptospira borgpetersenii*.

maneira interessante o bastante, as distâncias sinônimas para mamíferos são em média de 0.35 mudanças por local (esta média foi obtida de 50 alinhamentos diferentes de genes ortólogos incluindo quatro ordens de mamíferos distintas, tiradas de ALVAREZ-VALIN *et al*, 1998), isto é, menor do que um terço da média da distância sinônima observada entre as espécies *L. interrogans* e *L. borgpetersenii*. Em contraste, a distância média de aminoácidos entre ordens de mamíferos é de 0.21 mudanças por códon, notoriamente maior do que aquela entre *L. interrogans* e *L. borgpetersenii*. A razão das distâncias sinônimas/aminoácidos em mamíferos é de 1.67, embora um quarto disto entre leptospiros.

Nós também comparamos o total da distância de nucleotídeos em *Leptospira* versus mamíferos para pequenas subunidades de genes rRNA (isto é, 16S rRNA

em *Leptospira* e 18S rRNA em mamíferos). A ilustração que surge desta comparação é novamente inversa àquela exibida por distâncias sinônimas, isto é, as distâncias entre *L. interrogans* e *L. borgpetersenii* (0.79%; Tabela 2) são menores do que aquelas entre as ordens de mamíferos (1.1%, não mostrado).

À primeira vista, esta disparidade no comportamento entre posições silenciosas de um lado, e posições codificadoras de aminoácidos (ou genes ribossômicos) de outro, parece um tanto quanto enigmática. Contudo, é importante levar em conta que se espera que os fatores evolutivos que tendem a aumentar as taxas de substituição de nucleotídeos (e.g. tempo de geração, taxa mutacional etc.) produzam um efeito maior nas taxas sinônimas do que nas não-sinônimas (OHTA, 1995) ou em genes ribossômicos. Em outras palavras, a aceleração

nas taxas de substituição afetaram mais provavelmente o genoma como um todo, mas apareceram particularmente marcantes nas posições sinônimas.

Nós supomos que esta aceleração esteja em conexão com a mudança na composição de bases já descrita, e tenha ocorrido especificamente na ramificação que leva ao grupo *L. borgpetersenii*. Na próxima seção apresentamos evidências que sustentam esta hipótese.

Composição ancestral no *Leptospira*

Como descrito previamente, o *Leptospira interrogans* e o *L. borgpetersenii* diferem em seu conteúdo GC genômico, sendo esta diferença particularmente marcante na terceira posição dos códons. Três cenários evolutivos possíveis poderiam ser considerados. O primeiro é de que *L. interrogans* permanece na condição ancestral e sendo assim, *L. borgpetersenii* sofreu um aumento no GC genômico. Alternativamente, pode-se postular que *L. borgpetersenii* representa a condição ancestral e consequentemente, *L. interrogans* foi sujeito a uma diminuição no GC genômico. Finalmente, outro cenário, mas menos parsimonioso, é de que nenhum dos dois representa a condição ancestral.

A fim de ser capaz de discernir a direção da mudança na composição de bases, torna-se necessário determinar a condição ancestral. Como aparece na Figura 1, *Leptospira biflexa* se ramifica muito antes de *L. interrogans* e *L. borgpetersenii*, este pode então ser usado como um grupo externo para inferir a composição de GC ancestral dentro do gênero *Leptospira*. Uma inspeção visual desta figura imediatamente sugere que o primeiro cenário é o mais provável, isto é, *L. interrogans* representa a condição ancestral, uma vez que ambos *L. biflexa* (37% GC) e *L. interrogans* (entre 35 e 37% GC) têm conteúdos genômicos GC muito semelhantes, notoriamente inferiores ao do *L. borgpetersenii* (40.2 %).

Conclusões

O gênero *Leptospira* é particularmente interessante porque cepas de *L. borgpetersenii*, que perderam a habilidade ancestral de sobreviver fora do hospedeiro, são notavelmente mais ricas em GC do que seus parentes próximos do grupo *L. interrogans*. Esta observação está em clara contradição com o que seria o comportamento esperado do conteúdo GC em reduções de genoma associados ao parasitismo. Da comparação de genes ortólogos, mostramos que as diferenças no conteúdo GC entre ambos os grupos ocorrem no genoma inteiro, até nas posições de códons não-sinônimos, mas é particularmente marcante nas posições sinônimas.

Nós também analisamos a proporção entre distâncias sinônimas e não-sinônimas entre grupos de *Leptospira* e comparamos estes resultados com as proporções obtidas de ordens de mamíferos diferentes que divergiram entre si há cerca de 100 Mya. Surpreendentemente, a distância sinônima para as ordens de mamíferos é menor do que um terço daquela correspondente aos grupos de *Leptospira*. Em contraste, a média da distância entre aminoácidos de

ordens de mamíferos é notavelmente maior do que aquela entre os grupos de *L. interrogans* e *L. borgpetersenii*. Estas tendências opostas em distâncias de nucleotídeos levam a diferenças poderosas na proporção de taxas de substituição sinônimas para não-sinônimas, e comprovam uma grande aceleração do genoma nas taxas de substituições que afetaram especialmente a terceira posição dos códons em genes codificadores de proteína.

Devido à escassez de informações sobre dados de sequência de outras espécies deste gênero, é relativamente difícil discernir fielmente seu conteúdo GC ancestral. Entretanto, os dados fragmentados disponíveis parecem favorecer a hipótese de um conteúdo GC ancestral semelhante ao do atual *L. interrogans*.

Em resumo, a imagem global mostra mudanças notáveis nos níveis composicionais e taxas de substituição entre grupos do gênero *Leptospira*. Seria interessante aprofundar a investigação de possíveis forças evolutivas e biológicas dirigindo estes processos. Mais especificamente, seria interessante determinar se a mudança observada na composição poderia ser atribuída a mudanças no padrão básico de mutações ou a fatores funcionais (i.e. razões seletivas). A primeira possibilidade poderia ser testada determinando-se o padrão de substituições em pseudogenes que é amplamente reconhecido por refletir o espectro mutacional. Desde que as razões funcionais sejam consideradas, uma explicação possível para o aumento no GC genômico no grupo *L. borgpetersenii* poderia ser encontrada na hipótese termodinâmica proposta por BERNARDI e BERNARDI (1986) e mais recentemente sustentada por resultados de MUSTO et al. (2004; 2005). De acordo com esta hipótese, o aumento no GC seria uma resposta adaptativa à temperatura ambiente crescente. Esta hipótese se adequa aos resultados aqui presentes desde que *L. borgpetersenii* seja restrito ao hospedeiro de “sangue quente” (mamífero).

Agradecimentos

Somos gratos a Rosina Piovani pelas úteis discussões. Este trabalho foi parcialmente apoiado pela concessão FPTA-252 “Desenvolvimento de capacidades bioinformáticas na área de anotação genômica” do Instituto Nacional de Pesquisa Agropecuária (Uruguai).

Referências bibliográficas

- ALTSCHUL, S.F. et al. Basic local alignment search tool. *Journal of Molecular Evolution*, v.215, n.3, p.403-410, 1990.
- ALVAREZ-VALIN, F.; JABBARI, K.; BERNARDI, G. Synonymous and nonsynonymous substitutions in mammalian genes: intragenic correlations. *Journal of Molecular Evolution*, n.46, p.37-44, 1998.
- BERNARDI G.; BERNARDI G. Compositional constraints and genome evolution. *Journal of Molecular Evolution*, v.24, n.1-2, p.1-11, 1986.
- BULACH, D.M. et al. Genome reduction in *Leptospira borgpetersenii* reflects limited transmission potential.

PNAS, v.103, n.39, p.14560-14565, 2006. Disponível em: www.pnas.org/cgi/doi/10.1073/pnas.0603979103.

EDGAR, R.C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. **BMC Bioinformatics**, v.5, n.113, 2004. [doi:10.1186/1471-2105-5-113].

MORAN, N.A. Microbial minimalism: genome reduction in bacterial pathogens. **Cell**, v.108, p.583-586, 2002.

MUSTO, H. et al. Correlations between genomic GC levels and optimal growth temperatures in prokaryotes. **FEBS Letters**, v.573, n.1-3, p.73-77, 2004.

MUSTO, H. et al. The correlation between genomic G+C and optimal growth temperature of prokaryotes is robust: a reply to Marashi and Ghalanbor. **Biochem Biophys Res Commun**, v.330, p.357-360, 2005. [doi:10.1016/j.bbrc.2005.02.133].

NEI, M.; GOJOBORI, T. Simple methods for estimating the number of synonymous and nonsynonymous nucleotide substitutions. **Molecular Biology and Evolution**, v.3, n.5, p.418-426, 1986.

OLSEN, I.; PASTER, B.J.; DEWHIRST, F.E. Taxonomy of spirochetes. **Anaerobe**, n.6, p.39-57, 2000. [doi:10.1006/anae.1999.0319].

OHTA, T. Synonymous and nonsynonymous substitutions in mammalian genes and the nearly neutral theory. **Journal of Molecular Evolution**, v.40, n.1, p.56-63, 1995.

PASTER, B.J. et al. Phylogenetic analysis of the Spirochetes. **Journal of Bacteriology**, v.173, n.19, p.6101-6109, 1991.

ROCHA, E.P.; DANCHIN, A. Base composition bias might result from competition for metabolic resources. **Trends in Genetics**, v.18, n.6, p.291-294, 2002.

SESHADRI, R. et al. Comparison of the genome of the oral pathogen *Treponema denticola* with other spirochete genomes. **PNAS**, v.101, n.15, p.5646-5651, 2004. Disponível em: www.pnas.org/cgi/doi/10.1073/pnas.0307639101.

SPRINGER, M.S. et al. Placental mammalian diversification and the Cretaceous-Tertiary boundary. **PNAS**, v.100, n.3, p.1056-1061, 2003. Disponível em: www.pnas.org/cgi/doi/10.1073/pnas.0334222100.

ZHANG, J.; ROSENBERG, H.F.; NEI, M. Positive Darwinian selection after gene duplication in primate ribonuclease genes. **PNAS**, v.95, n.7, p.3708-3713, 1998.



Sobre os autores

Natalia Rego

Graduada em Biologia, da Universidad de la República – Uruguay, em 2006. Atualmente, desenvolve a tese de mestrado em zoologia, na área de bioinformática e genômica evolutiva, no Instituto Pasteur de Montevideo.

Hugo Naya

Graduado em Biologia em 1999, e obteve o diploma de doutorado em biologia (genômica computacional), da Universidad de la República – Uruguay, em 2004. Atualmente, o trabalho dele envolve genômica computacional e genética quantitativa. No momento, ele é chefe da Unidade de Bioinformática do Instituto Pasteur em Montevideo.