



**UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA
MOLECULAR
CURSO DE GRADUAÇÃO EM BIOTECNOLOGIA**

ANDREW LUNA RODRIGUES

**DESENVOLVIMENTO DE UMA GENOTIPAGEM BIOMARCADORA DE
PRODUTIVIDADE EM REBANHOS DE BOVINOS LEITEIROS NO CEARÁ**

**FORTALEZA
2025**

ANDREW LUNA RODRIGUES

DESENVOLVIMENTO DE UMA GENOTIPAGEM BIOMARCADORA DE
PRODUTIVIDADE EM REBANHOS DE BOVINOS
LEITEIROS NO CEARÁ

Trabalho de conclusão de curso apresentado ao
Curso de Graduação em Biotecnologia do
Centro de Ciências da Universidade Federal do
Ceará, como requisito parcial à obtenção do
grau de Bacharelado em Biotecnologia.

Orientador: Prof. Dr. André Luis Coelho da
Silva.

FORTALEZA

2025

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

R611d Rodrigues, Andrew Luna.

Desenvolvimento de uma genotipagem biomarcadora de produtividade em rebanhos de bovinos leiteiros no Ceará / Andrew Luna Rodrigues. – 2025.

60 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Ciências, Curso de Física, Fortaleza, 2025.

Orientação: Prof. Dr. André Luis Coelho da Silva.

Coorientação: Profa. Dra. Larissa Queiroz Pontes.

1. Kappa-caseína. 2. Rebanhos leiteiros. 3. qPCR. 4. Marcadores moleculares. I. Título.

CDD 530

ANDREW LUNA RODRIGUES

DESENVOLVIMENTO DE UMA GENOTIPAGEM BIOMARCADORA DE
PRODUTIVIDADE EM REBANHOS DE BOVINOS LEITEIROS NO CEARÁ

Trabalho de conclusão de curso apresentado ao
Curso de Graduação em Biotecnologia do
Centro de Ciências da Universidade Federal do
Ceará, como requisito parcial à obtenção do
grau de Bacharelado em Biotecnologia.

Aprovada em: ___/___/____.

BANCA EXAMINADORA

Prof. Dr. André Luis Coelho da Silva (Orientador)
Universidade Federal do Ceará (UFC)

Dra. Larissa Queiroz Pontes
Bio-Manguinhos/Fiocruz-CE

Prof. Dra. Cláudia Goulart de Abreu
Universidade Estadual Vale do Acaraú (UVA)

Aos meus pais, colegas de trabalho, amigos e a todos que me ajudaram direta ou indiretamente nestes árduos anos de curso.

AGRADECIMENTOS

Ao meu pai, André, por sempre estar presente nos momentos de incerteza quanto ao futuro e outras questões da vida para me aconselhar com suas palavras de sabedoria e experiência. A minha mãe, Sandra, por sempre me apoiar com amor incondicional, ajudando sempre que possível em todo momento de necessidade e desesperança. A minha irmã, Andressa, com a qual sempre pude contar para o que der e vier.

Aos meus familiares, como avós, tios e primos, pelas lembranças divertidas e pelos momentos de descontração.

Aos meus colegas de trabalho no laboratório, Luis, Kethlin, Vanessa e Gustavo, os quais sempre me apoiaram nos momentos difíceis, dando-me conselhos, orientações e tentando alegrar meu dia com conversas e atividades divertidas, mantendo minha sanidade durante essa etapa final do curso.

À equipe da TaqMol Diagnósticos, por terem me acolhido de braços abertos, fornecendo experiências esclarecedoras e conhecimentos importantes. Ao Wallady, por ser uma pessoa incrível, gentil, compreensível e didática, ajudando sempre nos momentos tensos da rotina trabalhosa. À Talita, mulher forte, exemplar, cheia de conhecimento e conselhos para dar não só sobre o trabalho, mas também sobre a vida. Ao Marcus Rafael, por ser uma das pessoas mais doces que já conheci, atencioso, prestativo e paciente com todas as minhas dúvidas. À Larissa, minha supervisora, pela orientação, pelos ensinamentos e toda experiência transmitida a mim nesses últimos anos, pela confiança no meu potencial e na minha capacidade.

Ao Professor Doutor André Luis Coelho da Silva, pela orientação e pelo exemplo de profissional incrível e dedicado, amante da área de Biotecnologia. A toda a equipe do Grupo de Biotecnologia Molecular e Estrutural (GBME).

A todos os amigos que conheci no decorrer do curso, por estarem ao meu lado, apoiando-me nos momentos de dificuldade e desespero.

À Universidade Federal do Ceará, por todas as oportunidades valiosas, pelo suporte financeiro e estrutura que permitiram a minha formação em Biotecnologia.

À Funcap, por fomentar minhas pesquisas e auxiliar na minha graduação.

RESUMO

Com o crescente conhecimento das características genéticas dos animais de produção, tem-se notado um aumento no interesse em estratégias genéticas que utilizem marcadores moleculares para análise de características de interesse econômico para a indústria. O gene da *kappa*-caseína está relacionado a características como produção de leite, teor de proteínas e gorduras e coagulação das micelas de caseína, sendo, assim, um marcador molecular de interesse para a indústria de queijos e derivados do leite. Atualmente, apesar de já existirem diversos testes de genotipagem para outros marcadores moleculares relacionados ao leite, como o gene de β -caseína, o mercado de testes moleculares de genotipagem para o gene de κ -caseína ainda não é muito explorado no Ceará. Tendo em vista isso, o presente trabalho, realizado em parceria com a empresa TaqMol Diagnósticos, buscou desenvolver um teste molecular baseado em técnicas de genotipagem por PCR quantitativa em tempo real (qPCR) para permitir o melhor manejo de rebanhos de bovinos leiteiros e a seleção de indivíduos com características produtivas economicamente desejáveis. Para isso, elaborou-se um protocolo de PCR em tempo real buscando otimizar a reação e os parâmetros de ciclagem. As amostras utilizadas foram obtidas através da empresa TaqMol Diagnósticos e tiveram gDNA extraído a partir de kit comercial. Para amplificação do gene de κ -caseína, foram desenhados e testados *primers* e sondas de hidrólise específicos de propriedade da empresa. O produto de PCR foi clonado em vetor de clonagem pGEM®-T Vector para obtenção de controle positivo. O sequenciamento de Sanger do gene clonado das colônias transformadas apontou homozigose AA para a amostra clonada. O teste de qPCR desenvolvido, dada à sua especificidade e sensibilidade devido ao uso de sondas de hidrólise específicas para a identificação dos SNPs do gene CSN3, apresenta fortes indícios de se uma técnica viável e confiável para genotipagem da κ -caseína.

Palavras-chave: *Kappa*-caseína; qPCR; Marcador Molecular; Rebanho leiteiro.

ABSTRACT

With the growing knowledge of the genetic characteristics of farm animals, there has been an increase in interest in genetic strategies that use molecular markers to analyze traits of economic interest to the industry. The kappa-casein gene is related to traits such as milk production, protein and fat content, and coagulation of casein micelles, and is therefore a molecular marker of interest to the cheese and dairy industry. Currently, although there are several genotyping tests for other molecular markers related to milk, such as β -casein gene, the market for molecular genotyping tests for κ -casein gene is still not widely explored in Ceará. In view of this, the present work, carried out in partnership with the company TaqMol Diagnósticos, sought to develop a qPCR-based genotyping test to allow better management of dairy cattle herds and the selection of individuals with economically desirable productive traits. For this purpose, a real-time PCR protocol was developed to optimize the reaction and cycling parameters. The samples used were obtained from TaqMol Diagnósticos and had gDNA extracted from a commercial kit. Specific primers and hydrolysis probes proprietary to the company were designed and tested to amplify the κ -casein gene. The PCR product was cloned into pGEM®-T Vector to obtain a positive control. Sanger sequencing of the cloned gene from the transformed colonies indicated AA homozygosity for the cloned sample. The qPCR test developed, given its specificity and sensitivity due to the use of specific hydrolysis probes for the identification of SNPs in the CSN3 gene, presents strong evidence of being a viable and reliable technique for κ -casein genotyping.

Keywords: *Kappa-casein; qPCR; Molecular marker; Dairy herd.*

LISTA DE FIGURAS

Figura 1 - Composição bioquímica geral do leite bovino.....	16
Figura 2 - Representação da estrutura da micela de caseína.....	18
Figura 3 - Estrutura do gene CSN3 no cromossomo bovino 6.....	19
Figura 4 - Classificação de marcadores moleculares e exemplos.....	21
Figura 5 - Mecanismo de funcionamento da PCR.....	23
Figura 6 - Procedimento detalhado do sequenciamento de Sanger.....	25
Figura 7 - Mecanismos de fluorescência de corantes intercalantes.....	26
Figura 8 - Mecanismo de fluorescência de sondas de hidrólise.....	27
Figura 9 - Representação do funcionamento do anelamento das sondas com diferentes fluoróforos aos alelos A e B.....	28
Figura 10 - Amplification plot da corrida de teste do par de primers.....	40
Figura 11 - Gráfico da curva de melt dos amplicons gerados na corrida de teste dos primers.....	41
Figura 12 - Gráfico de amplificação da sonda de hidrólise para o alelo A para as amostras 1, 2 e 3.....	42
Figura 13 - Gráfico de amplificação da sonda de hidrólise para o alelo B para as amostras 1, 2 e 3.....	42
Figura 14 - Allelic discrimination plot da corrida de teste das sondas de hidrólise.....	43
Figura 15 - Análise da corrida eletroforética em gel de agarose 2% para validação da amplificação correta do gene de kappa-caseína.....	45
Figura 16 - Análise da corrida eletroforética de gel de agarose 2% para validação da integridade do DNA purificado.....	46
Figura 17 - Análise eletroforética em gel de agarose 2% do DNA das colônias transformadas.....	48
Figura 18 - Alinhamento obtido após sequenciamento de Sanger do fragmento de κ -caseína da amostra sequenciada.....	48
Figura 19 - Alinhamento da sequência de aminoácidos de C10 com sequência de referência.....	49
Figura 20 - Gráfico de amplificação da sonda para o alelo A do teste de genotipagem de	

10 amostras.....	50
Figura 21 - Gráfico de amplificação da sonda para o alelo B do teste de genotipagem de 10 amostras.....	50
Figura 22 - Gráfico de discriminação alélica de 10 amostras de gDNA bovino de fazendas variadas.....	51

LISTA DE QUADROS

Quadro 1 - Diferença na sequência de aminoácidos entre as variantes alélicas A e B do gene CSN3.....	19
Quadro 2 - Exemplo da correlação entre os sinais de fluorescência e as sequências em uma amostra qualquer.....	29
Quadro 3 - Relação de reagentes, suas concentrações e volumes adicionados à reação..	
34	
Quadro 4 - Reagentes, suas concentrações finais e volumes adicionados à reação.....	34
Quadro 5 - Reagentes, concentrações finais e volumes adicionados à reação.....	35
Quadro 6 - Correlação das amostras com o nível de confidence e a call designada pelo software.....	43

LISTA DE ABREVIATURAS E SIGLAS

AFLP	<i>Amplified Fragment Length Polymorphism</i>
CE	Ceará
Cq	<i>Quantification cycle</i>
DArT	<i>Diversity Arrays Technology</i>
DNA	<i>Deoxyribonucleic Acid</i>
FAO	<i>Food and Agriculture Organization</i>
gDNA	DNA genômico
IBGE	Instituto Brasileiro de Geografia e Estatística
IPTG	Isopropil β -D-1-tiogalactopiranosídeo
LB	<i>Luria-Bertani</i>
NGS	<i>Next Generation Sequencing</i>
NTC	<i>No Template Control</i>
pb	Pares de base
PCR	<i>Polymerase Chain Reaction</i>
qPCR	<i>Quantitative Polymerase Chain Reaction</i>
RAPD	<i>Random Amplified Polymorphic DNA</i>
RFLP	<i>Restriction Fragment Length Polymorphism</i>
RPM	Rotações por minuto
SNP	<i>Single Nucleotide Polymorphism</i>
SSR	<i>Single Sequence Repeat</i>
X-Gal	5-bromo-4-cloro-3-indoxil- β -D-galactopiranosídeo

LISTA DE SÍMBOLOS

a	Alfa
β	Beta
C	Celsius
g	Grama
°	Grau
κ	<i>Kappa</i>
L	Litro
®	Marca registrada
™	Marca registrada
μ	Micro
M	Molar
%	Porcentagem

SUMÁRIO

1 INTRODUÇÃO.....	14
1.1 Composição bioquímica do leite.....	16
<i>1.1.1 Fração proteica do leite e sua importância para a indústria de derivados.....</i>	17
1.2 Caseínas e micelas de caseína.....	17
1.3 K-caseína e seus polimorfismos.....	18
1.4 Marcadores moleculares.....	20
1.5 Genotipagem e seu uso para detecção de SNPs.....	22
<i>1.5.1 Reação em cadeia da polimerase (PCR).....</i>	22
<i>1.5.2 Sequenciamento de Sanger.....</i>	24
<i>1.5.3 PCR quantitativa (qPCR) em tempo real.....</i>	26
<i>1.5.4 Abordagens atuais para genotipagem em rebanhos leiteiros.....</i>	29
2 OBJETIVOS.....	31
3 MATERIAIS E METODOLOGIA.....	32
3. 1 Materiais.....	32
<i>3.1.1 Reagentes.....</i>	32
<i>3.1.2 Kits de extração.....</i>	32
<i>3.1.3 Organismos utilizados.....</i>	32
3.2 Obtenção das amostras de sangue total bovino.....	32
3.3 Extração de gDNA.....	33
3.4 Desenho dos primers e sondas de hidrólise.....	33
3.5 PCR para teste dos primers e sondas de hidrólise.....	33
3.6 PCR para amplificação do fragmento de κ-caseína.....	35
3.7 Purificação do produto de PCR.....	36
3.8 Clonagem do produto de PCR purificado.....	36
<i>3.8.1 Ligação do produto de PCR purificado ao vetor pGEM.....</i>	36
<i>3.8.2 Transformação genética.....</i>	37
<i>3.8.3 PCR das colônias transformadas.....</i>	38
<i>3.8.4 Pré-inóculo das colônias transformadas.....</i>	38

3.8.5 Miniprep das colônias transformadas.....	38
3.8.6 Sequenciamento de Sanger.....	38
3.9 Testes de genotipagem com controle positivo da variante A.....	39
4 RESULTADOS E DISCUSSÃO.....	40
4.1 PCRs para teste dos primers e sondas de hidrólise.....	40
4.2 PCR para amplificação do fragmento de κ-caseína.....	44
4.3 Purificação do produto de PCR.....	45
4.4 Clonagem do produto de PCR purificado.....	46
4.4.1 Ligação do produto de PCR purificado ao vetor pGEM.....	46
4.4.2 PCR das colônias transformadas.....	47
4.4.3 Sequenciamento de Sanger.....	48
4.5 Testes de genotipagem com controle positivo da variante A.....	49
5 CONCLUSÃO.....	52
6 REFERÊNCIAS.....	53

1 INTRODUÇÃO

A atividade leiteira é uma das principais atividades econômicas do Brasil, uma vez que é uma grande fonte de renda e empregos, além de estar presente em 98% dos municípios do país, envolvendo mais de 1 milhão de produtores no campo, dentre os quais a maioria de pequenas e médias propriedades, além de gerar outros milhões de empregos nos demais segmentos da cadeia produtiva do leite (Rocha; Carvalho; Resende; 2020).

O leite, bem como seus derivados, é um dos alimentos mais consumidos no Brasil, com dados apontando um consumo per capita de 147 kg de leite por ano (FAO, 2022). Em 2023, a produção de leite nacional ultrapassou a marca dos 35,3 bilhões de litros, com as regiões Sul e Sudeste representando, respectivamente, 34% e 33% de toda a produção do país (CILEITE, 2025). Apesar dessa concentração da produção nessas duas regiões, o Nordeste tem apresentado notável crescimento na sua capacidade produtora, configurando-se como a 3^a região que mais produziu leite no país, com um crescimento de aproximadamente 10,7% em 2023 quando comparado a 2022, atingindo uma produção de mais de 6,2 bilhões de litros de leite (Vasconcelos, 2024).

Na região Nordeste, em 2023, um dos estados que mais se destacou foi o Ceará, o qual apresenta setor leiteiro de destaque no Brasil, configurando-se como o 9º maior produtor de leite no país e o 3º maior produtor da região Nordeste, com participação de 3,2% na produção leiteira nacional (IBGE, 2023). A indústria leiteira do Ceará produziu 53,6 milhões de litros de leite a mais em 2023 no comparativo com 2022, representando um aumento de 14,5% (Nobre, 2024). Esse desempenho foi impulsionado por uma série de fatores, incluindo políticas públicas com o objetivo de modernizar a cadeia produtiva do leite por meio de inovações tecnológicas, como a ordenha mecanizada, e a organização dos produtores em cooperativas, permitindo a eles a participação em programas que visam fortalecer o setor produtivo local, como o Programa para Aquisição de Alimentos (PAA Leite) (Porto, 2025).

O melhoramento genético de rebanhos é uma estratégia essencial para impulsionar a produção de leite no Ceará, tanto para as grandes indústrias quanto para a agricultura familiar. Na Portaria Interministerial MDA/MDS/MAPA nº 5, de 30 de agosto de 2024 (Brasil, 2024), que institui a estratégia de desenvolvimento da produção de leite na agricultura familiar, reforça essa necessidade ao incluir, em seu Art. 3º, inciso III, o estímulo ao melhoramento genético adequado aos diferentes biomas. No Ceará, isso significa investir em tecnologias como inseminação artificial e seleção de reprodutores que combinem alta produção e qualidade do leite, com adaptação dos animais ao clima semiárido (McManus *et al.*, 2022).

Além disso, a melhoria da composição do leite pode agregar valor ao produto, beneficiando tanto os produtores quanto a indústria local.

A influência dos marcadores moleculares em características de interesse econômico para indústrias específicas, como volume total de leite produzido, teor de proteínas e gorduras, já é conhecida na literatura (Silva; Demuner, 2024). Alguns marcadores moleculares, como o gene da κ -caseína estão relacionados a características como produtividade, composição e propriedades organolépticas do leite, apontando para a possibilidade desse marcador ser utilizado como marcadores genéticos para manejo dos rebanhos, seleção e cruzamento de animais de gado leiteiro (Barbosa *et al.*, 2019; Velmato *et al.*, 2020; Albazi *et al.*, 2021; Bangar *et al.*, 2021).

Dentre várias raças de bovinos especializadas para a produção leiteira, como Girolando, Holstein-Friesian e Jersey, observa-se a predominância do alelo A do gene de κ -caseína, e, em contrapartida, o alelo B, que está associado a maiores teores de proteínas e gordura e a formação de coágulos mais firmes, possui frequência relativamente baixa na maioria das raças, principalmente em Gir, Holstein e Girolando (Fox *et al.*, 2017; Bangar *et al.*, 2021). Azevedo *et al.* (2008) reportaram que vacas com o genótipo BB apresentam um rendimento de queijo 10% maior em comparação com o leite de vacas com o genótipo AA, o que pode impactar sobremaneira no retorno econômico na produção de laticínios.

A seleção de reprodutores tradicionalmente é feita pelo teste de progênie, onde os touros são selecionados pela produtividade (produção de leite, teor de gordura e de proteína do leite, etc) de suas filhas. Este método ainda é amplamente utilizado, mas possui muitas limitações, como o longo tempo necessário para obter dados confiáveis (uma vez que depende da lactação das filhas) e a incapacidade de prever o desempenho de animais jovens. Além disso, o teste de progênie é caro e limitado a características mensuráveis apenas em fêmeas (Perez *et al.*, 2018).

Os métodos de seleção baseados no uso de marcadores moleculares, quando comparados aos métodos convencionais, mostram-se como mais vantajosos, uma vez que permitem a obtenção de estimativas de valor genético mais acuradas para animais em estágios precoces de desenvolvimento, e assim, consequentemente, a diminuição do tempo entre gerações, ocasionando num ganho genético maior por geração (Perez *et al.*, 2018).

Dessa forma, torna-se bastante atrativo o uso de ferramentas e técnicas biotecnológicas que utilizem esses marcadores para auxiliar na seleção assistida de animais portadores de caracteres genéticos economicamente favoráveis. Essas ferramentas possibilitam uma agregação mais facilitada de valor à cadeia produtiva do leite, permitindo, dessa forma, a

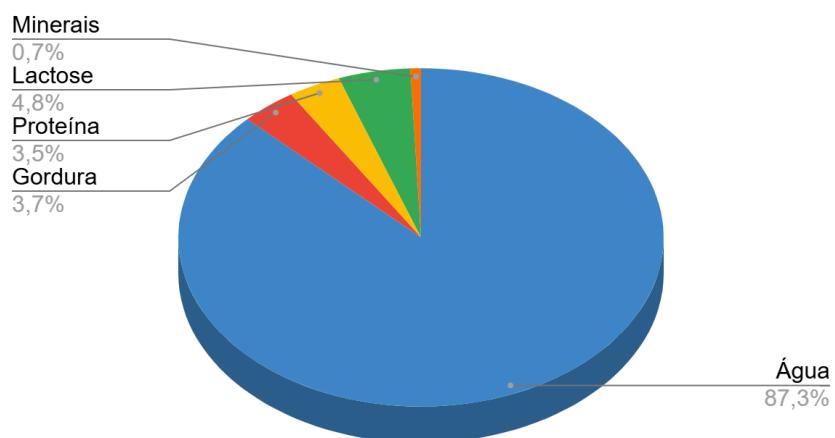
otimização da produtividade e da qualidade dos produtos lácteos, com a vantagem adicional de reduzir o intervalo de gerações, possibilitando maior progresso genético em menor tempo (Barbosa *et al.*, 2019).

1.1 Composição bioquímica do leite

O leite configura-se como um fluido de composição complexa, uma vez que é constituído por centenas de diferentes espécies de moléculas, com composição bioquímica geral formada por água (87,3%), carboidratos (4,8%), gorduras (3,7%), proteínas (3,5%), vitaminas, entre outros (Figura 1) (Goulding; Fox; O'Mahony., 2020). Possui a função de nutrir o recém-nascido até a etapa de desmame, sendo, dessa forma, rico em todos os nutrientes necessários para o desenvolvimento do organismo, como sais minerais, aminoácidos, hormônios, fatores de crescimento e agentes protetores (Fox *et al.*, 2015; Krishna *et al.*, 2021).

Entre seus constituintes, sua fração proteica é de grande importância por seus valores nutricionais e por seus atributos funcionais, além de ter destaque na fabricação de derivados do leite, uma vez que confere a ele as características fisico-químicas (espumação, gelificação, espessamento, etc) que o fazem ser utilizado na composição de diversos produtos, alimentando as demandas de diversos mercados globais (Goulding; Fox; O'Mahony, 2020).

Figura 1 - Composição bioquímica geral do leite bovino.



Fonte: Adaptado de Fox *et al.* (2015). Composição média dos diferentes constituintes do leite bovino: 87,3% de água leite e 12,7% de sólidos solúveis (0,7% minerais, 4,8% lactose, 3,5% proteínas e 3,7% gorduras). Essa composição

do leite varia dependendo da espécie do animal, da raça, da idade e de fatores ambientais.

1.1.1 Fração proteica do leite e sua importância para a indústria de derivados

A fração proteica é constituída principalmente por duas famílias de proteínas, com cerca de 80% sendo representados pelas caseínas e 20% por proteínas do soro do leite (predominantemente alfa-lactoalbumina e beta-lactoglobulina) (Goulding; Fox; O'Mahony, 2020). As propriedades físico-químicas das proteínas lácteas, como emulsificação, gelificação e capacidade de retenção de água, são de grande interesse para a indústria de alimentos, uma vez que possuem papel relevante na fabricação de uma diversidade de produtos alimentícios, como queijos, manteigas, iogurtes, sorvetes etc.

A maioria dos produtos alimentícios consistem em emulsões e espumas, e as proteínas lácteas possuem a capacidade de estabilizar esses produtos graças às suas propriedades anfifílicas, permitindo-as adsorver em sistemas termodinamicamente instáveis (Zhou *et al*, 2021). Algumas propriedades mais importantes para a indústria de derivados, como estabilidade em altas temperaturas e coagulação na presença de coalho são devido à parcela composta pelas caseínas, principalmente a *kappa*-caseína, razão pela qual recebem destaque sobre o impacto das proteínas do leite no seu processamento em alimentos. (Agregán *et al*, 2021).

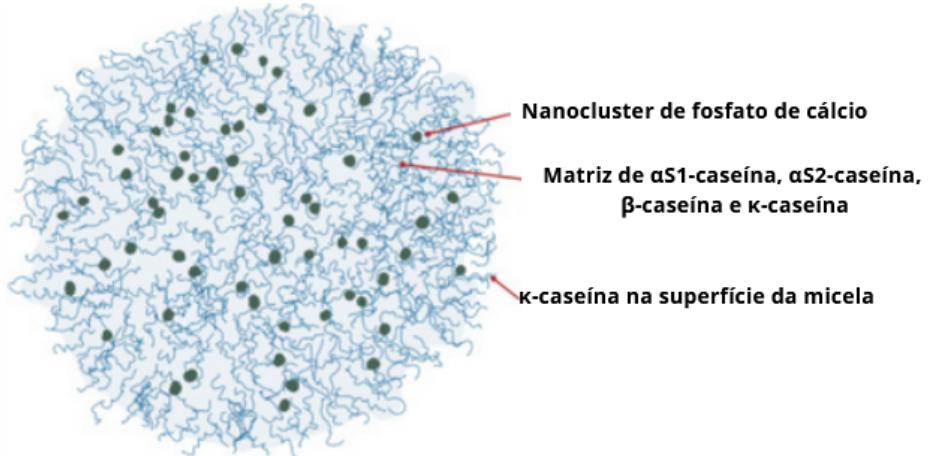
1.2 Caseínas e micelas de caseína

O grupo das caseínas, composto por 4 tipos de caseínas individuais (α_1 -, α_2 -, β -, e κ -caseína) são uma família de fosfoproteínas insolúveis, sintetizadas pelas células epiteliais da glândula mamária durante a lactação e secretadas no leite de todos os mamíferos. Elas são responsáveis pelo transporte do fosfato de cálcio para o animal recém-nascido, fornecendo minerais essenciais para o desenvolvimento dos ossos e dentes (Horne, 2020).

No leite, essas proteínas se encontram na forma de micelas de caseína, as quais são compostas majoritariamente por proteínas e uma parcela de minerais coletivamente referidos como fosfato de cálcio micelar (MCP) (Sadiq; Gill; Chadrapala, 2021). O MCP é composto principalmente de cálcio e fosfato, apresentando também traços citrato, magnésio e outros minerais. As micelas de caseína (Figura 2) são estruturas supramoleculares formadas pelos 4

tipos de caseína em interação com *nanoclusters* de fosfato de cálcio, os quais atuam como “nós” da estrutura que conectam as proteínas de caseína (Hettinga; Bijl, 2022).

Figura 2 - Representação da estrutura da micela de caseína



Fonte: Adaptado de Hettinga; Bijl (2022). A estrutura interna da micela de caseína é composta por uma matriz formada por 4 tipos de caseína, α s1-, α s2-, β -, e κ -caseína, em união à *nanoclusters* de fosfato de cálcio. A κ -caseína encontra-se na superfície das micelas auxiliando na estabilização da estrutura no leite.

O processamento e transformação do leite em produtos como iogurte e queijo dependem principalmente das propriedades interativas dessas estruturas de micelas, apesar dos outros componentes do leite, como as proteínas do soro, os lipídeos e os minerais também contribuírem para o processamento lácteo em alimentos (SMITH *et al.*, 2020).

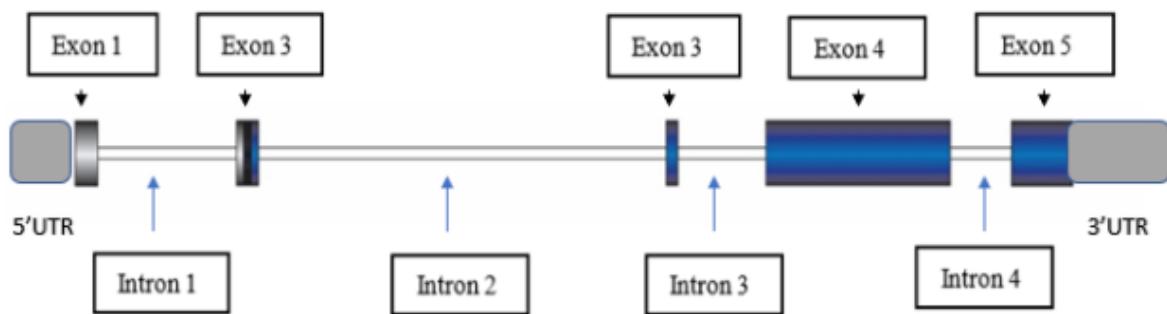
Dentre as diferentes caseínas constituintes do leite, a κ -caseína tem grande importância industrial, pois está diretamente relacionada com a formação, a estabilidade e agregamento das micelas de caseína, cuja desestabilização e formação de gel é o principal processo na fabricação de queijos (BRITTEN *et al.*, 2021).

1.3 K-caseína e seus polimorfismos

A produtividade e o rendimento do leite dos rebanhos dependem da expressão de diversos genes, a qual é influenciada por fatores ambientais e genéticos (Asmarasari *et al.*, 2021; Hani *et al.*, 2021; Albazi *et al.*, 2023). Dentre esses genes está o que codifica a proteína κ -caseína (CSN3), localizado no cromossomo bovino 6, e que é composto de uma sequência de

13 mil pares de base divididas em 5 exons (Figura 3), estando localizadas no exon 4 a maioria das sequências codificadoras das proteínas da κ -caseína (Asmarasari *et al.*, 2021).

Figura 3 - Estrutura do gene CSN3 no cromossomo bovino 6



Fonte: Asmarasari *et al.* (2021). O gene CSN3 está presente no cromossomo bovino 6. A maioria das sequências codificadoras das proteínas da κ -caseína estão localizadas no exon 4.

Foram encontradas até agora cerca de 9 variantes do gene da κ -caseína, contudo, na natureza, as variantes alélicas A e B são as mais comuns e importantes nos rebanhos de bovinos leiteiros (Asmarasari *et al.*, 2021; Deb *et al.*, 2014). A variante alélica B é semelhante à variante alélica A, diferindo-se apenas em dois resíduos de aminoácidos (Quadro 1): na posição 136 das sequências o aminoácido treonina é substituído por isoleucina enquanto na posição 148 ao ácido aspártico é substituído por uma alanina (Pishchan; Sylychenko, 2021; Suprovich; Mokhnachova, 2017).

Quadro 1 - Diferença na sequência de aminoácidos entre as variantes alélicas A e B do gene CSN3.

Variantes	Posição dos aminoácidos			
	135	136	148	155
A	Thr (T)	Thr (T)	Asp (D)	Ser (S)
B	Thr (T)	Ile (I)	Ala (A)	Ser (S)

Fonte: Adaptado de Habib *et al.* (2001). A diferença entre as variantes A e a variante B, ocorre em dois resíduos de aminoácidos específicos: nas posições 136 (treonina \rightarrow isoleucina) e 148 (ácido aspártico \rightarrow alanina).

A influência destes polimorfismos na composição do leite e nas suas propriedades físico-químicas (capacidade de emulsificação, de gelificação e de espumação), bem como a possibilidade de serem utilizados como marcadores moleculares para análise desses mesmos parâmetros já é bem documentada na literatura (Ceriotti *et al.*, 2004; Azevedo *et al.*, 2008; Volkandari; Indriawati; Margawati, 2017; Pishchan; Sylychenko, 2021). A variante A está associada a aumento da produção em volume de leite em vacas leiteiras, enquanto a variante B está associada a produção de leite com maior teor de gordura e proteínas, bem como a características desejáveis à produção de queijos duros, como menor tempo de coagulação do leite e formação de coágulos mais firmes (Huiam; Eltahir; Hamza, 2018; Pishchan; Sylychenko, 2021).

Segundo Fox et al (2017), o fator mais impactante no rendimento da produção de queijos é a composição do leite, mais especificamente nas concentrações de caseínas e gorduras. O leite produzido por vacas do genótipo BB do gene de k-caseína apresenta coagulação superior do coágulo e características de produção de queijo melhores do que o leite do genótipo AA. Isso se deve ao maior teor de caseínas na sua composição, especificamente de k-caseínas, bem como micelas menores e menor carga negativa.

1.4 Marcadores moleculares

Marcador molecular é todo fenótipo molecular oriundo de um gene expresso, ou de um segmento específico de DNA em regiões expressas ou não do genoma (Ferreira; Grattapaglia, 1998). Quando comparado aos outros tipos de marcadores genéticos, os marcadores moleculares apresentam uma série de vantagens que os configuram como um dos marcadores mais confiáveis disponíveis para determinação de características fenotípicas de interesse (Al-Samarai; Al-Karaz, 2015). Dentre as suas vantagens, tem-se o fato de não serem afetados por fatores ambientais e fisiológicos e a sua quantidade ilimitada disponível ao longo do genoma (Yang *et al.*, 2013).

Os marcadores moleculares são divididos em três grupos (Figura 4), dependendo do princípio em que são baseados: 1) baseados em hibridização, também conhecidos como de primeira geração, como *Restriction Fragment Length Polymorphism* (RFLP), 2) baseados em PCR, também conhecidos como de segunda geração, como *Random Amplified Polymorphic DNA* (RAPD), *Amplified Fragment Length Polymorphism* (AFLP), *Simple Sequence Repeat* (SSR) e 3) baseados em sequenciamento, também conhecidos como terceira geração, como

Single Nucleotide Polymorphism (SNP), *Next Generation Sequencing* (NGS) e *Diversity Arrays Technology* (DArt) (Grover; Sharma, 2016).

O sistema de marcador genético baseado na hibridização é embasado no pareamento de bases complementares, utilizando fragmentos pequenos de DNA como sondas para revelar polimorfismos apenas nas sequências homólogas a essa sonda (Turchetto-Zolet *et al.*, 2017). O principal exemplo desse sistema é a técnica RFLP, a qual consiste na clivagem do DNA por enzimas de restrição, gerando fragmentos de diversos tamanhos (Oliveira *et al.*, 2021). Os fragmentos gerados são, então, separados por eletroforese com base no tamanho deles e, em seguida, são hibridizados com sondas específicas que permitem a detecção de polimorfismos por meio de autorradiografia.

Com o advento da PCR, surgiram sistemas de marcadores moleculares mais sensíveis, específicos e rápidos, permitindo a amplificação de uma grande quantidade de uma sequência específica de DNA sem necessidade de clonagem, começando com poucas quantidades de material genético (Turchetto-Zolet *et al.*, 2017). A técnica de RAPD, um exemplo básico desse sistema de marcador molecular, consiste no uso de *primers* de sequência curta e aleatória para amplificação de diversas regiões do genoma de forma não específica. A diferença no padrão das bandas entre os indivíduos indicam os polimorfismos.

Os marcadores moleculares baseados em sequenciamento, como os SNPs, consistem na leitura direta da sequência de DNA alvo para identificação de variações genéticas entre indivíduos (Grover; Sharma, 2016). Essa leitura direta é feita por técnicas de sequenciamento de DNA, como Sanger e NGS, as quais permitem a detecção de alterações tão pequenas quanto a mutação de um nucleotídeo, ou a deleção e inserção de bases na sequência de DNA.

Figura 4 - Classificação de marcadores moleculares e exemplos.



Fonte: Adaptado de Grover; Sharma (2016). Na figura estão representados os 3 grupos em que são divididos os marcadores moleculares baseados no seu método de detecção e alguns exemplos de cada.

1.5 Genotipagem e seu uso para detecção de SNPs

Genotipagem é o processo utilizado para detecção de características de interesse em organismos por meio da identificação de pequenas variações no DNA de indivíduos da mesma espécie. Para o estudo genético de rebanhos, a técnica pode ser utilizada não só para estratégias de manejo de saúde animal e diagnósticos de doenças, mas também para seleção genômica de animais com características de interesse econômico, como genótipos de κ -caseína que determinam a produção de maior volume de leite em rebanhos de bovinos leiteiros (Embrapa Recursos Genéticos e Biotecnologia, 2021).

Os marcadores do tipo SNP são caracterizados por mudanças de um nucleotídeo em uma dada sequência de DNA, podendo ter duas possibilidades de nucleotídeo em uma dada posição. Para que uma posição de base com alternativas de sequência no DNA seja considerada um SNP, é preciso que o alelo menos frequente tenha uma frequência de 1% ou maior (Vignal *et al.*, 2002).

Diversas pesquisas têm mostrado o impacto desses marcadores nas funções biológicas, podendo ser utilizados em variadas áreas do conhecimento, como estudos evolutivos, filogenéticos, ecológicos e no melhoramento genético de animais e plantas (Turchetto-Zolet *et al.*, 2017). Em regiões de codificação, os SNPs podem ocasionar alterações na sequência das proteínas codificadas e, por conseguinte, na sua função desempenhada, levando a mudanças nos fenótipos (Kim; Misra, 2007).

A genotipagem pode ser realizada utilizando técnicas e metodologias diversas, destacando-se aquelas baseadas em: 1) Hibridização, como a hibridização alelo específica e a hibridização de sondas de hidrólise; 2) Métodos de ligação de oligonucleotídeo baseado em PCR, como a PCR-RFLP, e 3) Métodos baseados em NGS, os quais permitem a análise paralela de milhares a milhões de *loci* simultaneamente (Turchetto-Zolet *et al.*, 2017).

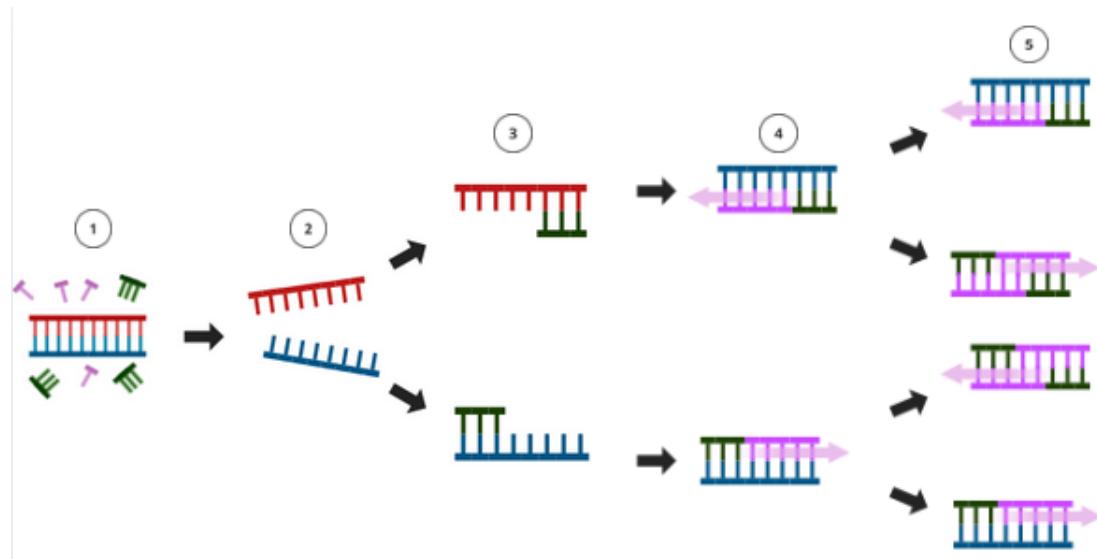
1.5.1 Reação em cadeia da polimerase (PCR)

A técnica consiste na amplificação de regiões específicas de interesse do DNA por meio de uma reação enzimática, sendo capaz de produzir grande número de cópias de material genético a partir de concentrações pequenas (McDonald; Taylor; Linacre, 2024).

Ela necessita de reagentes como trifosfatos de desoxirribonucleotídeos (dNTPs), enzima DNA polimerase termoestável, *primers* e tampão contendo magnésio e potássio (Lorenz, 2012).

A PCR é constituída por 3 etapas consecutivas dentro de um ciclo (Figura 5): 1) a etapa de desnaturação, na qual o DNA é submetido a altas temperaturas para a separação das duplas fitas do DNA, 2) a etapa de anelamento, na qual os *primers* se anelam às suas regiões complementares das fitas separadas do DNA, e 3) a etapa de extensão, na qual a reação é submetida a temperatura ótima para atividade da polimerase, de modo a permiti-la sintetizar a nova molécula de DNA (Johnson, 1991). Essas 3 etapas se repetem diversas vezes durante o processo, duplicando a quantidade de material genético a cada ciclo.

Figura 5 - Mecanismo de funcionamento da PCR.



Fonte: Elaborado pelo autor. 1: Início da reação com *primers*, dNTPs, tampão contendo magnésio e potássio e enzima DNA polimerase. 2: Etapa de desnaturação, necessária para separar as duplas fitas e facilitar o anelamento dos *primers*. 3: Etapa de anelamento, em que os *primers* se unem à região complementar específica do fragmento de DNA *template*. 4: Etapa de extensão, em que a enzima DNA polimerase sintetiza as novas fitas de DNA. 5: Representação dos ciclos posteriores ao primeiro, em que cada nova fita passa pelas mesmas etapas, amplificando o fragmento inicial de forma exponencial.

Apesar de ser uma técnica valiosa para as pesquisas moleculares, a PCR apresenta algumas limitações: 1) devido a sua alta sensibilidade, a técnica é bastante suscetível a contaminações, de modo que mesmo pequenas concentrações de DNA contaminante pode gerar resultados enganosos ou ambíguos, sendo necessário uso de controles para confirmação da ausência de contaminação; 2) as enzimas DNA polimerase são propensas a erros durante a incorporação de nucleotídeos à sequência sendo sintetizada, embora a taxa muito baixa;; 3) a técnica requer conhecimentos prévios sobre a sequência desejada, pois os *primers* são desenhados com base em informações pré-existentes sobre a sequência de interesse no estudo e 4) é suscetível a inibidores da reação, como reagentes residuais da extração de DNA, e a excesso de DNA, uma vez que grandes quantidades de DNA *template* sequestram o Mg²⁺, cofator da DNA polimerase, possivelmente comprometendo os resultados obtidos ao final do processo (Garibyan; Avashia, 2013).

Em certos protocolos de genotipagem para identificação de variantes genéticas, como a PCR-RFLP, a PCR configura-se como um dos primeiros passos de uma cadeia de processos, tendo a função de amplificar a sequência de interesse de modo a garantir quantidade suficiente de material genético (Crossley *et al.*, 2020)..

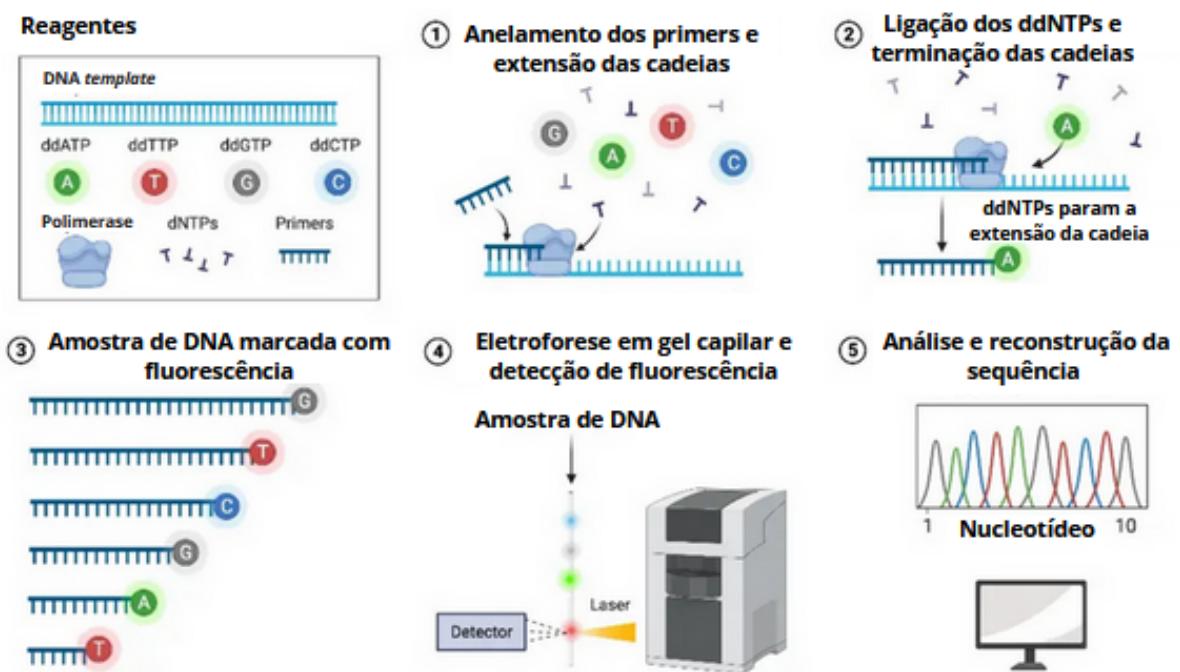
1.5.2 Sequenciamento de Sanger

O sequenciamento de Sanger é uma técnica de sequenciamento de DNA amplamente utilizada devido a sua precisão, sendo usada em sequenciamento de fragmentos de DNA, em diagnósticos clínicos e em validação de variantes genéticas (Valencia *et al.*, 2013), sendo considerada padrão ouro para determinação de sequências de DNA, sejam elas de ocorrência natural ou produzidas sinteticamente (Crossley *et al.*, 2020).

A técnica baseia-se num processo de replicação do DNA com a adição de nucleotídeos especiais chamados de didesoxinucleotídeos (ddNTPS), os quais são nucleotídeos modificados que possuem um hidrogênio (-H) no carbono 3' em vez da hidroxila (-OH) presente em nucleotídeos normais (Figura 6) (*Sanger Sequencing and Next-Generation Sequencing: Detailed Procedures and Applications*, 2024). Quando esses nucleotídeos são adicionados à cadeia de DNA sendo sintetizada, a síntese é interrompida, impedida pela ausência da hidroxila a qual é ligada o próximo nucleotídeo. Dessa forma, são gerados fragmentos de diversos tamanhos, permitindo assim a identificação da posição exata de cada nucleotídeo da sequência (Valencia *et al.*, 2013). Após a síntese dos fragmentos, é necessário separá-los baseados no seu tamanho para que seja feita a leitura da sequência. Para isso, um

método muito utilizado é a eletroforese capilar, a qual permite que os fragmentos migrem do capilar em ordem crescente de tamanho, passando, posteriormente, por um detector de fluorescência. O equipamento sequenciador de DNA, então, registra os dados coletados pelo detector e gera um eletroferograma, um gráfico que mostra picos fluorescentes correspondentes a cada nucleotídeo (Crossley *et al.*, 2020).

Figura 6 - Procedimento detalhado do sequenciamento de Sanger.



Fonte: Adaptado de PREACHBIO (2024). 1) Anelamento dos *primers* na sequência do DNA *template* e extensão da sequência pela polimerase; 2) Ligação dos ddNTPs à fita molde, interrompendo a síntese pela polimerase e gerando fragmentos de diversos tamanhos; 3) Alguns ddNTPs possuem marcadores que geram fluorescência; 4) os fragmentos são separados baseados no seus tamanhos por meio de eletroforese capilar; 5) O equipamento de análise escaneia a eletroforese capilar e detecta os sinais de fluorescência, gerando um eletroferograma da sequência analisada.

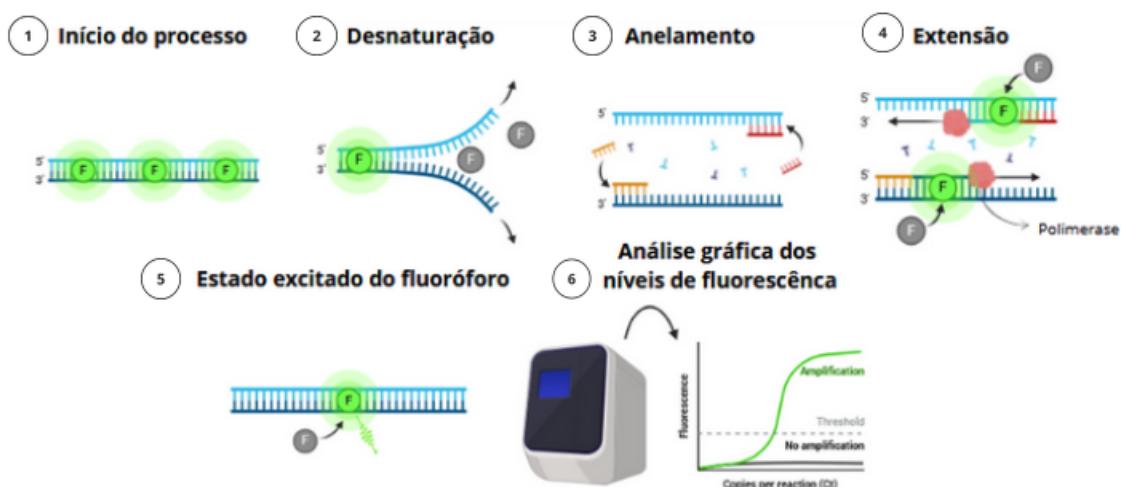
O sequenciamento de Sanger, apesar de ser o padrão ouro para determinação de sequências de DNA devido a sua alta confiabilidade e precisão, apresenta algumas desvantagens com relação a outros métodos de identificação de polimorfismos: 1) o alto custo para grandes volumes de amostra, quando comparado a outros métodos de sequenciamento, como NGS; 2) a limitação quanto ao tamanho dos fragmentos, podendo sequenciar com qualidade até um máximo de 900 pb, e 3) seu baixo paralelismo, não sendo capaz de sequenciar muitos fragmentos simultaneamente, como o NGS (Shaffer, 2023).

1.5.3 PCR quantitativa (qPCR) em tempo real

Com o advento de um método de monitoramento do processo de PCR no decorrer dos ciclos, desenvolvido inicialmente em 1993 por Higuchi et al (1993), tornou-se possível acompanhar a amplificação dos fragmentos de DNA em tempo real por meio de fluorescência, descartando a necessidade de etapas de detecção após a técnica, como eletroforese ou hibridização por *Southern blot*, diminuindo as chances de contaminação por manipulação das amostras durante essas etapas (Nascimento; Suarez; Pinhal, 2015).

A detecção em tempo real da qPCR pode ser feita por dois métodos de fluorescência distintos, cada um com suas peculiaridades, vantagens e desvantagens. O primeiro deles baseia-se no uso de moléculas intercalantes fluorescentes (Figura 7), como o SYBR Green, os quais são moléculas capazes de se ligar em regiões da dupla fita de DNA e emitem fluorescência quando excitados em comprimentos de onda específicos (Kubista et al., 2006). Quanto mais amplicons formados ao longo dos ciclos de amplificação, maior o sinal de fluorescência emitido pela reação, até que ultrapasse um nível de fluorescência, o *threshold*, acima do ruído de fundo. Suas vantagens consistem no menor custo do método de detecção e na sua simplicidade. Apesar de sua praticidade, o método não é específico, uma vez que o corante se liga a qualquer DNA de fita dupla, podendo se ligar à dímeros de *primers* e produtos não-específicos (Sergeevich et al., 2024).

Figura 7 - Mecanismos de fluorescência de corantes intercalantes.

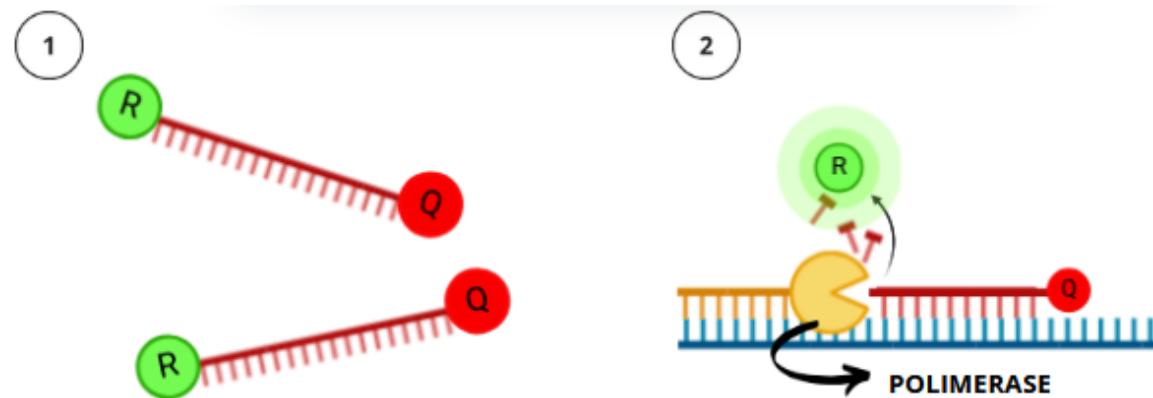


Fonte: Elaborado pelo autor. Na etapa de desnaturação (2), aumenta-se a temperatura até que ocorra a separação das fitas do DNA, liberando o fluoróforo no meio em seu estado inibido. Após o anelamento dos *primers* (3), a enzima polimerase se liga à sequência e inicia a etapa de polymerização do fragmento de interesse (4), e, durante esse processo, os fluoróforos se ligam de forma intercalada à dupla fita

formada, passando a exibir estado de excitação (5). Conforme ocorrem mais ciclos, a quantidade de DNA de dupla fita com fluoróforos intercalados aumenta, ocasionando um aumento do sinal de fluorescência, o qual será captado pelo equipamento (6).

O segundo método de detecção baseia-se no uso de sondas de hidrólise, como as sondas TaqMan (Figura 8). Estas consistem em pequenas sequências de DNA sintéticas desenhadas para anelar a uma região específica do DNA alvo (Kubista *et al.*, 2006). Elas são formadas por duas partes importantes ligadas em extremidades opostas do oligonucleotídeo: o fluoróforo (repórter), o qual é responsável pela emissão da fluorescência que é captada pelo software acoplado ao equipamento, e o *quencher* (em português, extintor), o qual tem o papel de inibir a fluorescência do repórter enquanto estiverem ligados pelo oligonucleotídeo. Esse método tem a vantagem de ser mais específico, pois as sondas de hidrólise são desenhadas para anelar especificamente a uma região do DNA alvo, minimizando a detecção de produtos inespecíficos e garantindo maior confiabilidade nos resultados, sendo, contudo o custo de desenvolvimento das sondas relativamente elevado (Sergeevich *et al.*, 2024).

Figura 8 - Mecanismo de fluorescência de sondas de hidrólise.

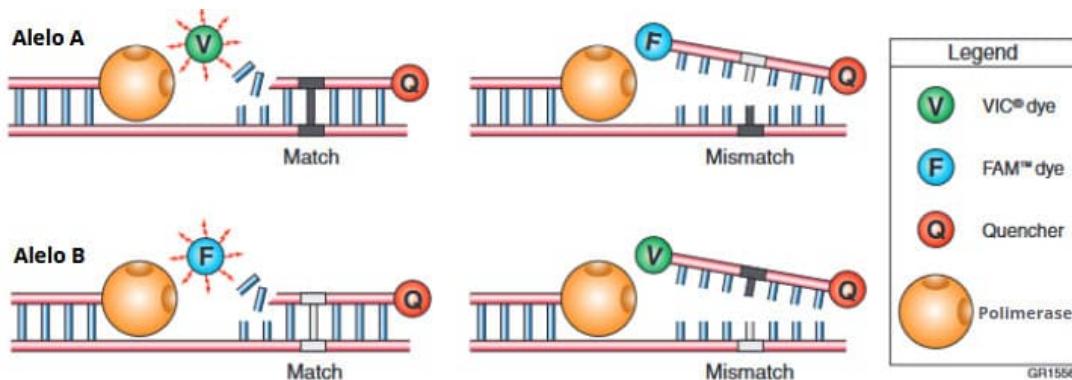


Fonte: Elaborado pelo autor. 1) Esquema de sonda de hidrólise TaqMan, composta de fluoróforo, ou repórter, na extremidade 5' e *quencher* na extremidade 3'. 2) A degradação da sonda pela polimerase durante a etapa de extensão libera o fluoróforo, emitindo sinal de fluorescência que será captado pelo detector.

O método de detecção da qPCR por meio de sondas de hidrólise também se configura como uma opção para genotipagem de sequências de interesse por meio do uso da plataforma TaqMan® (Applied Biosystems), a qual utiliza a amplificação da reação em cadeia da polimerase e a discriminação alélica para gerar dados de genotipagem (Schleinitz *et al.*, 2011).

O seu princípio de funcionamento baseia-se no uso de um par de sondas com dois corantes fluorescentes distintos, como, por exemplo, duas sondas TaqMan® que têm como alvo um local de SNP (Applied Biosystems, 2010). Uma das sondas com corantes fluorescentes específicos, por exemplo VIC®, anela-se perfeitamente com o SNP do alelo A enquanto a outra sonda com corante fluorescente distinto do primeiro, por exemplo FAM™, anela-se perfeitamente com o SNP do alelo B (Figura 9). Caso a sonda não seja perfeitamente complementar à região da SNP, o anelamento da sonda será imperfeito e de baixa frequência, gerando assim menores níveis de fluorescência quando comparado à sonda que se anelou com perfeição.

Figura 9 - Representação do funcionamento do anelamento das sondas com diferentes fluoróforos aos alelos A e B.



Fonte: Adaptado de Applied Biosystems (2010). No alelo A, a sonda marcada com o fluoróforo VIC® anela-se perfeitamente na região do SNP, formando um *match*, enquanto a sonda marcada com o fluoróforo FAM™ não se anela ou anela de forma imperfeita, ocorrendo *mismatch*. No alelo B ocorre o contrário, ocorrendo *match* para a sonda com fluoróforo FAM™ e *mismatch* para a sonda com fluoróforo VIC®.

A genotipagem pelo método de qPCR pelo uso de sondas de hidrólise é feita pela medida dos níveis de fluorescência de cada fluoróforo realizada pelo *software* do termociclador (Quadro 2). É a discrepância entre os níveis de fluorescência de cada corante que será analisada pelo *software* do equipamento de qPCR para que seja feita a genotipagem. O *software* registra os níveis de emissão de fluorescência de cada fluoróforo e gera um gráfico de discriminação alélica, também conhecido como gráfico de dispersão, em que cada eixo representa um corante fluorescente. Cada poço com amostra é registrado como um ponto individual no gráfico, agrupando-os em *clusters*, ou conjuntos, dependendo do genótipo atribuído (heterozigoto ou homozigoto para o alelo A ou para o alelo B) pelo *software* (APPLIED BIOSYSTEMS, 2010).

Quadro 2 - Exemplo da correlação entre os sinais de fluorescência e as sequências em uma amostra qualquer.

Um aumento substancial de...	Indica...
Apenas corante VIC®	Homozigose para o alelo A
Apenas corante FAM™	Homozigose para o alelo B
Ambos os sinais de fluorescência	Heterozigose para o alelo A e alelo B

Fonte: Adaptado de Applied Biosystems (2010). Correlação utilizada pelo programa de genotipagem para discriminação alélica das amostras.

1.5.4 Abordagens atuais para genotipagem em rebanhos leiteiros.

Com o crescente conhecimento das características genéticas de animais de produção, como os bovinos leiteiros, têm surgido um interesse no desenvolvimento de métodos e técnicas que permitam manejar os rebanhos buscando agregar valor à cadeia produtiva do leite, por meio de mapeamento e melhoramento dos caracteres genéticos dos rebanhos (Silva; Demuner, 2024).

Atualmente são utilizadas diversas técnicas para analisar marcadores de produtividade, como o gene da κ -caseína, em rebanhos de bovinos leiteiros. Para esses marcadores, há testes imunológicos, como os baseados em ELISA (*Enzyme-Linked Immunosorbent Assay*), que permitem a inferência de genes a partir da análise da detecção de proteínas relacionadas a esses marcadores, como o *Bovine Casein Kappa ELISA Kit* (Antibodies.com). No caso da κ -caseína, por exemplo, esse teste permite inferir indiretamente o genótipo por meio de testes da concentração de κ -caseína na amostra de leite analisado, uma vez que, segundo Fox et al. (2017), as variantes B do gene de κ -caseína estão associados a maiores teores de κ -caseína na composição do leite. Contudo, essas técnicas não permitem distinguir entre animais homozigotos e heterozigotos, uma vez que não analisam os genes em si. Outra limitação dos testes baseados em imunoensaios é que somente fêmeas que estão em processo de produção leiteira podem ser analisadas, uma vez que o método detecta biomarcadores específicos no leite e os machos não produzem leite.

A seleção genética de machos na bovinocultura leiteira, nesse contexto, é especialmente impactante, principalmente devido ao seu potencial de multiplicação genética. Enquanto uma fêmea pode produzir, em média, um bezerro por ano, um único touro, por meio

da inseminação artificial, pode gerar milhares de descendentes anualmente, influenciando significativamente o progresso genético do rebanho. O investimento na seleção de machos é estratégico e importante, uma vez que eles possuem impacto significativo em um grande número de fêmeas ao mesmo tempo, tornando-o uma ferramenta essencial para a eficiência e sustentabilidade da bovinocultura leiteira. A seleção de touros com alto valor genético pode resultar em ganhos superiores em características como produção de leite, composição do leite e saúde do rebanho, em comparação com a seleção de fêmeas (Pryce *et al.*, 2017; Brito *et al.*, 2020), além de aumentar em até 50% a taxa de progresso genético quando comparada aos métodos tradicionais (García-Ruiz *et al.*, 2016).

Nesse contexto, os testes moleculares, como a qPCR, configuram-se como uma opção mais viável para a análise e estudo de caracteres polimórficos relacionados a produtividade, uma vez que, por terem os genes como objeto direto da análise não apresentam influência do ambiente e podem ser detectados independente de sexo e estágio de desenvolvimento dos animais (Faleiro, 2007).

Apesar disso, o mercado de testes moleculares para o gene de κ -caseína no Ceará ainda não é bastante explorado, criando assim uma oportunidade para o desenvolvimento de novas ideias e tecnologias que permitam aos produtores explorarem melhor o potencial genético do seu rebanho e gerar produtos de qualidade.

Dessa forma, neste trabalho propôs-se o desenvolvimento de um teste de genotipagem por qPCR para os polimorfismos do gene da κ -caseína. Dessa forma, busca-se melhorar o manejo genético dos animais no rebanho cearense, oferecendo uma nova ferramenta robusta e precisa aos Zootecnistas e Médicos Veterinários locais.

2 OBJETIVOS

2.1 Objetivo geral

Desenvolver um teste de genotipagem dos polimorfismos do gene da *kappa*-caseína através de qPCR em rebanhos leiteiros cearenses.

2.2 Objetivos específicos

- Desenhar *primers* e sondas de hidrólise específicas para o gene CSN3;
- Padronizar a reação da genotipagem por qPCR;
- Clonar o produto de PCR em vetor de clonagem apropriado;
- Sequenciar os clonados para obter controles positivos para as variantes A e B do gene da κ -caseína.

3 MATERIAIS E METODOLOGIA

3.1 Materiais

3.1.1 Reagentes

- *Primers* e sondas de hidrólise TaqMan para amplificação de fragmento do gene de κ-caseína de propriedade da empresa TaqMol Diagnósticos;
- *Primers* universais M13;
- *DreamTaq™ Green PCR Master Mix* (Thermo Fisher);
- *GoTaq® qPCR Master Mix 2X* (Promega);
- *qPCR ProbesMaster 2X* (Cellco);
- *BenchTop 100 pb DNA ladder* (Promega);
- *pGEM®-T Easy Vector Systems* (Promega);
- IPTG 100 mM;
- X-gal 120 mM;
- Carbenicilina 100 mM

3.1.2 Kits de extração

- *Blood-Animal-Plant DNA Preparation Kit* (Cellco);
- *QIAprep Spin Miniprep Kit* (QIAGEN);
- *QIAquick® Gel Extraction Kit* (QIAGEN).

3.1.3 Organismos utilizados

Células quimiocompetentes de *Escherichia coli* DH5- α preparadas seguindo as instruções de Green e Sambrook (2012).

3.2 Obtenção das amostras de sangue total bovino

As amostras de sangue total bovino, de animais sabidamente saudáveis, utilizadas nos experimentos foram cedidas pela empresa TaqMol Diagnósticos.

3.3 Extração de gDNA

As extrações foram realizadas nas dependências do Laboratório de Biotecnologia Molecular (LabBMoL) que possui CQB mantido pela CTNBio, utilizado o *Blood Animal-Plant DNA Preparation Kit* (Cellco), seguindo as instruções do fabricante. Primeiro, as amostras de sangue foram submetidas à tampão de lise para rompimento da membrana das hemárias. Em seguida, adicionou-se tampão de lise para rompimento das membranas dos leucócitos e liberação do DNA, simultaneamente ,adicionou-se proteinase K para degradação de proteínas estruturais e histonas que estão associados ao DNA genômico. Logo após, procedeu-se a adição de RNase A para degradação do RNA presente na amostra. Por fim, o DNA foi lavado em coluna de purificação fornecida pelo fabricante e eluído com tampão de eluição também fornecido pelo fabricante.

3.4 Desenho dos *primers* e sondas de hidrólise

O desenho dos *primers* e das sondas de hidrólise utilizados nos experimentos foram feitos por meio de ferramentas especializadas, como *Primer3Plus* e *PrimerQuest™*, os quais forneceram os oligonucleotídeos específicos para o gene *CSN3* de ID: 281728, obtida no banco de dados GenBank, codificante do gene da κ-caseína bovina.

Parâmetros, como temperatura de melting, composição GC, quantidade de pares de bases, formação de dímeros (homo e heterodímeros), bem como a formação de estruturas secundárias chamadas de *hairpins* foram analisados no desenho dos *primers* e sondas para que se tenha um melhor aproveitamento nas etapas de padronização subsequentes.

Os *primers* e sondas de hidrólise utilizados neste estudo são de propriedade da empresa TaqMol Diagnósticos e suas sequências são de confidencialidade da empresa.

3.5 PCR para teste dos *primers* e sondas de hidrólise

Foram realizadas no equipamento *QuantStudio™ 3 Real-Time PCR System* (Applied Biosystems). Para o teste do par de *primers* desenhados, fez-se as reações com duplicatas conforme o Quadro 3:

Quadro 3 - Relação de reagentes, suas concentrações e volumes adicionados à reação.

Reagente	Concentração final	Volume (μ L)
<i>Primer forward</i>	200 nM	0,4
<i>Primer reverse</i>	200 nM	0,4
<i>GoTaq® qPCR Master Mix 2X</i> (Promega)	1X	10
gDNA bovino	\approx 20 ng/ μ L	1
Água ultrapura	-	q.s.p. 20

Fonte: Elaborado pelo autor. No quadro, a lista de reagentes utilizados, as suas concentrações finais na reação e o volume utilizado em μ L.

As condições de ciclagem do equipamento para 40 ciclos foram as seguintes: desnaturação inicial de 95 °C por 2 minutos, seguida de 95°C por 15 segundos, anelamento e extensão a 60° C por 1 minuto. Após o término dos 40 ciclos, houve uma etapa de dissociação das fitas duplas (*melt curve*): 95°C por 15 segundos, 60°C por 1 minuto e 95 °C por 1 segundo. Essa etapa permite a análise da especificidade dos produtos amplificados, possibilitando a detecção de amplificação de dímeros de *primers* ou outros produtos indesejados. Como controle negativo, usou-se duplicata de uma reação com adição de 1 μ L de água ultrapura em substituição ao gDNA.

Para o teste das sondas de hidrólise, fez-se a seguinte reação com duplicatas conforme o Quadro 4:

Quadro 4 - Reagentes, suas concentrações finais e volumes adicionados à reação.

Reagente	Concentração final	Volume (μ L)
<i>Primer forward</i>	150 nM	0,3
<i>Primer reverse</i>	150 nM	0,3
Sonda de hidrólise K1	400 nM	0,08
Sonda de hidrólise K2	400 nM	0,8
Corante ROX	50 nM	0,04

<i>qPCR ProbesMaster 2X</i> (Cellco)	1X	10
gDNA bovino	$\approx 20 \text{ ng}/\mu\text{L}$	1
Água ultrapura	-	q.s.p. 20

Fonte: Elaborado pelo autor. No quadro, a lista de reagentes utilizados, as suas concentrações finais na reação e o volume utilizado em μL .

O corante ROX é um corante fluorescente inerte com um nível de fluorescência que não se altera durante todo o processo de PCR, não sendo afetado pela amplificação do produto de PCR. Contudo, ele é afetado por qualquer outro fator que altere as leituras de fluorescência do equipamento, como a presença de bolhas na reação, erros de pipetagem, entre outros. Ele é usado como uma referência passiva que permite a normalização fluorescente para dados de qPCR, fornecendo uma maneira de contabilizar variações potenciais que distorceram as leituras de amplificação (Thermo Fisher, 2021).

As condições de ciclagem do equipamento para 40 ciclos foram as seguintes: desnaturação inicial a 95 °C por 2 minutos, seguida de 95°C por 15 segundos, anelamento e extensão a 60 °C por 1 minuto. Após o término dos 40 ciclos, houve uma etapa de pós-leitura a 60 °C por 30 segundos, para análise dos dados de fluorescência coletados pelo equipamento. Como controle negativo, usou-se duplicata de uma reação com adição de 1 μL de água ultrapura em substituição ao gDNA.

A análise dos dados gerados foi feita por *software Design & Analysis* versão 2.8.0 (*Applied Biosystems*).

3.6 PCR para amplificação do fragmento de κ -caseína

A PCR foi realizada no equipamento *Mastercycler* Eppendorf®. A reação foi feita em duplicatas, com volume final de 50 μL , conforme o Quadro 5:

Quadro 5 - Reagentes, concentrações finais e volumes adicionados à reação.

Reagente	Concentração final	Volume (μL)
<i>Primer forward</i>	200 nM	1
<i>Primer reverse</i>	200 nM	1

<i>DreamTaq™ Green PCR Master Mix 2X</i> (Thermo Fisher)	1X	10
gDNA bovino	$\approx 20 \text{ ng}/\mu\text{L}$	2
Água ultrapura	-	q.s.p. 50

Fonte: Elaborado pelo autor. No quadro, a lista de reagentes utilizados, as suas concentrações finais na reação e o volume utilizado em μL .

A reação foi feita com uma desnaturação inicial a 95°C por 3 minutos, seguida de 34 ciclos de PCR sob as seguintes condições: desnaturação a 95°C por 30 segundos, anelamento a 60 °C por 30 segundos, extensão a 72 °C por 30 segundos e extensão final a 72 °C por 10 minutos. Em seguida, os 50 μL do produto de PCR foram analisados em corrida eletroforética em gel de agarose 2% em tampão TAE, utilizando-se marcador de peso molecular *BenchTop 100 pb DNA Ladder* (Promega), a fim de confirmar a amplificação do DNA observando-se o padrão de banda formada. Após análise do gel de eletroforese, excisou-se as bandas com bisturi para posterior purificação do produto de PCR.

3.7 Purificação do produto de PCR

Para purificação do DNA do produto de PCR, utilizou-se o *QIAquick® Gel Extraction Kit* (QIAGEN) seguindo as instruções do fabricante. O DNA extraído foi eluído em água ultrapura aquecida a 60 °C. Após a extração, foi feita uma corrida eletroforética em gel de agarose 2% em tampão TAE utilizando marcador de peso molecular *BenchTop 100 pb DNA Ladder* (Promega) para análise do padrão de bandas e confirmação da purificação do produto de PCR.

3.8 Clonagem do produto de PCR purificado

3.8.1 Ligação do produto de PCR purificado ao vetor *pGEM*

Utilizou-se o *pGEM®-T Easy Vector Systems* (Promega) para ligar o DNA a um vetor de clonagem. Adaptou-se as instruções do fabricante da seguinte forma: para a reação de 10 μL de ligação usou-se 1 μL de *T4 DNA Ligase*, 0,5 μL de *pGEM®-T Vector*, 5 μL de 2X

Rapid Ligation Buffer, T4 DNA Ligase, 0,5 µL de água ultrapura e 3 µL de DNA do produto de PCR purificado. Após a mistura dos componentes, incubou-se a reação à temperatura ambiente por 1 hora. Em seguida, armazenou-se a reação em geladeira a 4 °C *overnight*.

A enzima polimerase presente no mastermix utilizado para amplificação do fragmento adiciona uma adenina (A) nas extremidades 3' do DNA amplificado, o que permite que o vetor, o qual é fabricado com extremidades timina (T) coesivas, ligue-se de forma eficiente ao produto de PCR, impedindo a recircularização do vetor (*pGEM®-T and pGEM®-T Easy Vector Systems Technical Handbook*, 2025).

O vetor além de possuir gene de resistência a ampicilina, permitindo a seleção de colônias transformadas, também contém uma região lacZ, que codifica uma subunidade da enzima β-galactosidase. Caso o vetor seja religado sem o inserto, a β-galactosidase será expressa pelas células, formando colônias de cor azul quando cultivadas em meios contendo IPTG e X-Gal. Se o inserto for clonado com sucesso, ele interrompe o gene lacZ, resultando em colônias brancas, indicando a presença do DNA de interesse (*pGEM®-T and pGEM®-T Easy Vector Systems Technical Handbook*, 2025).

3.8.2 Transformação genética

Para a transformação genética, utilizaram-se células quimiocompetentes de *Escherichia coli* DH5-a para transformação. Primeiramente, fez-se uma reação de volume 100 µL: 10 µL de MgCl₂ 0,1 M, 5 µL CaCl₂ 0,1 M, 30 µL de água ultrapura, 50 µL de células quimiocompetentes *E. coli* DH5-a e 5 µL de DNA ligado ao pGEM. Depois incubou-se a reação em gelo por 30 minutos, logo em seguida aqueceu-se a reação a 42°C em banho seco por 2 minutos para o choque térmico, por fim incubou-se novamente em gelo por 1 minuto. Após as incubações, adicionou-se 900 µL de meio Luria-Bertani (LB), preparado segundo Miller (1972), ao tubo de reação e submeteu-se o tubo a agitação por 1 hora, a 37 °C e 180 rotações por minuto (RPM) para regeneração das células. Posterior à regeneração, centrifugou-se o tubo a 14500 RPM por 1 minuto, descartando o sobrenadante e ressuspensendo as células em 100 µL de meio LB. Em seguida, plaqueou-se os 100 µL do tubo em meio LB com ágar, adicionado de carbenicilina 100 mM, IPTG 100 mM e X-gal 120 mM por meio do uso de esferas de vidro, *beads*, até espalhar todo o líquido pela placa. Por fim incubou-se as placas em estufa a 37 °C *overnight*.

3.8.3 PCR das colônias transformadas

Fez-se PCR de 7 colônias transformadas de *E. coli* DH5- α para averiguar o resultado da transformação. A reação de 25 μ L foi feita da seguinte forma: 0,5 μ L de cada primer M13, 11,5 μ L de água ultrapura e 12,5 μ L de *DreamTaq™ Green PCR Master Mix* (Thermo Fisher). Com a ponta de ponteiras de 10 μ L, coletou-se as colônias selecionadas e usou-as para amplificação de DNA. Os parâmetros de termociclagem para 35 ciclos foram os seguintes: desnaturação inicial a 95 °C por 5 minutos, desnaturação a 95 °C por 30 segundos, anelamento a 50 °C por 30 segundos, extensão a 72 °C por 30 segundos e extensão final a 72 °C por 5 minutos. Em seguida, 10 μ L dos produtos de PCR das colônias foi analisado em corrida eletroforética em gel de agarose 2%, utilizando-se marcador de peso molecular *BenchTop 100 pb DNA Ladder* (Promega), a fim de confirmar a correta amplificação do DNA por meio da observação das bandas formadas.

3.8.4 Pré-inóculo das colônias transformadas

Selecionou-se 3 colônias brancas para preparo de pré-inóculos. Cultivou-se as 3 colônias selecionadas da placa de petri em tubos de centrifugação 15 mL contendo 5 mL de meio LB líquido adicionado de carbenicilina, e incubadas a 37°C sob agitação a 180 RPM, *overnight*.

3.8.5 Miniprep das colônias transformadas

Foi feita a extração do DNA do pré-inóculo foi feita por meio da utilização do *QIAprep Spin Miniprep Kit* (QIAGEN), seguindo as instruções do fabricante. O DNA foi eluído em 50 μ L água ultrapura a 60°C.

3.8.6 Sequenciamento de Sanger

Para o sequenciamento das amostras de DNA das colônias transformadas extraídas, fez-se parceria com a Fundação Oswaldo Cruz - CE (Fiocruz) para a realização de sequenciamento de Sanger.

3.9 Testes de genotipagem com controle positivo da variante A

Após a obtenção do controle positivo, foram feitas novas corridas de qPCR para testar a genotipagem de amostras de gDNA bovino. Para essas corridas, fez-se a seguinte reação com duplicatas para 10 amostras de bovinos, com volume final de 20 µL: 0,3 µL de cada primer a 10 mM, 0,08 µL de cada sonda de hidrólise, 8,2 µL de água ultrapura, 0,04 µL de corante ROX, 10 µL de *qPCR ProbesMaster* (Cellco) e 1 µL de gDNA bovino. As condições de ciclagem do equipamento para 40 ciclos foram as seguintes: desnaturação inicial a 95 °C por 2 minutos, seguida de 95 °C por 15 segundos, anelamento e extensão a 60 °C por 1 minuto. Após o término dos 40 ciclos, houve uma etapa de pós-leitura a 60 °C por 30 segundos para análise dos dados de fluorescência pelo equipamento.

A análise dos dados gerados foi feita por *software Design & Analysis* versão 2.8.0 (Applied Biosystems).

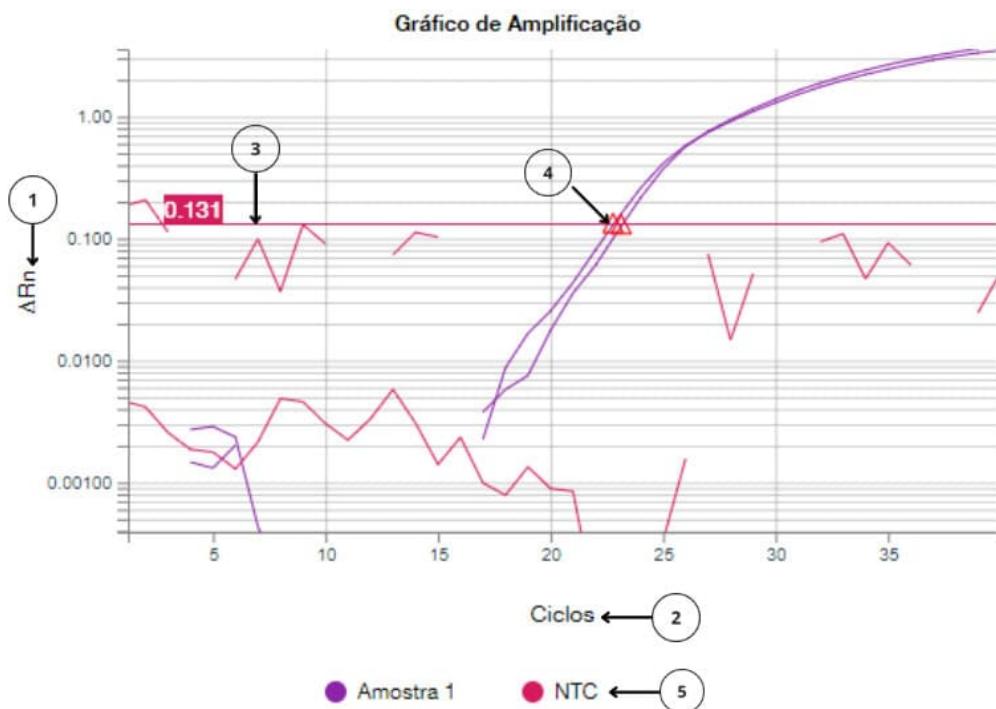
4 RESULTADOS E DISCUSSÃO

4.1 PCRs para teste dos *primers* e sondas de hidrólise

A Figura 10 mostra um gráfico de amplificação gerado pelo *software Design & Analysis v. 2.8.0* após a corrida. De acordo com a Figura 10, gráfico de amplificação é composto pelos seguintes elementos: (1) Eixo Y: representa a fluorescência registrada, que reflete a quantidade de DNA amplificado na reação; (2) Eixo X: representa os ciclos da reação de PCR; (3) *Threshold*: representa um nível de fluorescência acima do ruído de fundo; (4) Cq (*quantification cycle*): representa o ciclo no qual a fluorescência gerada pelo produto amplificado atinge ou excede o *threshold* e (5) NTC (*no template control*): reação que contém todos os componentes de qPCR sem a adição de DNA molde.

Os *primers* desenhados foram capazes de amplificar o fragmento de κ -caseína desejado. As duplicatas de NTC não mostram amplificação, aparentando não haver sinais de contaminação.

Figura 10 - Amplification plot da corrida de teste do par de *primers*.

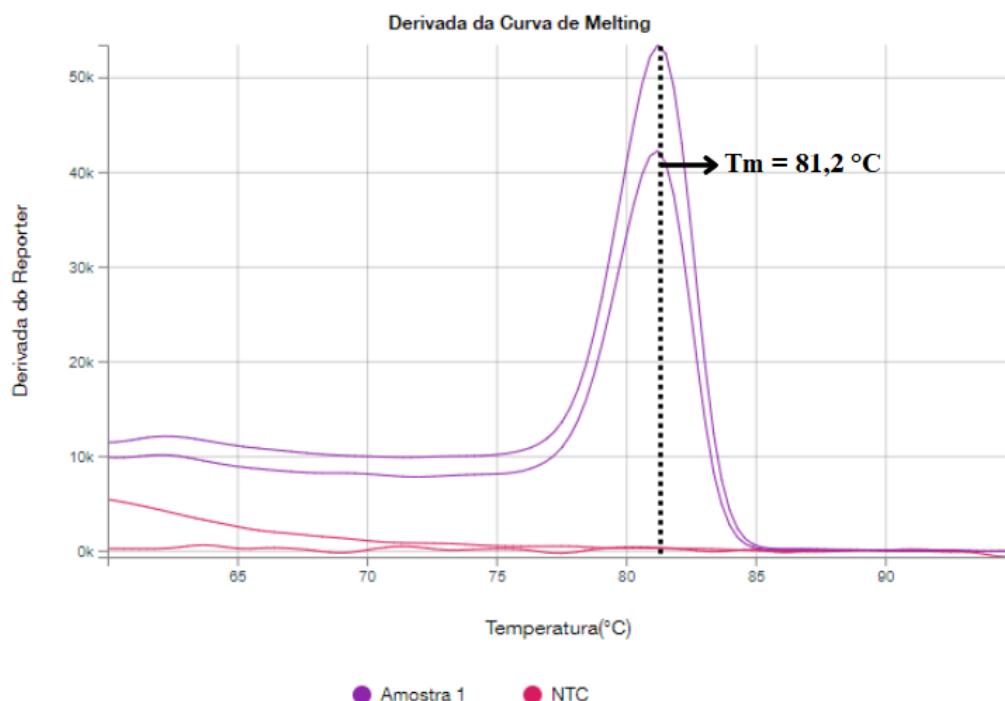


Fonte: *Design & Analysis v2.8.0* (Applied Biosystems). Gráfico de amplificação da amostra 1 para teste dos *primers*. Nota-se a amplificação do fragmento de κ -caseína. (1) Eixo Y, mede o nível de fluorescência a cada ciclo. (2) Eixo X, mostra o número de ciclos. (3) *Threshold*, etapa

exponencial da amplificação. (4) C_q, ciclo em que o gráfico intercepta o *threshold*. (5) NTC, reação com todos os componentes de qPCR exceto o DNA molde.

Após a mesma corrida, foi gerado também uma curva de *melting* (temperatura na qual 50% das moléculas de DNA encontram-se desnaturadas) (Figura 11). observa-se a presença de apenas um pico acentuado na amostra 1 de gDNA bovino testada, indicando a formação de amplicon único formado durante a corrida em temperatura de 81,2 °C condizente com a esperada para o amplicon de κ-caseína gerado pelos *primers* desenhados, confirmando a especificidade da amplificação. A ausência de picos no NTC (*no template control*) indica que não houve contaminação ou amplificação inespecífica nas reações.

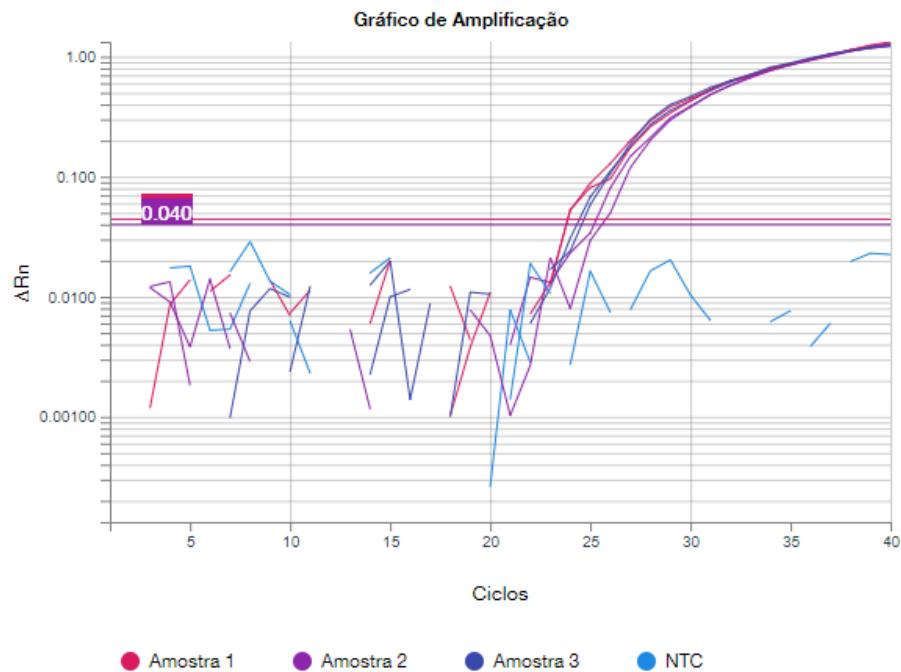
Figura 11 - Gráfico da curva de *melt* dos amplicons gerados na corrida de teste dos *primers*.



Fonte: *Design & Analysis v2.8.0* (Applied Biosystems). Gráfico da curva de *melting* do fragmento amplificado. A T_m de *melting* observada foi de 81,2°C. Nota-se a presença de um único pico definido. A amostra 1 é gDNA extraído de sangue total de bovino.

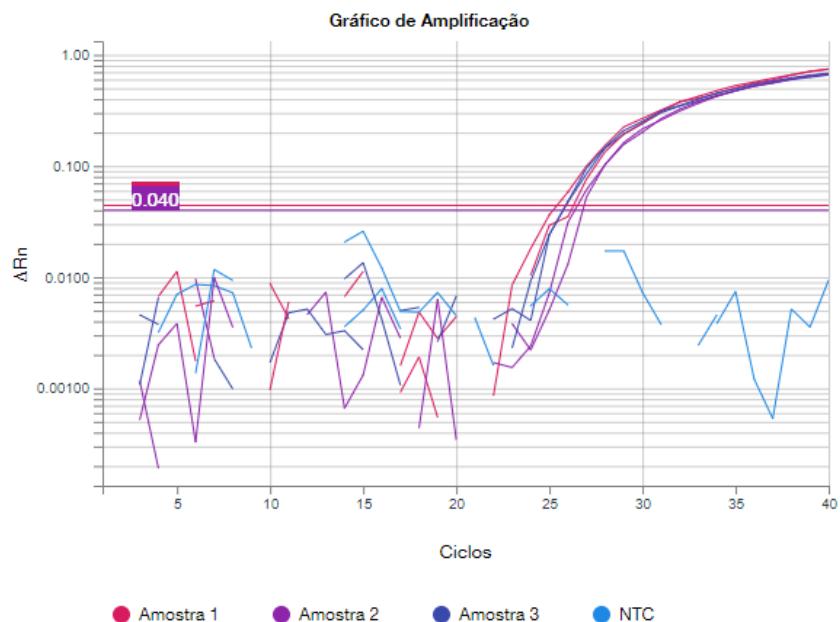
As Figuras 12 e 13 mostram as curvas de amplificação da corrida de teste das sondas de hidrólise KA2 e KB2 desenhadas, geradas pelo software *Design & Analysis v. 2.8.0* após a corrida. Todas as amostras, com exceção do NTC, apresentaram amplificação acima da linha de *threshold*, implicando num correto anelamento de ambas as sondas aos fragmentos específicos de cada uma.

Figura 12 - Gráfico de amplificação da sonda de hidrólise para o alelo A para as amostras 1, 2 e 3.



Fonte: *Design & Analysis v2.8.0* (Applied Biosystems). Nota-se a amplificação da sonda KA2 em todas as amostras testadas.

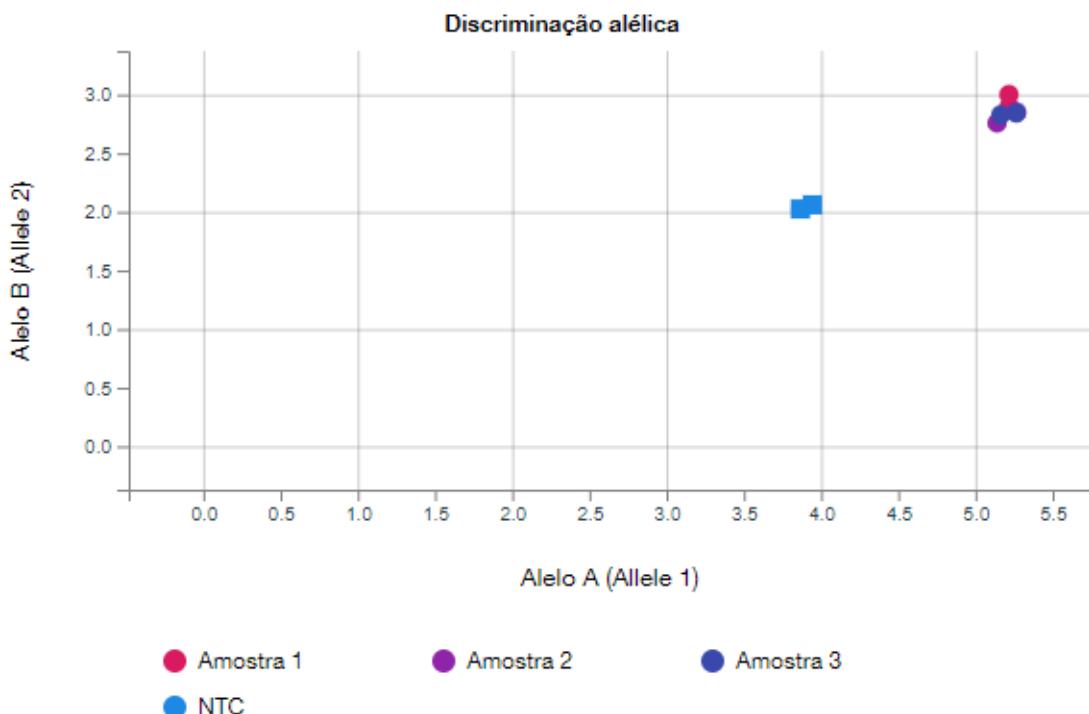
Figura 13 - Gráfico de amplificação da sonda de hidrólise para o alelo B para as amostras 1, 2 e 3.



Fonte: *Design & Analysis v. 2.8.0*. Nota-se a amplificação da sonda KB2 para todas as amostras testadas.

O software *Design & Analysis* também gerou ao final da mesma corrida uma curva de discriminação alélica das amostras (Figura 14). O programa não foi capaz de determinar um resultado (*call*) com a confiabilidade necessária, atribuindo *undetermined* (indeterminado, em português) como resultado (Quadro 6). Isso se deve ao fato de o nível de *confidence*, um parâmetro de qualidade do programa *Design & Analysis* (*Applied Biosystems*) que expressa a probabilidade de um *call* de genotipagem atribuída a cada poço do equipamento, ter dado um valor abaixo do padrão estabelecido pelo programa (*THERMO FISHER*, 2023).

Figura 14 - Allelic discrimination plot da corrida de teste das sondas de hidrólise.



Fonte: *Design & Analysis* v2.8.0 (Applied Biosystems). Na figura, observa-se o agrupamento das amostras feito pelo software com base no genótipo atribuído por ele a cada uma. A falta de controles positivos para as duas variantes faz com que o programa não tenha informações suficientes para gerar uma *call* confiável para as amostras. Como não há controles que ele possa usar de padrão, o software agrupa todas as amostras próximas umas das outras em um mesmo *cluster*.

Quadro 6 - Correlação das amostras com o nível de confidence e a call designada pelo software.

Sample	Confidence	Allele 1	Allele 2	Call
Amostra 1	0.947	5.218	2.901	Undetermined
Amostra 2	0.947	5.14	2.758	Undetermined
Amostra 3	0.947	5.164	2.829	Undetermined
Amostra 1	0.947	5.217	3.002	Undetermined
Amostra 2	0.947	5.265	2.852	Undetermined
Amostra 3	0.947	5.265	2.843	Undetermined
NTC	1	3.944	2.06	No Amplification
NTC	1	3.866	2.025	No Amplification

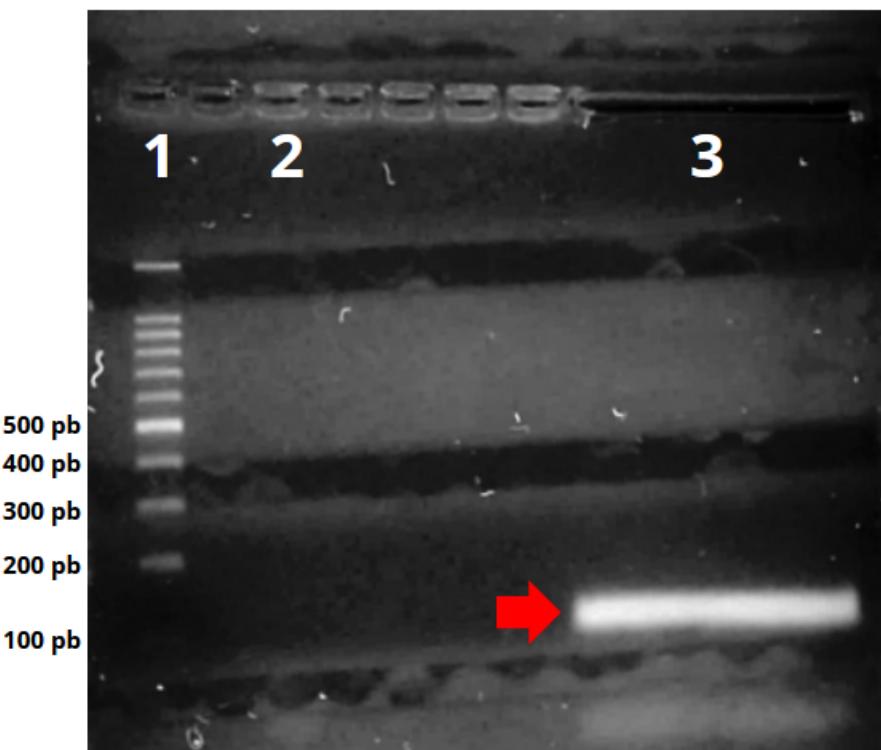
Fonte: *Design & Analysis v2.8.0* (Applied Biosystems). Imagem dos parâmetros de qualidade da qPCR de genotipagem, o valor de *confidence* < 95% calculado pelo programa impossibilita ser atribuído um genótipo às amostras testadas.

Com o intuito de aumentar o valor desse parâmetro durante as corridas e, dessa forma, melhorar a discriminação alélica feita pelo programa, propôs-se clonar o produto de PCR de uma das amostras e sequenciá-lo para obtenção de controle positivo para uma das variantes.

4.2 PCR para amplificação do fragmento de κ-caseína

A Figura 15 mostra o gel de eletroforese após a corrida com amostra 2 de gDNA bovino. Nele, vê-se que o fragmento amplificado contém entre 100-200 pb, tamanho previsto para amplificação pelos *primers* desenhos, o qual é de aproximadamente 125 pb. Após confirmação da amplificação do fragmento de κ-caseína, excisou-se as bandas de DNA do gel para que pudessem ser purificadas por meio do kit *QIAquick® Gel Extraction Kit* (QIAGEN), seguindo suas instruções.

Figura 15 - Análise da corrida eletroforética em gel de agarose 2% para validação da amplificação correta do gene de *kappa*-caseína.



Fonte: Elaborado pelo autor. 1): Marcador de peso molecular BenchTop *100 pb DNA ladder* (Promega); 2): NTC; 3): Amplicon de κ -caseína da amostra 2.

4.3 Purificação do produto de PCR

A Figura 16 mostra o gel de eletroforese do produto de PCR purificado. O tamanho das bandas aparenta conter entre 100-200 pb, estando no intervalo de tamanho previsto de aproximadamente 125 pb. Apesar de ser possível observar duas bandas de DNA, foi aplicado nos poços a mesma amostra, pois para a purificação do produto de PCR foi necessário excisar a banda em duas bandas de tamanho menores, seguindo as orientações do kit *QIAquick® Gel Extraction Kit* (QIAGEN).

Figura 16 - Análise da corrida eletroforética de gel de agarose 2% para validação da integridade do DNA purificado.



Fonte: Elaborado pelo autor. 1: Marcador de peso molecular *BenchTop 100 pb DNA ladder* (Promega). 2.1 e 2.2: Produto de PCR purificado da amostra 2. Há duas bandas, pois, a banda de gel utilizada para extração foi excisada em duas conforme orientações do kit de purificação utilizado.

4.4 Clonagem do produto de PCR purificado

4.4.1 Ligação do produto de PCR purificado ao vetor pGEM

Após a verificação do sucesso da purificação do produto de PCR, decidiu-se ligar a amostra a um vetor de clonagem *pGEM®-T Easy Vector Systems* (Promega) para que fosse feita a clonagem do produto de PCR purificado. A enzima polimerase presente no mastermix utilizado para amplificação do fragmento adiciona uma adenina (A) nas extremidades 3' do DNA amplificado, o que permite que o vetor, o qual é fabricado com extremidades timina (T)

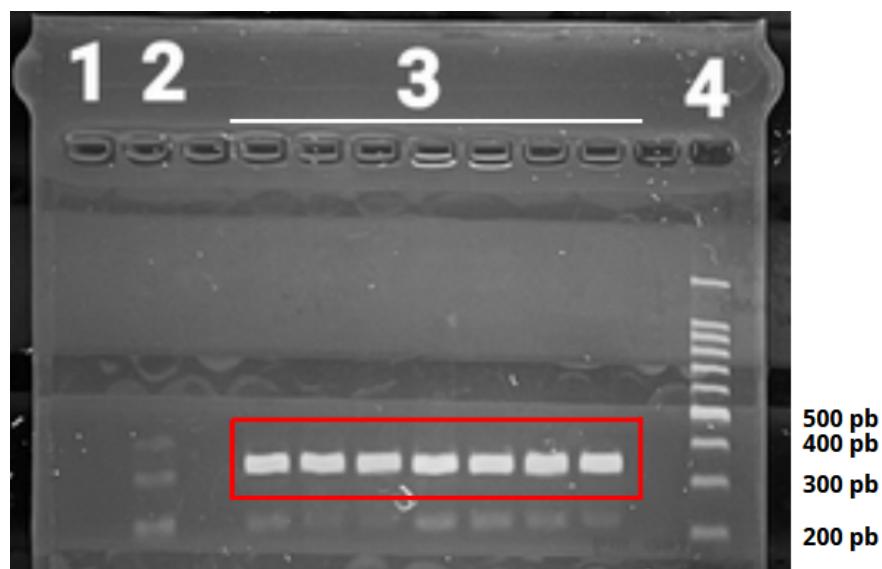
coesivas, ligue-se de forma eficiente ao produto de PCR, impedindo a recircularização do vetor (*pGEM®-T and pGEM®-T Easy Vector Systems Technical Handbook*, 2025).

O vetor além de possuir gene de resistência a ampicilina, permitindo a seleção de colônias transformadas, também contém uma região lacZ, que codifica uma subunidade da enzima β -galactosidase. Caso o vetor seja religado sem o inserto, a β -galactosidase será expressa pelas células, formando colônias de cor azul quando cultivadas em meios contendo IPTG e X-Gal. Se o inserto for clonado com sucesso, ele interrompe o gene lacZ, resultando em colônias brancas, indicando a presença do DNA de interesse (*pGEM®-T and pGEM®-T Easy Vector Systems Technical Handbook*, 2025).

4.4.2 PCR das colônias transformadas

Após a seleção de 7 colônias brancas transformadas, fez-se uma PCR de colônia para confirmação da transformação. Na Figura 17, há a análise eletroforética em gel de agarose 2% das bandas de DNA das colônias transformadas, mostrando um tamanho de entre 300-400 pb do amplicon.

Figura 17 - Análise eletroforética em gel de agarose 2% do DNA das colônias transformadas.

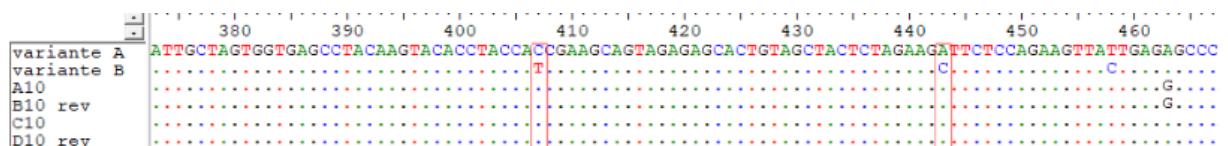


Fonte: Elaborado pelo autor. 1): NTC; 2): Controle positivo, clonado de pGEM β -caseína; 3): Produto de PCR das 7 colônias brancas escolhidas; 4): Marcador de peso molecular BenchTop 100 pb DNA ladder (Promega).

4.4.3 Sequenciamento de Sanger

Na Figura 18, vê-se o alinhamento das sequências de DNA das 4 colônias sequenciadas, evidenciando um possível genótipo de homozigose AA para as amostra sequenciadas. Cada colônia tem duas sequências, uma *forward* e uma *reverse*, sendo A10 e B10 as sequências *forward* e *reverse* de uma colônia, respectivamente, e C10 e D10 sequências *forward* e *reverse* de outra colônia, respectivamente. O alinhamento das sequências de nucleotídeos foi feito no software *BioEdit* versão 7.7.1.

Figura 18 - Alinhamento obtido após sequenciamento de Sanger do fragmento de κ -caseína da amostra sequenciada.



Fonte: Elaborado pelo autor. As duas primeiras linhas são as sequências referências conhecidas para as variantes A e B; A10 é a sequência *forward* da colônia 1 e B10 é a sequência reverso-complemento da mesma colônia, enquanto C10 é a sequência *forward* da colônia 2 e D10 é o reverso-complemento da mesma colônia. O sequenciamento aponta para o genótipo homozigoto AA para a amostra 2, sendo possível identificar os dois SNPs que diferenciam a variante A da variante B do gene de κ -caseína.

Na Figura 19, observa-se o alinhamento da sequência de aminoácidos de C10 com a sequência de aminoácidos referência da variante A de κ -caseína. O alinhamento da sequência de aminoácidos foi feito na ferramenta *BioEdit* versão 7.7.1.

Figura 19 - Alinhamento da sequência de aminoácidos de C10 com sequência de referência.

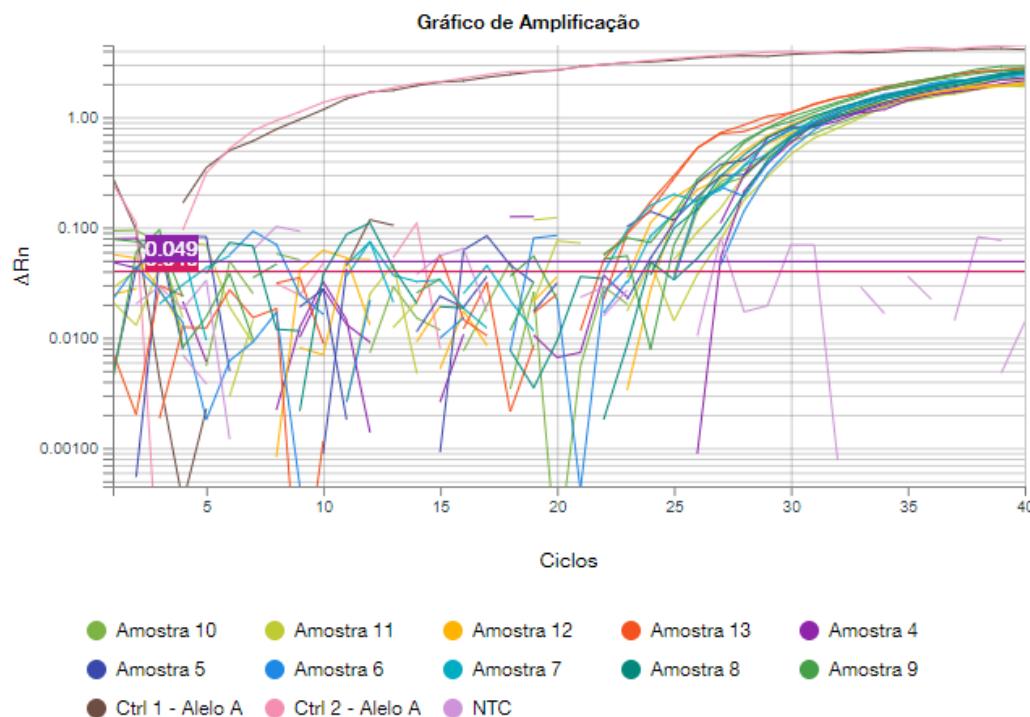


Fonte: Autor. Na figura, observa-se o alinhamento das sequências de referências para ambas as variantes de κ -caseína com as sequências das colônias sequenciadas

4.5 Testes de genotipagem com controle positivo da variante A

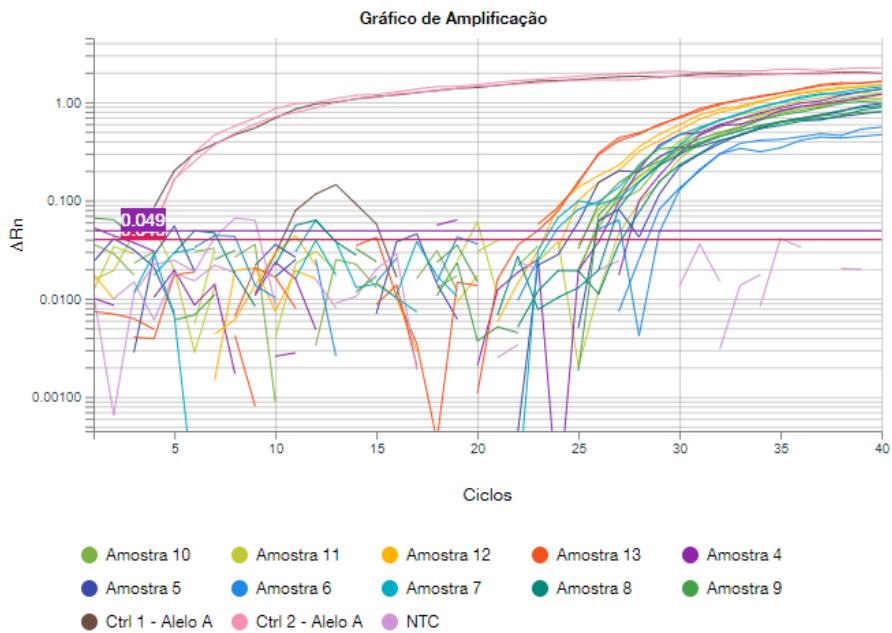
Após a obtenção do controle positivo da variante A para a κ -caseína, foi feita uma nova corrida para teste de genotipagem com 10 amostras diferentes de gDNA bovino, dessa vez com a adição de controle positivo para variante A (Figuras 20 e 21). Foi possível observar, após o término da corrida, um aumento do parâmetro de *confidence*, o qual ultrapassou o mínimo necessário de 95% para que o programa realizasse *calls* de genotipagem para as amostras. Dessa forma, o *software Design & Analysis v. 2.8.0*. atribuiu o genótipo homozigoto AA para todas as amostras testadas.

Figura 20 - Gráfico de amplificação da sonda para o alelo A do teste de genotipagem de 10 amostras.



Fonte: *Design & Analysis v2.8.0* (Applied Biosystems). Observa-se no gráfico o padrão de amplificação da sonda para o alelo A nas amostras testadas.

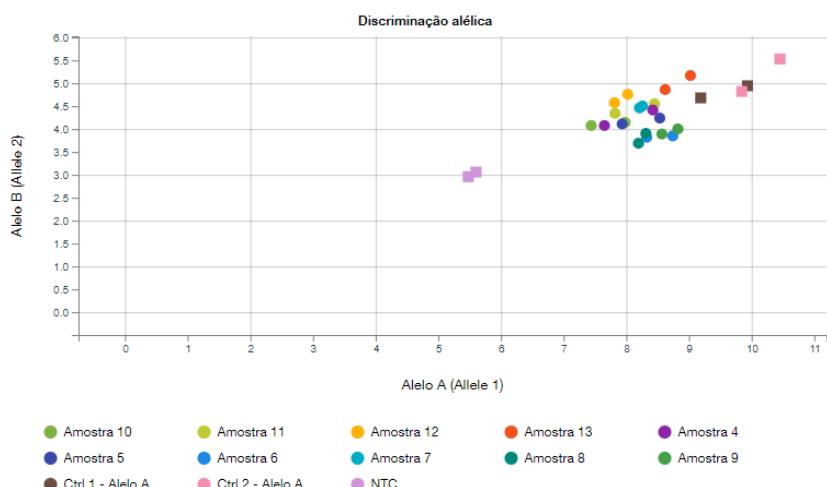
Figura 21 - Gráfico de amplificação da sonda para o alelo B do teste de genotipagem de 10 amostras.



Fonte: *Design & Analysis v2.8.0* (Applied Biosystems). Observa-se no gráfico o padrão de amplificação da sonda para o alelo B nas amostras testadas.

Com a adição do controle positivo da variante A, o software *Design & Analysis v2.8.0* já foi capaz de gerar uma *call* de genotipagem para cada amostra, sendo todas as amostras testadas atribuídas o genótipo de homozigose para o alelo A (Figura 22).

Figura 22 - Gráfico de discriminação alélica de 10 amostras de gDNA bovino de fazendas variadas.



Fonte: *Design & Analysis v2.8.0* (Applied Biosystems). Nota-se uma distribuição espacial melhor das amostras feita pelo programa. Contudo, a falta de controle positivo para variante B ainda dificulta a discriminação das amostras com mais confiabilidade.

Essa aparente predominância da variante A está de acordo com os estudos de Deb et al (2014), Djedović et al (2015), Barbosa et al (2019) e Bangar et al (2021) sobre a frequência alélica dos polimorfismos de κ -caseína em rebanhos de bovinos, observando uma alta frequência do alelo A nas populações das raças estudadas por eles, como a raça Girolando, a qual é responsável por cerca de 80% do leite produzido no Brasil (Gonsales, 2021). A baixa frequência do alelo B do gene de κ -caseína entre as amostras testadas também corrobora com os estudos de Azevedo et al (2008) e Barbosa et al (2019), os quais registraram baixa frequência dessa variante nas populações testadas em seus estudos.

Técnicas moleculares como a desenvolvida neste estudo permitem gerar um panorama do perfil genético de rebanhos de bovinos leiteiros cerarenses de forma mais facilitada do que as técnicas tradicionais.

Barbosa et al (2019), por exemplo, afirma que o conhecimento das proporções dos diferentes genótipos e alelos dos genes de interesse, como o da κ -caseína, pode auxiliar no desenvolvimento de seleções assistidas por marcadores (SAM), permitindo o aumento da produtividade leiteira e a melhoria das características físico-químicas e tecnológicas do leite. Ademais, segundo Drogemuller et al (2001), a combinação da SAM com as ferramentas tradicionais de seleção permitem a obtenção de progresso genético de forma mais rápida, acelerando

Já existem outros estudos na literatura que obtiveram sucesso ao utilizar a técnica de PCR em tempo real para genotipagem do gene de κ -caseína (Kovalchuk; Tagmazyan; Klimov, 2019; Ladyka et al., 2022), além de validarem seus resultados com outras técnicas de genotipagem, como PCR-RFLP, comprovando a possibilidade do desenvolvimento de teste de genotipagem por qPCR para análise dos polimorfismos de κ -caseína como marcadores de produtividade. Dentre as principais vantagens dessa técnica, quando comparados a outros métodos de genotipagem, como PCR-RFLP, pode-se citar: 1) rapidez, ; 2) praticidade, asdasd; 3) custo-benefício; 4) sensibilidade, uma vez que consegue detectar quantidades mínimas de material genético (20 ng/ μ L, por exemplo) e 5) especificidade, devido às *primers* e sondas de hidrólise que são desenhados para anelamento específico ao fragmento do gene de interesse.

Dessa forma, recomenda-se a obtenção de controle positivo para a variante B, de modo a oferecer ao *software* mais informações para que seja feita uma discriminação alélica mais específica e eficiente. Assim, devem ser desenhados *primers* mutagênicos para a realização de uma mutagênese sítio-dirigida que modifique o DNA controle da variante A em B.

Após obtenção do controle positivo para variante B, recomenda-se também a realização de mais testes de genotipagem, buscando otimizar o processo e analisar o padrão de discriminação alélica formada pelo programa.

5 CONCLUSÃO

A extração do DNA genômico da amostra de sangue total dos bovinos e a amplificação do fragmento de κ -caseína nos procedimentos de PCR e qPCR foram bem-sucedidas, com bandas de tamanho esperado. A clonagem do controle positivo de variante alélica A do gene de κ -caseína foi realizada com êxito, sendo confirmada pelo sequenciamento de Sanger posterior ao procedimento. A adição de um controle positivo às reações de genotipagem por qPCR aumentou as taxas de *confidence* do software utilizado, proporcionando maior confiabilidade às *calls* de genótipos atribuídas às amostras de gDNA bovino testadas. Avaliando os resultados obtidos, é possível observar que, dada à especificidade e sensibilidade conferida à técnica por meio do uso de sondas de hidrólise específicas para a identificação dos SNPs do gene CSN3, a PCR em tempo real apresenta indícios de ser uma técnica viável e confiável para a genotipagem da κ -caseína, contudo ainda faltam mais etapas de averiguação e confirmação da eficiência da genotipagem desses polimorfismos por qPCR.

6 REFERÊNCIAS

- AGREGÁN, R. et al. Proteomic Advances in Milk and Dairy Products. **Molecules**, v. 26, n. 13, p. 3832, 23 jun. 2021.
- AL-SAMARAI, F. R.; AL-KAZAZ, A. A. Molecular Markers: an Introduction and Applications. **European Journal of Molecular Biotechnology**, v. 9, n. 3, p. 118–130, 2015.
- ALBAZI, W. et al. Association of prolactin gene polymorphism with some biochemical and lactation traits in dairy cow in Karbala province. **Turkish Journal of Physiotherapy and Rehabilitation**, v. 32, n. 3, 2021.
- ALBAZI, W. J. et al. Association of Kappa casein gene polymorphism with milk production traits in crossbred dairy cows. **Ciência Animal Brasileira**, v. 24, 2023.
- AND. **Sanger Sequencing and Next-Generation Sequencing: Detailed Procedures and Applications.** Disponível em: <http://www.preachbio.com/2024/07/sanger-sequencing-and-next-generation.html>. Acesso em: 25 fev. 2025.
- APPLIED BIOSYSTEMS. **Allelic Discrimination Getting Started Guide.** Disponível em: https://assets.thermofisher.com/TFS-ArtsAssets/LSG/manuals/cms_042315.pdf. Acesso em: 11 fev. 2025.
- ASMARASARI, S. A. et al. Kappa casein (CSN3) gene polymorphism and its effect on cumulative milk yields of Holstein Friesian dairy cattle. **IOP Conference Series: Earth and Environmental Science**, v. 902, n. 1, p. 012047, 2021.
- AZEVEDO, A. L. S. et al. Genetic polymorphism of the kappa-casein gene in Brazilian cattle. **Genetics and Molecular Research**, v. 7, n. 3, p. 623–630, 2008.
- BANGAR, Y. C. et al. Genetic polymorphisms of kappa casein gene and its association with milk and composition traits in cows: An updated meta-analysis. **Meta Gene**, p. 100948, 2021.
- BARBOSA, S. B. P. et al. Genetic association of variations in the kappa-casein and β-lactoglobulin genes with milk traits in girolando cattle. **Revista Brasileira de Saúde e Produção Animal**, v. 20, 2019.
- BRASIL. Portaria Interministerial MDA/MDS/MAPA nº 5, de 30 de agosto de 2024.** Institui a Estratégia de Desenvolvimento da Produção de Leite na Agricultura Familiar. Diário Oficial da União, Brasília, DF, 30 ago. 2024. Disponível em: <https://frentemineira.com.br/2024/09/portaria-interministerial-mdm-mds-mapam-no-5-de-30-de-agosto-de-2024/>. Acesso em: 20 fev. 2025.
- BRITO, L. F. et al. Genetic selection for milk production and its impact on dairy cattle health: A review. **Journal of Dairy Science**, v. 103, n. 1, p. 1-15, 2020.
- BRITTEN, M.; GIROUX, H. J. Rennet coagulation of heated milk: A review. **International Dairy Journal**, v. 124, p. 105179, 2021.

CHANOKPHAT PHADUNGATH. Casein micelle structure: a concise review. **Songklanakarin Journal of Science and Technology (SJST)**, v. 27, n. 1, p. 201–212, 2005.

CILEITE - CENTRO DE INTELIGÊNCIA DO LEITE. **Leite em Números: produção e produtividade**. 2025. Disponível em: https://www.cileite.com.br/leite_numeros_producao. Acesso em: 03 mar. 2025.

CROSSLEY, B. M. et al. Guidelines for Sanger sequencing and molecular assay monitoring. **Journal of Veterinary Diagnostic Investigation**, v. 32, n. 6, p. 767–775, 2020.

DALGLEISH, D. G.; CORREDIG, M. The Structure of the Casein Micelle of Milk and Its Changes During Processing. **Annual Review of Food Science and Technology**, v. 3, n. 1, p. 449–467, 2012.

DE KRUIF, C. G.; HOLT, C. Casein Micelle Structure, Functions and Interactions. **Advanced Dairy Chemistry—1 Proteins**, p. 233–276, 2003.

DEB, Rajib. et al. Genetic polymorphism and association of kappa-casein gene with milk production traits among Frieswal (HF × Sahiwal) cross breed of Indian origin. **Iranian journal of veterinary research**, v. 15, n. 4, p. 406–8, 2014.

DJEDOVIC, R. et al. Relationship between genetic polymorphism of κ-casein and quantitative milk yield traits in cattle breeds and crossbreds in Serbia. **Genetika**, v. 47, n. 1, p. 23–32, 2015.

DROGEMULLER, C.; HAMANN, H.; DISTL, O. Candidate gene markers for litter size in different German pig lines. **Journal of Animal Science**, v. 79, n. 10, p. 2565, 2001.

EMBRAPA RECURSOS GENÉTICOS E BIOTECNOLOGIA. **Genotipagem**. 2021. Disponível em: <https://www.embrapa.br/en/agencia-de-informacao-tecnologica/tematicas/reproducao-animal/technologias/genotipagem>. Acesso em: 08 fevereiro 2025.

FALEIRO, F. G. **Marcadores genético-moleculares aplicados a programas de conservação e uso de recursos genéticos**. [s.l.] Planaltina, DF: Embrapa Cerrados, 2007., 2007. Disponível em: <http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/570287>. Acesso em: 14 fev. 2025.

FARRELL, H. M. et al. Casein micelle structure: What can be learned from milk synthesis and structural biology? **Current Opinion in Colloid & Interface Science**, v. 11, n. 2-3, p. 135–147, jun. 2006.

FERREIRA, M.E.; GRATTAPAGLIA, D. Introdução ao uso de marcadores moleculares em análise genética. 3^a ed. Brasília: EMBRAPA-CENARGEN, 1998, pp. 220.

FOX, P. F. et al. **Dairy Chemistry and Biochemistry**. 2. ed. Cham: Springer International Publishing, 2015. p. 149–239

FOX, P. F. et al. Cheese Yield. **Springer eBooks**, p. 279–331, 23 ago. 2017.

G. CERIOTTI et al. Milk protein loci polymorphism in taurine (*Bos taurus*) and zebu (*Bos indicus*) populations bred in hot climate. **Journal of Animal Breeding and Genetics**, v. 121, n. 6, p. 404–415, 24 nov. 2004.

GARCÍA-RUIZ, A. et al. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. **Proceedings of the National Academy of Sciences**, v. 113, n. 28, p. E3995-E4004, 2016.

GARIBYAN, L.; AVASHIA, N. Research Techniques Made Simple: Polymerase Chain Reaction (PCR). **Journal of Investigative Dermatology**, v. 133, n. 3, p. 1–4, 2013.

GONSALES, S. A. **Girolando: produção de leite, características e história | MilkPoint**. Disponível em: <https://www.milkpoint.com.br/artigos/producao-de-leite/raca-girolando-historia-caracteristica-s-e-producao-224447/>. Acesso em: 05 mar. 2025.

GOULDING, D. A.; FOX, P. F.; O'MAHONY, J. A. Milk proteins: An overview. **Milk Proteins**, p. 21–98, 2020.

GREEN. M. R; SAMBROOK, J. **Molecular Cloning: A Laboratory Manual**. 4. ed. New York. Cold Spring Harbor Laboratory Press, 2012.

GROVER, A.; SHARMA, P. C. Development and use of molecular markers: past and present. **Critical Reviews in Biotechnology**, v. 36, n. 2, p. 290–302, 2014.

HANI, H. A.; ALBAZI, W. J.; ALI, H. Association of prolactin gene polymorphism with some biochemical and lactation traits in dairy cows in Karbala province. **Turkish Journal of Physiotherapy and Rehabilitation**. v. 32, n. 3, 2021.

HETTINGA, K.; BIJL, E. **Can recombinant milk proteins replace those produced by animals?** Current Opinion in Biotechnology, v. 75, p. 102690, 2022.

HIGUCHI, R. et al. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. **Bio/technology (Nature Publishing Company)**, v. 11, n. 9, p. 1026–30, 1993.

HOLT, C. Structure and Stability of Bovine Casein Micelles. **Advances in Protein Chemistry**, v. 43, p. 63–151, 1992.

HOLT, C. et al. Invited review: Caseins and the casein micelle: Their biological functions, structures, and behavior in foods. **Journal of Dairy Science**, v. 96, n. 10, p. 6127–6146, out. 2013.

HORNE, D. S. Casein micelle structure and stability. **Milk Proteins**, p. 213–250, 2020.

HUIAM, M. E.; ELTAHIR, S. S.; HAMZA, A. E. Molecular Characterization of Genetic Variability among Sudanese Baggara Cattle within Kappa Casein Csn3 Gene (Exon V). **Molecular Biology**, v. 7, n. 1, 2017.

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produção de origem animal, por tipo de produto, segundo o Brasil, as Grandes Regiões e as Unidades da Federação**. 2023. Disponível em:

<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9107-producao-da-pecuaria-municipal.html?=&t=resultados>. Acesso em 04 mar. 2025.

JOHNSON, W. M. The Polymerase Chain Reaction: An Overview and Development of Diagnostic PCR Protocols at the LCDC. **Canadian Journal of Infectious Diseases**, v. 2, n. 2, p. 89–91, 1991.

KIM, S.; MISRA, A. SNP Genotyping: Technologies and Biomedical Applications. **Annual Review of Biomedical Engineering**, v. 9, n. 1, p. 289–320, 15 ago. 2007.

KOVALCHUK, S.; TAGMAZYAN, A.; KLIMOV, E. A Novel Test System for Genotyping rs43703016 Single-nucleotide Substitutions in the Bovine CSN3 Gene. **Annual Research & Review in Biology**, p. 1–5, 30 jul. 2019.

KRISHNA, T. C. et al. Influence of Ultra-Heat Treatment on Properties of Milk Proteins. **Polymers**, v. 13, n. 18, p. 3164, 2021.

KUBISTA, M. et al. The real-time polymerase chain reaction. **Molecular Aspects of Medicine**, v. 27, n. 2-3, p. 95–125, 2006.

LADYKA, V. et al. Evaluation of cow genotypes by kappa-casein of dairy breeds. **Acta fytotechnica et zootechnica/Acta fytotechnica et zootechnica**, v. 25, n. 1, 2022.

LORENZ, T. C. Polymerase Chain Reaction: Basic Protocol plus Troubleshooting and Optimization Strategies. **Journal of Visualized Experiments**, v. 63, n. 63, 2012.

M. HABIB YAHYOUI et al. Genetic polymorphism of the caprine kappa casein gene. **Journal of Dairy Research**, v. 68, n. 2, p. 209–216, 2001.

MCDONALD, C.; TAYLOR, D.; LINACRE, A. PCR in Forensic Science: A Critical Review. **Genes**, v. 15, n. 4, p. 438, 2024.

MCMAHON, D. J.; OOMMEN, B. S. Supramolecular Structure of the Casein Micelle. **Journal of Dairy Science**, v. 91, n. 5, p. 1709–1721, 2008.

MCMANUS, C. et al. O uso da genética na adaptação dos bovinos aos ambientes sob estresse. **Revista Brasileira de Buiatria**, v. 3, n. 1, p. 1–25, 2022.

ME, H.; SS, E.; AE, H. Molecular Characterization of Genetic Variability among Sudanese Baggara Cattle within Kappa Casein Csn3 Gene (Exon V). **Molecular Biology**, v. 7, n. 1, 2017.

MILLER, J. H. **Experiments in molecular genetics**. [Cold Spring Harbor, N.Y.]: Cold Spring Harbor Laboratory, 1972.

NASCIMENTO, S.; SUAREZ, E. R.; PINHAL, M. A. DA S. Tecnologia de PCR e RT-PCR em tempo real e suas aplicações na área médica. **RBM rev. bras. med**, 2010.

NOBRE, R. **Ceará expandiu produção de leite em 53,5 milhões de litros em 2023**. Disponível em: <https://www.opovo.com.br/noticias/economia/2024/03/14/ceara-expandiu-producao-de-leite-e-m-535-milhoes-de-litros-em-2023.html>. Acesso em: 3 mar. 2025.

OLIVEIRA, A. J. et al. Principais marcadores moleculares. **Research, Society and Development**, v. 10, n. 15, p. e562101523633–e562101523633, 26 nov. 2021.

PANOUILLÉ, M.; NICOLAI, T.; DURAND, D. Heat induced aggregation and gelation of casein submicelles. **International Dairy Journal**, v. 14, n. 4, p. 297–303, 17 dez. 2003.

PEREZ, B. C. et al. Seleção genômica aplicada a pequenas populações de gado de leite. In: BALIEIRO, J. C. et al. (org). **Novos Desafios da Pesquisa em Nutrição e Produção Animal**, Pirassununga: 5D Editora, 2018. 294 p. il.

pGEM ® -T and pGEM ® -T Easy Vector Systems. [s.l: s.n]. Disponível em: <<https://www.promega.com.br/-/media/files/resources/protocols/technical-manuals/0/pgem-t-and-pgem-t-easy-vector-systems-protocol.pdf?rev=42b7715403bf452cb4487ac142d2c61c&la=en>>. Acesso em: 13 fev. 2025.

PISHCHAN, S. G.; SYLYCHENKO, K. A. Characteristics of kappa-casein gene polymorphism in cows of Swiss breed and their productive qualities at a large dairy. **Regulatory Mechanisms in Biosystems**, v. 12, n. 3, p. 513–518, 28 ago. 2021.

PORTO, R. Pecuária leiteira no Ceará: tradição e tecnologia impulsionam produção no Sertão Central. 2025. Disponível em: <<https://ibandce.com.br/pecuaria-leiteira-no-ceara-tradicao-e-tecnologia-impulsionam-producao-no-sertao-central/>>. Acesso em: 3 mar. 2025.

PRYCE, J. E. et al. Invited review: Opportunities for genetic improvement of metabolic diseases. **Journal of Dairy Science**, v. 100, n. 8, p. 6855–6873, 2017.

ROCHA, D. T. da; CARVALHO, G. R.; RESENDE, J. C. de. **Cadeia produtiva do leite no Brasil: produção primária.** 2020. Embrapa Gado de Leite. Disponível em: <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1124858/1/CT-123.pdf>. Acesso em: 03 de mar. 2025.

SADIQ, U.; GILL, H.; CHANDRAPALA, J. Casein Micelles as an Emerging Delivery System for Bioactive Food Components. **Foods**, v. 10, n. 8, p. 1965, 23 ago. 2021.

SCHLEINITZ, D.; DISTEFANO, J. K.; KOVACS, P. Targeted SNP Genotyping Using the TaqMan® Assay. **Methods in Molecular Biology**, p. 77–87, 7 nov. 2010.

SCHMIDT, D. **Association of caseins and casein micelle structure.** Disponível em: <<https://www.semanticscholar.org/paper/Association-of-caseins-and-casein-micelle-structure-Schmidt/1235665f7484f9e7725ad259d0123c81d203c8f4>>. Acesso em: 13 fev. 2025.

SERGEEVICH, O. A. et al. **Real-time PCR: advantages and limitations.** In: Publisher. agency: Proceedings of the 6th International Scientific Conference «Reviews of Modern Science» (May 16-17, 2024). Zürich, Switzerland, 2024. 394p. Universität Luzern, 2024. p. 285.

SHAFFER, C. **Challenges with Sanger Sequencing.** 2023. Disponível em: <<https://www.news-medical.net/life-sciences/Challenges-with-Sanger-Sequencing.aspx>>. Acesso em: 11 fev. 2025.

SILVA, H. P.; DEMUNER, L. F. MELHORAMENTO GENÉTICO DE BOVINOS LEITEIROS: ESTUDO DAS TÉCNICAS E IMPACTOS DO MELHORAMENTO GENÉTICO NA PRODUTIVIDADE E QUALIDADE DO LEITE. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, v. 10, n. 12, p. 988–1003, 2024.

SMITH, G. N. et al. Casein micelles in milk as sticky spheres. **Soft Matter**, v. 16, n. 43, p. 9955–9963, 11 nov. 2020.

SUPROVICH, T.; MOKHNACHOVA, N. Gene polymorphism of economically-useful traits in Ukrainian gray cattle breed. **The Animal Biology**, v. 19, n. 1, p. 111–118, mar. 2017.

THERMO FISHER. **The Purpose of ROX Reference Dye in Real-Time PCR (qPCR)**. 2021. Disponível em: <https://www.thermofisher.com/blog/behindthebench/the-purpose-of-rox-reference-dye-in-real-time-pcr-qpcr/>. Acesso em: 04 mar. 2025.

THERMO FISHER. **Genotyping Analysis Module User Guide**. 2023. Disponível em: https://assets.thermofisher.com/TFS-Arts/LSG/manuals/MAN0018749_GenotypingAnalysisModule_QSDASW_v2_UG.pdf. Acesso em: 11 fev. 2025.

TURCHETTO-ZOLET, A. C. et al. Marcadores Moleculares na Era Genômica: metodologias e aplicações. Ribeirão Preto: Sociedade Brasileira de Genética, 2017. 181 p. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/206115/001056160.pdf>. Acesso em: 12 fev. 2025.

VALENCIA, C. A. et al. Sanger Sequencing Principles, History, and Landmarks. **SpringerBriefs in Genetics**, p. 3–11, 2013.

VASCONCELOS, H. (Brasil). **IBGE: 2023 apresentou crescimento na produção brasileira total de leite**. 2024. MilkPoint. Disponível em: <https://www.milkpoint.com.br/noticias-e-mercado/panorama-mercado/ibge-2023-apresentou-crescimento-na-producao-brasileira-total-de-leite-237479/>. Acesso em: 03 mar. 2025.

VELMATOV, A. et al. KAPPA-CASEIN GENE POLYMORPHISM AND ITS EFFECT ON MILK OF RED AND WHITE COWS FOR CHEESEMAKING. v. 20, n. 2, p. 972–5210, 2020.

VIGNAL, A. et al. A review on SNP and other types of molecular markers and their use in animal genetics. **Genet. Sel. Evol.**, v. 34, p. 275, 2002.

VOLKANDARI, S. D.; INDRIAWATI, I.; MARGAWATI, E. T. Genetic polymorphism of kappa-casein gene in Friesian Holstein: a basic selection of dairy cattle superiority. **Journal of the Indonesian Tropical Animal Agriculture**, v. 42, n. 4, p. 213, 19 dez. 2017.

WALSTRA, P. Casein sub-micelles: do they exist? **International Dairy Journal**, v. 9, n. 3-6, p. 189–192, mar. 1999.

WUSIGALE; LIANG, L.; LUO, Y. Casein and pectin: Structures, interactions, and applications. **Trends in Food Science & Technology**, v. 97, p. 391–403, 1 mar. 2020.

YANG, W. et al. Review on the development of genotyping methods for assessing farm animal diversity. **Journal of Animal Science and Biotechnology**, v. 4, n. 1, 23 jan. 2013.

ZHOU, B. et al. Interfacial properties of milk proteins: A review. **Advances in Colloid and Interface Science**, v. 295, p. 102347, 1 set. 2021.