Positive Selection in Homologous Genes of Alphaherpesviruses

Giovanni Marques de Castro, Jorge Augusto Hongo, Francisco Pereira Lobo

Embrapa Informática Agropecuária, Instituto de Computação, UNICAMP, Embrapa Informática Agropecuária

Alphaherpesviruses are pathogens that infect humans and a large variety of other animals. One of the most studied Alphaherpesviruses is the Human Simplex virus 1 (HSV-1), which causes oral Herpes and is prevalent in more than 50% of the human population. The continuous coevolutionary fight between hosts and pathogens, with hosts genes evolving to escape pathogen infection and pathogens genes evolving to escape host defences, often generate an 'molecular arms race', where the genes involved in host-parasite biological interaction tend to evolve faster than the genomic background. This phenomenon is named positive selection and, at the codon level, genes evolving under positive selection present a higher-than-expected ratio of non-synonymous mutations when compared with the ratio of synonymous mutations. The positively selected genes on parasite genomes are usually directly involved in modulating host immune system and, therefore, the search for such genes can point new drug targets, as well as can help the comprehension of parasitism at molecular level. This study comprises the first genomic search for positively selected genes on HSV-1 genome. The coding regions of 22 Alphaherpesviruses were translated in silico and used as input to ORTHOMCL 1.4 to establish homology relationships among them. After this step we removed groups with paralogous genes, as well as individual genes with evidence of errors, and further analyzed the group of homologous genes that contained a gene from the HSV-1 genome and had at least four genes. Each group of homologous proteins was aligned using the MUSCLE program. The protein alignments were then used to generate a codon alignment of the respective coding sequences. The phylogenetic relationships of each group of homologous proteins were established using the proml program from Phylip package with 1000 bootstraps. We used the codon alignment and the phylogenetic trees to search for positive selection using a site-based model using the codeml program from PAML 4.5 package. Each group was evaluated using four distinct codon evolution models: M1(neutral), M2(selection), M7(beta) and M8(beta+ω). The likelihood values of models M1/M2 and M7/M8 was evaluated to detect possible groups of genes under positive selection. We found that almost 50% of HSV-1 genes were found to be under positive selection, with 37 genes identified by M8 model and 2 genes by both M8 and M2 models. The genes identified by both models were the pUL36, the largest tegument protein and highly conserved in Alphaherpesviruses, and the glycoprotein C, that has a high antigenic response and an important role on immune evasion. Of special interest, the HSV-1 virus contains 11 glycoproteins, 7 of which presented signs of positive selection. These glycoproteins are at the surface of the virion and under great selective pressure to continuously change their antigenic profile.

Keywords: Positive selection, Alphaherpesviruses, homology, HSV-1, coevolution

Concentration area: Genomics Evolution