



Simpósio de Integração Acadêmica

“Bicentenário da Independência: 200 anos de ciência, tecnologia e inovação no Brasil e 96 anos de contribuição da UFV”

SIA UFV 2022



Análise filogenética de Calicreínas da família *Equidae*, do gênero *Bos* e de *Homo sapiens*

COSTA, Isabella Cristina Tolêdo¹; BARACAT-PEREIRA, Maria Cristina²; MENDES, Tiago Antônio de Oliveira²; GUIMARÃES, José Domingos³; LÓPEZ, Camilo José Ramírez⁴; ARAÚJO, Ana Karolina Ferreira⁵

¹Pós-graduação, Bioquímica Aplicada, UFV; ²Professor(a) DBB, UFV; ³ Professor, DVT, UFV; ⁴Pós-graduação, Medicina Veterinária, UFV; ⁵Graduação, Bioquímica, UFV

Calicreínas, Filogenia, Bioinformática

Introdução

As Calicreínas (do inglês *Kallikreins*, ou KLKs) são proteínas encontradas na parte não-celular do sêmen de mamíferos, envolvidas no equilíbrio da liquefação/coagulação seminal. São serino-proteases que possuem potencial para biomarcadores de distúrbios no organismo, como a KLK do tipo 3, ou *Plasma seminal antigen* (PSA), que é um biomarcador de câncer de próstata em humanos. KLK1E2 vem sendo estudada com a possibilidade de ser indicador de fertilidade em garanhões (*Equus caballus*) e touros (*Bos taurus*).

Objetivos

Este trabalho tem o objetivo de realizar a análise filogenética de KLKs do gênero *Bos*, da família *Equidae* e de *Homo sapiens*, visando estudar a conservação e as divergências evolutivas das KLKs no conjunto de dados e utilizar os resultados para trabalho com diferentes espécies do gênero *Equus*.

Material e Métodos

A sequência-modelo escolhida foi a de Calicreína 1E2 (KLK1E2) de *Equus caballus* (UniProtKB – Q6H321 – KLK2_HORSE). Foram utilizadas sequências nos formatos FASTA, depositadas no banco de dados do *National Center for Biotechnology Information* (NCBI), adquiridas por meio de confronto da sequência de aminoácidos da KLK1E2 na ferramenta tBLASTn, utilizando-se os genomas de referência escolhidos. Como critérios de seleção, as sequências de nucleotídeos utilizados deveriam possuir *e-value* menor que 10^{-10} , cobertura maior que 70% e identidade maior que 30%. Para obtenção da sequência completa do gene codificador de cada proteína, foram selecionadas sequências contendo 2.000 pares de bases antes e depois da sequência de nucleotídeos, que foram traduzidos pelo programa *online* Augustus, versão 3.3.3. A curadoria das sequências seguiu-se pela confirmação da presença dos domínios de tripsina em cada uma, assim como na KLK utilizada como

modelo (domínio da classe CL0124), utilizando-se a base de dados Pfam. Em seguida, foi feito um alinhamento das sequências curadas no Clustal W utilizando os parâmetros-padrão pela ferramenta MEGA X (versão 10.2.6 para macOS), a fim de se observar a conservação das sequências. Por fim, a árvore filogenética foi gerada pelo MEGA X.

Resultados e Discussão

A primeira análise envolveu 19 sequências de aminoácidos, que não gerou resultado significativo em função de valores de *bootstrap* inferiores a 50. Seguiu-se análises de conservações das sequências, e quatro delas foram retiradas da árvore por divergirem das outras na conservação do sítio de glicosilação, sítio ativo e cisteínas envolvidas nas ligações dissulfeto. A segunda análise, então realizada com as 15 sequências restantes, resultou em uma árvore significativa, com valores de *bootstrap* maiores que 50.

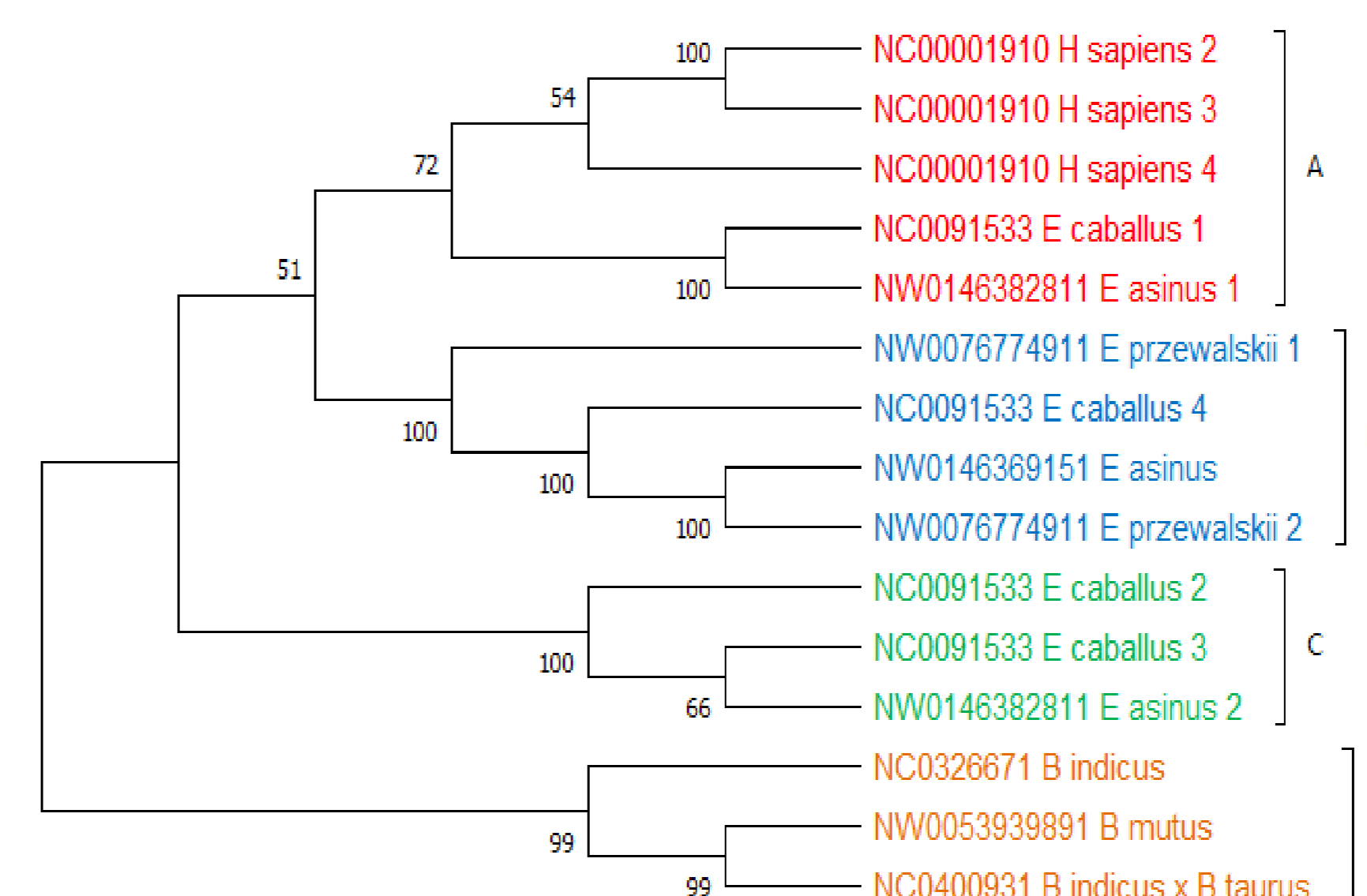


Figura 1. Árvore filogenética gerada no programa MEGA X com 15 sequências de calicreínas buscadas em genomas de referência de *Equidae*, *Bos* e *Homo sapiens*. A árvore apresenta 4 clados, identificados com as letras A, em vermelho, B, em azul, C em verde e D em laranja.

Conclusões

Há separação das KLKs por gêneros e as sequências da família *Equidae* são mais divergentes evolutivamente de *Bos* do que de *Homo sapiens*, evidenciando o potencial como biomarcador destas proteínas.

Apoio Financeiro

