

Análise *in silico* da Beta-Caseína de *Bos taurus*

In silico analysis of Beta Casein from Bos taurus

Sarah Pereira dos Santos*, Aline Helena da Silva Cruz

Centro Universitário Araguaia. Departamento de Ciências Biológicas. Goiânia.
Goiás. Brasil. sarahslsantos@gmail.com

RESUMO

A Beta-caseína é uma proteína presente no leite bovino codificada pelo gene CSN2, das suas 13 variantes, entre elas A1 e A2 apresentam maior enfoque em pesquisas, pois o alelo A1 está associado ao desenvolvimento de patologias em seres humanos. A diferença entre os alelos A1 e A2 está na mutação de um par de bases que resulta na substituição de uma prolina por uma histidina. Essa mutação favorece a clivagem enzimática e ocorre a liberação do peptídeo bioativo BCM-7. A Bioinformática é a metodologia que utiliza meios computacionais para explorar e realizar experimentos científicos utilizando dados de plataformas de biotecnologia. O presente estudo utilizou ferramentas de bioinformática na análise da proteína codificada pelo gene CSN2 de *Bos taurus*, correlacionando proteína precursora; proteína secretada, e as variantes A1 e A2. As análises foram realizadas através da utilização de diferentes programas de bioinformática como Blastp, UniProt, Gene Runner e STRING. As análises possibilitaram a predição de peptídeos, sítios de modificações pós-traducionais e caracterização de similaridade e identidade com a Beta-caseína de outros mamíferos. Por meio do software STRING identificou-se a interação proteica da Beta-caseína com proteínas do leite. Por meio deste estudo foi possível verificar a importância e aplicabilidade da bioinformática e da internet no processamento de dados de forma mais rápida e eficiente, contribuindo para a inferência de funções de proteínas caseínas de sequência conhecida, ou mesmo preditas, a partir de sequência de gênicas disponibilizadas em bancos de dados.

Palavras-chave: bioinformática, CSN2, variantes A1 e A2.

ABSTRACT

Beta-casein is a protein present in bovine milk encoded by the CSN2 gene, of its 13 variants, among them A1 and A2 have greater focus in research, as the A1 allele is associated with the development of pathologies in humans. The difference between the A1 and A2 alleles lies in the mutation of a base pair that results in the replacement of a proline by a histidine. This mutation favors enzymatic cleavage and the release of the bioactive peptide BCM-7 occurs. Bioinformatics is the methodology that uses computational means to explore and carry out scientific experiments using data from biotechnology platforms. The present study used bioinformatics tools to analyze the protein encoded by the CSN2 gene of *Bos taurus*, correlating precursor protein; secreted protein, and the A1 and A2 variants. Analyzes were performed using different bioinformatics programs such as Blastp,

UniProt, Gene Runner and STRING. The analyzes enabled the prediction of peptides, sites of post-translational modifications and characterization of similarity and identity with beta-casein from other mammals. The STRING software identified the protein interaction of beta-casein with milk proteins. Through this study, it was possible to verify the importance and applicability of bioinformatics and the internet in data processing in a faster and more efficient way, contributing to the inference of functions of casein proteins of known sequence, or even predicted, from gene sequences available in databases.

Keywords: bioinformatics, CSN2, A1 and A2 variants.

INTRODUÇÃO

Os seres humanos, animais e plantas são compostos por moléculas, que trabalham de forma complexa e organizados para a formação das proteínas que são, em grande parte, responsáveis por exercer funções metabólicas. A Biotecnologia tem como objetivo estudar o funcionamento dessas moléculas e identificar formas de aprimoramento do conhecimento destas⁽¹⁾.

A bioinformática ou biocomputação associa conhecimentos em distintas áreas do conhecimento, reúne dados biológicos, que possibilitam extrair informações precisas para o desenvolvimento de uma pesquisa sem a necessidade de realizar experimentos laboratoriais, graças aos avanços da tecnologia⁽²⁾.

O estudo *in silico* é o nome dado a essa metodologia, que utiliza meios computacionais para explorar e realizar experimentos científicos simulando o que ocorre em um organismo vivo⁽³⁾. Estes estudos computacionais também são relacionados com a nomenclatura bioinformática, que é a pesquisa, desenvolvimento ou aplicação de ferramentas computacionais e abordagens para expandir o uso de dados biológicos, médicos, comportamentais ou de saúde, incluindo aqueles para adquirir, armazenar, organizar, arquivar, analisar ou visualizar tais dados⁽⁴⁾.

A bioinformática surgiu na década de 1980, mesmo antes deste período e até a atualidade, diferentes estudos laboratoriais têm possibilitado a compreensão do funcionamento e estrutura de diferentes moléculas. No caso da Beta-caseína, é uma proteína do grupo das caseínas presente no leite bovino. Das proteínas que compõem o leite, cerca de 80% são caseínas e essas são divididas em quatro grupos:

alfa S1 caseína e alfa S2 caseína (45-55%), Beta-caseína (25-35%) e Kappa-caseína (8-15%)⁽⁵⁾.

O *CSN2* é o gene que codifica a Beta-caseína e de suas 13 variantes (A1, A2, A3, A4, B, C, D, E, F, H1, H2, I, G), se destacam os alelos A1 e A2, por serem os mais comuns a serem encontrados no leite bovino⁽⁶⁾. O leite bovino apresenta aproximadamente quatro vezes mais Beta-caseína quando comparado ao leite humano⁽⁷⁾.

O consumo de leite bovino contendo a variante A1 está associado ao desenvolvimento de patologias em seres humanos, como APLV (Alergia à Proteína do Leite)⁽⁸⁾, e agravamento de outras doenças como: diabetes mellitus tipo 1⁽⁹⁾, Doenças Cardiovasculares⁽¹⁰⁾, Síndrome da Morte Súbita Infantil⁽¹¹⁾, Esquizofrenia e Autismo⁽¹²⁾.

A APLV é a doença mais comum causada pela variante A1 e usualmente é confundida com a intolerância à lactose. É importante ressaltar as diferenças entre elas, de acordo com ALVES e MENDES⁽¹³⁾ a APLV é classificada como uma doença inflamatória acometendo o trato intestinal e também a pele e se manifesta como uma resposta imunológica em combate com algumas proteínas do leite, entre elas, a Beta-caseína. Já a intolerância à lactose é promovida pela deficiência da enzima lactase, na qual sua função é caracterizada pela quebra das moléculas de lactose.

Em relação aos alelos A1 e A2, os animais podem ser genotipicamente homozigotos (A1A1 ou A2A2) ou heterozigotos (A1A2) para o gene da beta-caseína⁽¹⁴⁾. O fator determinante para o genótipo do animal depende da raça e país de origem. Estudos buscam caracterizar a frequência destes genótipos em animais da raça Gir Leiteiro⁽¹⁵⁾, e outras raças como Holstein, Caracu e Sindi⁽¹⁶⁾.

A genotipagem dos animais reprodutores auxilia o pecuarista a aumentar a população de indivíduos homozigóticos para o alelo A2. Essa alternativa possibilita vários benefícios, como: reduzir a frequência de gado com o alelo A1 no rebanho e produção de produtos com certificação de qualidade. Mas é importante ter cuidado em relação à seleção artificial destes animais em busca de melhorias genéticas, pois

o alelo A1 pode estar correlacionado a processos fisiológicos importantes para a fisiologia bovina ainda não identificadas⁽¹⁵⁾.

A diferença entre as variantes A1 e A2 está na mutação por substituição de um nucleotídeo da cadeia de aminoácidos, na qual o nucleotídeo C (citosina) é substituído pelo nucleotídeo A (adenina). Durante a tradução, para formar a proteína, ocorre à substituição de um aminoácido prolina (observada nos indivíduos com alelo A2) por um aminoácido histidina (observada nos indivíduos contendo o alelo A1) na posição 67 da cadeia de aminoácidos⁽¹⁷⁾, para uma cadeia parcial da beta caseína de 209 aminoácidos⁽¹⁸⁾.

Portanto, indivíduos heterozigotos (A1A2) produzem os dois tipos de Beta-caseína⁽¹⁴⁾. As variantes genéticas são denominadas mutações na estrutura primária das caseínas na qual ocorre substituição em um ou mais aminoácidos na cadeia polipeptídica da sequência primária⁽⁷⁾.

A variante A1 favorece a clivagem enzimática (corte), devido à mutação por substituição, a cadeia se torna vulnerável por não ter a prolina impedindo que a clivagem ocorra. Com isso, causa a liberação do peptídeo opióide beta casomorfinina-7 (BCM-7)⁽⁹⁾.

O BCM-7 é um peptídeo bioativo que pode causar no organismo um efeito semelhante à morfina, atuando no sistema nervoso central e sistemas periféricos⁽¹⁹⁾. Está classificado com um peptídeo opióide agonista e desempenha funções farmacológicas⁽²⁰⁾ que inibem ou estimulam substâncias do funcionamento da atividade celular de diferentes sistemas regulatórios no organismo, como por exemplo, no sistema cardiovascular age como um anti-hipertensivo por promover a inibição da enzima conversora de angiotensina (ECA)⁽²¹⁾.

Essa substância ao entrar no trato gastrointestinal humano causa reações como a diminuição da motilidade intestinal, inibição da secreção gástrica, aumento da absorção de água e maior estímulo para contração da vesícula biliar⁽²²⁾. No prisma comercial, as caseínas em geral apresentam importantes propriedades nutricionais, sensoriais e de textura no leite fluido e seus derivados⁽²³⁾.

No que se referem às funções biológicas, as caseínas têm função de transportar cálcio, fosfato e proteína para o recém-nascido, uma vez que a estrutura micelar da caseína sequestra as moléculas de fosfato de cálcio das glândulas mamárias⁽²⁴⁾. O presente estudo utilizou ferramentas de bioinformática na análise da proteína codificada pelo gene *CSN2* de *Bos taurus*, correlacionando proteína precursora; proteína secretada, e as variantes A1 e A2.

METODOLOGIA

Para realização da análise dos alelos A1 e A2, as sequências de aminoácidos destes e da beta caseína de *Bos taurus* foram obtidas no banco de dados do GenBank disponível no National Center for Biotechnology Information (NCBI - <https://www.ncbi.nlm.nih.gov/>). As sequências foram alinhadas através do software gratuito Gene Runner versão 6.5.52x64 Beta (<http://www.generunner.net/>)⁽²⁵⁾ e dos softwares disponíveis no site do NCBI: Blastx, Blastp (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Nas análises também foram utilizados os programas UniProt, KEGG Kyoto Encyclopedia of Genes and Genomes (<https://www.genome.jp/kegg/>) e STRING (<https://www.expasy.org/resources/string>). O blast do UniProt é um recurso de dados essenciais do ELIXIR (<https://www.uniprot.org/blast/>). O ELIXIR reúne as principais organizações de ciências da vida da Europa na gestão e salvaguarda do crescente volume de dados gerados por pesquisas com financiamento público. KEGG é um recurso de banco de dados para a compreensão de funções de alto nível e utilidades do sistema biológico e possibilita a interpretação biológica de conjuntos de dados moleculares em grande escala, como sequências de genoma e metagenoma⁽²⁶⁾. Para predição da interação da proteína da Beta-caseína com outras proteínas do leite foi utilizado o software STRING.

RESULTADOS E DISCUSSÃO

Através da busca da sequência proteica de Beta-caseína de *Bos taurus* no Genbank foram identificadas e selecionadas as sequências apresentadas na Tabela 1. A sequência da beta caseína (Genbank: AAA30431.1, 26-APR-1993) apresenta 224 aminoácidos. Esta sequência de aminoácido foi traduzida a partir da sequência referenciada como sequência completa de Beta-caseína bovina⁽²⁷⁾.

Portanto, com um tamanho diferente de 209 aminoácidos como descrito previamente⁽²⁸⁾⁽²⁹⁾. A sequência parcial de 209 aminoácidos é identificada no Genbank: RF: 1314242A. Esta sequência é referenciada como uma estrutura primária de cDNA para Beta-caseína de bovinos⁽¹⁸⁾, tendo atualização no GenBank em 23 de outubro de 1996.

Tabela 1. Sequência proteica de Beta-caseína de *Bos taurus*.

Beta-caseína [<i>Bos taurus</i>]: GenBank: AAA30431.1 (224 aa)
MKVLILACLVALALARELEELNVPGEIVESLSSEESITRINKKIEKFQSEEQQTEDELQDKIHFP AQTQSLVYFPFGPIHNSLPQNIPPLTQTPVVVPPFLQPEVMGVSKVKEAMAPKHKEMPFPKYPVE PFTESQSLTLTDVENLHLPLLLQSWMHQPHQPLPPTVMFPPQSVLSLSQSKVLPVPQKAVPYQ RDMPIQAFLLYQEPVLPVVRGPFPIIV
Beta-caseína [<i>Bos taurus</i>]: GenBank: PRF: 1314242A (209 aa)
RELEELNVPGEIVESLSSEESITRINKKIEKFQSEEQQTEDELQDKIHFPFAQTQSLVYFP FPGPIHNSLPQNIPPLTQTPVVVPPFLQPEVMGVSKVKEAMAPKHKEMPFPKYPVEPFT ESQSLTLTDVENLHLPLLLQSWMHQPHQPLPPTVMFPPQSVLSLSQSKVLPVPQKAV PYPQRDMPIQAFLLYQEPVLPVVRGPFPIIV
Alelo A1 - Beta-caseína, parcial [<i>Bos taurus</i>]: GenBank: AFR67539.1 (77 aa)
DELQDKIHFPFAQTQSLVYFPFGPIHNSLPQNIPPLTQTPVVVPPFLQPEVMGVSKVKEA MAPKHKEMPFPKYPVEPF
Alelo A2 - Beta-caseína, parcial [<i>Bos taurus</i>]: GenBank: QOW17440.1 (77 aa)
DELQDKIHFPFAQTQSLVYFPFGPIPNSLPQNIPPLTQTPVVVPPFLQPEVMGVSKVKEA MAPKHKEMPFPKYPVEPF

Fonte: GenBank.

A utilização do blast do UniProt para análise da Beta-caseína (AAA30431.1) possibilita uma comprovação *in silico* de que é uma proteína secretada no leite. Essa informação, apesar de parecer redundante, é interessante para análise de proteínas hipotéticas de outros organismos com informações ainda desconhecidas. Análises através da Bioinformática possibilitam informações importantes que podem complementar análises realizadas por outros pesquisadores

como a realizada por Ribadeau Dumas et al que caracterizaram a estrutura primária da Beta-caseína de bovinos⁽³⁰⁾.

A análise de modificações pós-traducionais e eventos de processamento PTM/Processing do UniProt (<https://www.uniprot.org/uniprot/P02666>) mostra que a diferença no tamanho das sequências de AAA30431.1 e PRF: 1314242A está relacionada ao fato da beta caseína apresentar um peptídeo sinal de secreção N-terminal nos 15 primeiros aminoácidos da proteína (Tabela 2).

Peptídeos de sinal são encontrados em proteínas que são direcionadas ao retículo endoplasmático e, eventualmente, destinadas a serem secretadas. A sequência de sinal é geralmente removida na proteína madura; nesses casos, indica-se a forma madura de uma proteína derivada do processamento de uma proteína precursora⁽³¹⁾. A análise através do KEGG (código bta049170) indica a Beta-caseína como uma proteína participante da via de sinalização da prolactina.

Tabela 2. Identificação de peptídeos presentes na proteína Beta-caseína.

Característica	Posição	Descrição	Visão gráfica na sequência	Comprimento
Peptídeo Sinal	1-15	Secreção		15 aa
Cadeia (PRO_0000004470)	16-224	Beta-caseína		209 aa
Peptídeo (PRO_0000292031)	41-45	Casoparan		5 aa
Peptídeo (PRO_0000320153)	113-120	Peptídeo antioxidante		8 aa
Peptídeo (PRO_0000308464)	129-136	Casohipotensina		8 aa

Fonte: KEGG.

As caseínas são as principais constituintes das proteínas do leite e podem ser classificadas em duas famílias: a primeira consiste nas kappa-caseínas e a segunda agrupa as α -s1, α -s2 e caseínas. As caseínas α/β apresentam duas regiões conservadas: um agrupamento de resíduos de serina fosforilados e a sequência de sinal. O padrão de assinatura ou padrão de consenso para esta família de proteínas

(Casein_Alpha_Beta, PS00306) é baseado nos últimos oito resíduos da sequência de sinal: C-L-[LV]-A-x-A-[LVF]-A (<https://prosite.expasy.org/cgi-bin/prosite/prosite-search-ac?PDOC00277#ref1>).

A análise da Beta-caseína [*Bos taurus*] (AAA 30431.1) através do PROSITE também possibilitou a identificação de sítios de fosforilação. Este software descreve domínios de proteínas, famílias e locais funcionais, bem como padrões e perfis associados para identificá-los. O peptídeo sinal de secreção é identificado por retângulo, a chave identifica a assinatura da família da beta caseínas (pfam 00363), os sítios de fosforilação para proteína caseína quinase II são identificados pelos asteriscos (Figura 1).



Figura 1. Alinhamento da sequência completa e parcial da beta caseína, segundo PROSITE.

O alinhamento é o processo ou resultado da correspondência de nucleotídeos ou resíduos de aminoácidos de duas ou mais sequências biológicas para atingir níveis máximos de identidade e, no caso de sequências de aminoácidos, conservação, com o objetivo de avaliar o grau de similaridade e a possibilidade de homologia (FASSLER, COOPER., 2011)⁽⁴⁾.

O alinhamento (Figura 1), através do software Gene Runner, das sequências completa e parcial da beta caseína apresentadas no Quadro 1 possibilitou identificar que a mutação de substituição característica do alelo A1 é referente ao aminoácido 82 da sequência completa da beta caseína disponível no GenBank: AAA 30431.1 e referente ao aminoácido 67 da sequência parcial da beta caseína (proteína madura) disponível como GenBank: PRF: 1314242A, que equivale ao aminoácido 82 na sequência completa. A representação esquemática do alinhamento da sequência

parcial da beta caseína (Figura 2) mostra a diferença entre as variantes A1 e A2, sendo possível identificar o peptídeo correspondente ao BCM-7 (Figura 2).

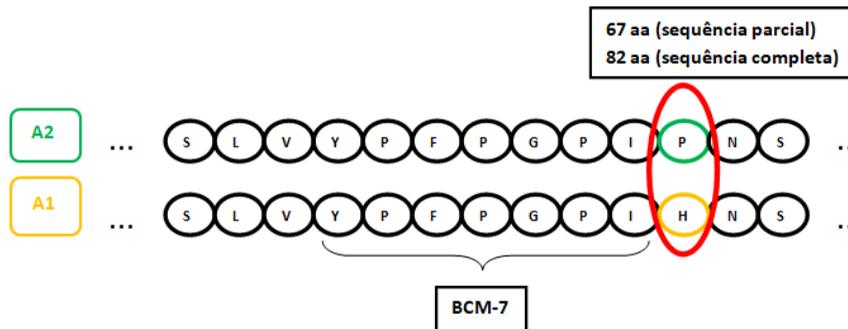


Figura 2. Diferença entre as variantes de beta caseína A1 e A2, segundo GenBank.

A Ferramenta básica de pesquisa de alinhamento local ou BLAST - Basic Local Alignment Search Tool⁽³²⁾ disponível no site do NCBI possibilitou o alinhamento de sequências através do Blastp. Conforme informações disponibilizadas no próprio site (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), o BLAST encontra regiões de similaridade local entre sequências. O programa compara sequências de nucleotídeos ou proteínas a bancos de dados de sequências e calcula a significância estatística das correspondências.

O BLAST pode ser usado para inferir relações funcionais e evolutivas entre sequências, bem como ajudar a identificar membros de famílias de genes e também de proteínas. Alinhamento de sequência de proteína com sequências de proteínas depositadas no GenBank (Figura 3), apresenta similaridade e identidade de sequências com Beta-caseína de outros mamíferos.

						
Organismo	<i>Bos taurus</i>	<i>Bos indicus</i> (<i>Bos taurus indicus</i>)	<i>Bison bison</i>	<i>Bos grunniens</i>	<i>Bos mutus</i>	<i>Bubalus bubalis</i>
Nome popular	Gado bovino	Zebu	Bisão-americano	Iaque Doméstico	Iaque Selvagem	Búfalo-Asiático
GenBank	AAA30431.1	AQY78355.1	XP_010850446.1	AXE74295.1	MXQ92033.1	QHB80269.1
Proteína	Beta-caseína	Beta casein precursor variant-1	Beta-casein <i>Bison bison bison</i>	Beta-casein <i>Bos grunniens</i>	hypothetical protein [Bos mutus]	Casein beta <i>Bubalus bubalis</i>
Score	438 bits (1127)	436 bits (1121)	434 bits (1115)	430 bits (1106)	434 bits (1116)	427 bits (1097)
Expect(e-value)	7e-155	6e-154	4e-153	4e-151	5e-153	3e-150
Identidade	224/224 (100%)	223/224 (99%)	223/224 (99%)	221/224 (99%)	223/224 (99%)	220/224 (98%)
GAPs	0/224 (0%)	0/224 (0%)	0/224 (0%)	0/224 (0%)	0/224 (0%)	0/224 (0%)

Figura 3. Dados comparativos do alinhamento de sequências de *Bos taurus* com outros mamíferos, segundo dados disponíveis em GenBank.

A compreensão entre a identidade da Beta-caseína entre esses diferentes mamíferos se torna interessante principalmente devido ao progresso da agropecuária e a possibilidade de utilização do leite de diferentes formas. Incluindo a utilização do leite de outras espécies no consumo humano, como é o caso do leite de *Bubalus bubalis*. Além de contribuir para despertar o interesse na identificação de diferentes mutações ainda desconhecidas que possam impactar na saúde pública, como é o caso da variante A1.

O e-value (Expectation value ou Expect value) representa o número de alinhamentos diferentes com pontuações equivalentes ou melhores que S (bit score) que se espera que ocorram em uma pesquisa de banco de dados por acaso. Quanto menor o valor E, mais significativo é a pontuação e o alinhamento⁽⁴⁾. O e-value diminui quando o score aumenta.

Bit score ou pontuação de bit, é derivado da pontuação de alinhamento bruto, S, levando em consideração as propriedades estatísticas do sistema de pontuação. Como as pontuações de bits são normalizadas em relação ao sistema de pontuação, elas podem ser usadas para comparar as pontuações de alinhamento de diferentes pesquisas⁽⁴⁾.

A identidade é a extensão em que duas sequências (de nucleotídeos ou aminoácidos) têm os mesmos resíduos nas mesmas posições em um alinhamento, frequentemente expressa como uma porcentagem. A similaridade indica até que

ponto as sequências de nucleotídeos ou proteínas estão relacionadas. A similaridade entre duas sequências pode ser expressa como porcentagem de identidade de sequência e/ou porcentagem de substituições positivas. Além disso, a homologia é a similaridade atribuída à descendência de um ancestral comum. Os componentes biológicos homólogos (genes, proteínas, estruturas) são chamados de homólogos⁽⁴⁾.

É interessante a identificação de identidades entre as caseínas de diferentes espécies, principalmente porque as proteínas do leite também estão integradas, como pode ser observado através da análise com o software STRING (Figura 4). Esta é uma base de conhecimento e a ferramenta de software para identificação de interações proteína-proteína conhecidas e previstas.

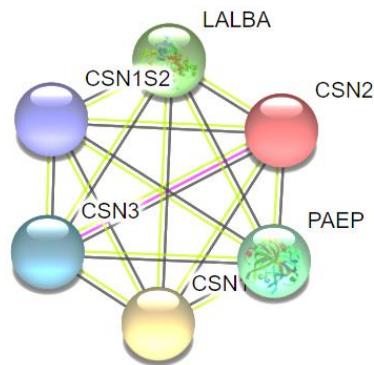


Figura 4. Interação da beta caseína com outras proteínas, segundo dados disponíveis em GenBank.

Na figura 4 é possível observar interação proteica da Beta-caseína (CSN2) com outras proteínas. O CSN2 possui papel importante na determinação das propriedades de superfície das micelas da caseína (224 aa); A alfa-S1-caseína (CSN1S1), apresenta papel importante na capacidade do leite de transportar fosfato de cálcio; A alfa-lactobumina (LABLA) é a subunidade reguladora da síntese da lactose altera a especificidade do substrato da galactosiltransferas; O PAEP, precursor da beta-lactoglobulina, componente primário do soro de leite, ele se liga ao retinol e provavelmente está envolvido no transporte dessa molécula; A kappa-caseína (CSN3), estabiliza a formação de micelas, evitando a precipitação da caseína no leite; Alfa-S2-caseína (CSN1S2), apresenta papel importante na capacidade do leite de transportar fosfato de cálcio.

Estas análises foram possíveis com o uso da biotecnologia. Mesmo diante da ampliação dos recursos tecnológicos ainda se percebe a falta ou dificuldade em obter ou acessar dados para o desenvolvimento de pesquisas, pois é lento o processo de alimentar o banco de dados com informações precisas. Araújo e colaboradores⁽²⁾ destacavam há 10 anos esta necessidade: promover o desenvolvimento de tecnologias, em favor da humanidade, promovendo avanços nos mais variados setores da sociedade.

CONCLUSÃO

As análises de bioinformática permitem uma redução na duração e nas despesas em relação a análises proteicas. A partir do presente trabalho foi possível agrupar informações e disponibilizá-las na língua portuguesa de forma contextualizada e integrada mostrando a importância da continuidade das análises da Beta-caseína, mostrando a similaridade entre a sequência proteica em relação à *Bos taurus* com outras espécies. Além disso, as análises também permitem identificar uma relação dessa proteína com outras proteínas a partir de análises de bioinformática, como já tem sido observada na secreção do leite. Por meio deste foi possível obter um novo olhar, através de análises *in silico*, da Beta-caseína, suas características, estruturas e interações. Verificando-se a importância da bioinformática e da ampliação de estudos no processamento de dados a fim de obter resultados de forma mais rápida e eficiente, contribuindo para a inferência de funções de proteínas de sequência conhecida, ou mesmo preditas, a partir de sequência de genes disponibilizada em bancos de dados.

REFERÊNCIAS

- (1) Borém, A. 2005. A história da Biotecnologia. Lago Norte- Brasília- DF. Revista Biotecnologia, Ciência e Desenvolvimento. Edição Especial-Meio Ambiente. Número.34. p. 10-12.
- (2) Araújo, ND; Farias, RP; Pereira, PB; Figueiredo, FM; Moraes, AMB; Saldanha, LC; Gabriel, JE. A era da bioinformática: seu potencial e suas

implicações para as ciências da saúde. *Estud Biol.* v.30, p.143-8, 2008. DOI: <https://doi.org/10.7213/reb.v30i70/72.22819>

(3) Tamasco, G. Análises in silico: a bioinformática para predição de rotas metabólicas. *Blof do Profissão Biotec.* V. 7, março de 2021. Disponível em: www.profissaobiotec.com.br/analise-in-silico-bioinformatica-predicao-de-rotas-metabolicas

(4) Fassler, J; Cooper.P. BLAST Glossary. Bookshelf ID: NBK62051. National Center for Biotechnology Information, U.S. National Library of Medicine. 2011. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK62051/>. Acesso em: 06/11/2021.

(5) Vercesi Filho, AE.; Camargo, GMF.; Cardoso, DF; El Faro, L; Fernandes, AR; Tonhati, H. Identificação de alelos A1 e A2 para o gene da Beta-caseína na raça Gir Leiteiro. In: SIMPÓSIO BRASILEIRO DE MELHORAMENTO ANIMAL, 9, 2012. ANAIS SBMA. João Pessoa, Paraíba, Brasil.

(6) Rahimi Z, Gholami M, Rahimi Z, Yari K. 2015. Evaluation of beta-casein locus for detection of A1 and A2 alleles frequency using allele specific PCR in native cattle of Kermanshah, Iran. *Biharean Biologist* 9: art.141136

(7) Sgarbieri, VC. "Revisão: Propriedades estruturais e físico-químicas das proteínas do leite." *Brazilian Journal of Food Technology* 8.1 (2005): 43-56.

(8) F. Rancé, G. Dutau, Actualités sur l'exploration et la prise en charge de l'allergie aux protéines du lait de vache (APLV), *Revue Française d'Allergologie*, Volume 49, Supplement 1,2009, Pages S28-S33, [https://doi.org/10.1016/S1877-0320\(09\)72481-9](https://doi.org/10.1016/S1877-0320(09)72481-9). Acesso em out.2020> (<https://www.sciencedirect.com/science/article/pii/S1877032009724819>)<

(9) Elliot RB.; Harris, DP.; Hill, J.P.; Bibby, N.J.; Wasmuth, H.E. Type I (insulindependent) diabetes mellitus and cow milk: casein variant consumption.*Diabetologia*, v. 42, p. 292-296, 1999.

(10) Mclachlan, C.N. Beta-casein A1, ischaemic heart disease mortality, and other illnesses. *Medical Hypotheses*, 2001.

(11) Sun, Z.; Zhang, Z.; Wang, X.; Cade, R.; Elmir, Z.; et al. Relation of betacasomorphin to apnea in sudden infant death syndrome. *Peptides*, v. 24, p. 937-943, 2003.

(12) Woodford, K. A beta-casein, type 1 diabetes and links to other modern illnesses. *IDF Congress*, n. April, p. 1–20, 2008.

- (13) Alves, J.Q.N; Mendes, J.F.R. Consumo dietético e estado nutricional em crianças com alergia à proteína do leite de vaca. *Com. Ciências Saúde*. V. 24, n. 1, p. 65-72, 2013.
- (14) Lima, TCC. Polimorfismo no gene da Beta-caseína em rebanhos zebuínos leiteiros no estado do Rio Grande do Norte. Universidade Federal do Rio Grande do Norte. Macaíba-RN. 2014. Disponível em: https://repositorio.ufrn.br/bitstream/123456789/22746/1/TabattaCristineChavesDeLima_DISSERT.pdf. Acesso em: 17/11/2021
- (15) Sousa, NFMS; Celestino, CS; Gomes, TC; Lacorte, GA; Bastos, RT; Steinberg, RS. Avaliação das frequências alélicas e genotípicas do polimorfismo A1/A2 no gene da Beta-caseína em animais da raça gir leiteiro. VIII Seminário de Iniciação Científica do IFMG – 12 a 14 de agosto de 2019, Campus Ribeirão das Neves.
- (16) Lima, ACJ.; Lara, MAC. Polimorfismo do gene β -caseína em bovinos. *Actas Iberoamericanas de Conservación Animal AICA*, v. 6, p. 280-285, 2015.
- (17) Jaiswal, K.; De, S.; Sarsavan, A. Detection of single nucleotide polymorphism by T-ARMS PCR of cross bred cattle Karan Fries for A1, A2 beta casein types. *International Journal of Biological Sciences Research*, v. 1, p. 18-20, 2014.
- (18) Baev AA, Smirnov IK, Gorodetskiĭ SI. Pervichnaia struktura kDNK beta-kazeína korovy [Primary structure of cDNA for bovine beta-casein]. *Mol Biol (Mosk)*. 1987 Jan-Feb;21(1):255-65. Russian. PMID: 3033475.
- (19) Brantl V, Teschemacher H, Henschen A, Lottspeich F, 1979. Novel opioid peptides derived from casein (b-casomorphins). I. Isolation from bovine casein peptone. *Hoppe Seylers Z Physiol Chem* 360: 1211–1216.
- (20) Tidona, F. et al. Bioactive peptides in dairy. *Italian Journal of Animal Science*, v.8, n.3, p.315-340, 2008.
- (21) SILVA, SV.; MALCATA, FX. Caseins as source of bioactive peptides. *International Dairy Journal*, v.15, n.1, p.1-15, 2005
- (22) Barbosa, MG. Souza, AB. Tavares, GM; Antunes, AEC. . Leites A1 e A2: revisão sobre seus potenciais efeitos no trato digestório. *Segurança alimentar e nutricional*, v. 26, p. 1-11, 2019
- (23) De Kruif, CG. et al. Casein micelles and their internal structure. *Advances in Colloid and Interface Science*, v. 171–172, p. 36– 52, 2012.

- (24) De Kruif, CG. Holt, C. Casein Micelle Structure, Functions and Interactions Advanced Dairy Chemistry, Volume 1, Proteins, Third Edition A, Chapter 5 Ed. P. F. Fox and PLH. McSweeney Kluwer Academic, New York, 2003.
- (25) Edgar, RC. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32(5):1792-1797
- (26) Kanehisa, M, Sato, Y. KEGG Mapper for inferring cellular functions from protein sequences. *Protein Science.* 2020; 29: 28– 35. <https://doi.org/10.1002/pro.3711>
- (27) Bonsing J, Ring JM, Stewart AF, Mackinlay AG. Complete nucleotide sequence of the bovine beta-casein gene. *Aust J Biol Sci.* 1988;41(4):527-37. doi: 10.1071/bi9880527. PMID: 3271384.
- (28) Fontes, P. Tudo o que você precisa saber sobre o leite A2. *Revista Leite integral.* 2019. <https://www.revistaleiteintegral.com.br/noticia/tudo-o-que-voce-precisa-saber-sobre-leite-a2>.
- (29) Sulimova, G, Azari, MA., Rostamzadeh, J, Abadi, MM, Lazebny, O. 2007. κ -casein gene (CSN3) allelic polymorphism in Russian cattle breeds and its information value as a genetic marker. *Russian Journal of Genetics* 43: 73-79.
- (30) Ribadeau Dumas B, Brignon G, Grosclaude F, Mercier JC. Structure primaire de la caséine beta bovine. Séquence complète [Primary structure of bovine beta casein. Complete sequence]. *Eur J Biochem.* 1972 Feb;25(3):505-14. French. doi: 10.1111/j.1432-1033.1972.tb01722.x. PMID: 4557764.
- (31) UniProt. Signal peptide. UniProt Consortium. National Institutes of Health. 2019. Disponível em: <https://www.uniprot.org/help/signal>
- (32) Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990 Oct 5;215(3):403-10. doi: 10.1016/S0022-2836(05)80360-2. PMID: 2231712.