

Busca computacional por grupos de genes homólogos sob evidência de seleção positiva em *Alphaherpesvirinae*

Giovanni Marques de Castro¹
Francisco Pereira Lobo²

Introdução

Alphaherpesvirinae são vírus patogênicos que infectam uma ampla variedade de animais, tais como humanos, bovinos, caprinos, suínos e aves (SIKALLIS et al., 2009). Uma das espécies virais mais estudadas dentre os *Alphaherpesvirinae* é o *Human Simplex Virus 1* (HSV-1), responsável pela herpes oral e prevalente em mais de 50% da população humana.

O estudo molecular da ação da seleção natural em regiões codificadoras evidencia um claro viés na frequência de mutações não-sinônimas quando comparada à frequência de mutações sinônimas. De maneira geral, alinhamentos múltiplos de códons de um dado grupo de genes homólogos possuem a vasta maioria das colunas do alinhamento sem variação no aminoácido codificado; já as mutações para códons sinônimos ocorrem em frequências consideravelmente maiores. Esse fenômeno ocorre porque mutações não-sinônimas usualmente reduzem a eficiência funcional da proteína codificada em comparação ao alelo não mutante fixado anteriormente. Assim, mutações não-sinônimas geralmente diminuem a aptidão evolutiva dos organismos que as possuem, e esses alelos menos funcionais são rapidamente removidos das populações pela seleção negativa ou purificadora (AGUILETA et al., 2009).

Entretanto, algumas posições em poucos genes podem apresentar uma frequência de mutações não-sinônimas significativamente mais elevada

¹ Biotecnologia, UFSCar, giovanni.castro@colaborador.embrapa.br

² Embrapa Informática Agropecuária, francisco.lobo@embrapa.br

do que o observado no restante dos alinhamentos de códons em análise, indicando a preferência pela fixação de novos alelos em detrimento aos antigos. Esse fenômeno é denominado seleção positiva ou seleção Darwiniana, sendo observado em códons e genes que codificam proteínas nas quais ocorre pressão seletiva para a variação ao invés da conservação do aminoácido na posição em análise quando comparada ao restante das posições, ou em uma dada sequência quando comparada ao restante das sequências.

Dentre os fenômenos biológicos que comumente possuem grupos de genes homólogos evoluindo sob pressão seletiva positiva destacam-se genes envolvidos em percepção sensorial, reprodução, imunidade e na relação parasita-hospedeiro (AGUILETA et al., 2009). A contínua luta coevolucionária entre hospedeiros e parasitas, na qual os genes dos hospedeiros evoluem para escapar da infecção parasitária e os genes de parasitas evoluem para escapar das defesas dos hospedeiros, frequentemente produzem uma corrida armamentista molecular, onde os genes envolvidos na relação parasita-hospedeiro evoluem mais rapidamente do que o restante dos genes encontrados nesses genomas. Assim, genes sob evidência de seleção positiva advindos de genomas de parasitas compreendem ótimos candidatos para alvos de drogas, bem como para permitir uma maior compreensão da relação ecológica parasita-hospedeiro em nível molecular. Este trabalho visa realizar a primeira busca genômica por genes sob evidência de seleção positiva em vírus do grupo *Alphaherpesvirinae*.

As regiões codificadoras dos 22 *Alphaherpesvirinae* depositados no banco de dados NCBI foram extraídas e traduzidas *in silico* para serem utilizadas como entrada no programa ORTHOMCL 1.4 (Li et al., 2003), de modo a se estabelecer as relações de homologia entre elas. Depois desse passo removemos os genes parálogos, bem como genes individuais com evidências de erros, e procedemos as análises dos grupos de homólogos que possuíam ao menos quatro genes, sendo um deles um gene de HSV-1. Cada grupo de proteínas homólogas foi alinhada utilizando o software MUSCLE (EDGAR, 2004). Os alinhamentos proteicos foram utilizados para se construir um alinhamento de códons a partir das respectivas regiões codificadoras. As relações filogenéticas entre as sequências proteicas foram estabelecidas utilizando-se o programa Proml do pacote Phylip utilizando-se 1000 bootstraps. Os alinhamentos de códons e as árvores

filogenéticas foram utilizadas para a busca por seleção positiva utilizando-se o programa Codeml do pacote PAML (YANG, 2007) pela avaliação de quatro modelos de evolução de códons: M1(neutro),M2(seleção positiva), M7(beta) e M8(beta+seleção positiva). A verossimilhança dos diferentes modelos foi então comparada (M1 comparado a M2 e M7 a M8), de modo a se selecionar os grupos de genes que possuísem verossimilhança significativamente maior dos modelos que permitem seleção positiva (M2 e M8) quando comparados a modelos que não permitem tal fenômeno (M1 e M7, respectivamente).

Dos 77 genes de HSV-1, foi verificado que quase 50% dos genes estão sob evidência de seleção positiva. 37 genes foram identificados pelo modelo M8, dos quais dois foram identificados também pelo modelo M2. Os genes identificados por ambos os modelos foram pUL36, a maior proteína do tegumento, altamente conservada entre os alfa herpesvírus, e a glicoproteína C, a qual é altamente antigênica e possui importante papel na evasão imune por parte dos vírus. Adicionalmente, dentre as 11 glicoproteínas de HSV-1, 7 apresentaram evidências de seleção positiva. Essas proteínas encontram-se na superfície das partículas virais e sob intensa pressão seletiva para mudar continuamente o seu perfil imunológico.

Este estudo detectou diversos genes virais sabidamente relacionados à modulação do sistema imune dos hospedeiros sob evidência de seleção positiva. Foram encontrados alguns genes sem função biológica conhecida sob evidência de seleção positiva, os quais podem ser futuros candidatos a alvos para o desenvolvimento de drogas antivirais.

Agradecimentos

À Embrapa Informática Agropecuária, por fornecer a infraestrutura computacional para a realização deste trabalho.

Referências

AGUILETA, G.; REFRÉGIER, G.; YOCKTENG, R.; FOURNIER, E.; GIRAUD, T. Rapidly evolving genes in pathogens: methods for detecting positive selection and examples among fungi, bacteria, viruses and protists. **Infection, Genetics and Evolution**, v. 9, n. 4, p. 656-70, Jul. 2009.

EDGAR, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. **BMC Bioinformatics**. v. 19, n. 5, p. 113, Aug. 2004.

Li, L.; STOECKERT JUNIOR, C. J.; ROOS, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. **Genome Research**, v. 13, n. 9, p. 2178-89, Sept. 2003.

SIKALLIS, G.; SPANDIDOS, D. A.; SOURVINOS, G. Herpesviridae and novel inhibitors. **Antiviral Therapy**, v. 14, n. 8, p. 1051-64, 2009.

YANG, Z. PAML 4: phylogenetic analysis by maximum likelihood. **Molecular Biology Evolution**, v. 24, n. 8, p. 1586-91, Aug. 2007.