Received: 6 July 2022 | Revised: 12 March 2023 | Accepted: 10 April 2023

DOI: 10.1111/jbg.12777

ORIGINAL ARTICLE

Journal of Animal Breeding and Genetics WILEY

Phenotypic variation in milk fatty acid composition and its association with stearoyl-CoA desaturase 1 (*SCD1*) gene polymorphisms in Gir cows

Carolina Guimarães Ramos Matosinho[1] | **Pablo Augusto de Souza Fonseca**[2] | **Maria Gabriela Campolina Diniz Peixoto**[3] | **Izinara Cruz Rosse**[4] | **Fernando César Ferraz Lopes**[3] | **Thalia Zózimo**[1] | **Aníbal Eugênio Vercesi Filho**[5] | **Frank Ângelo Tomita Bruneli**[3] | **Maria Raquel Santos Carvalho**[1] | **Marco Antonio Sundfeld Gama**[3,#]

[1]Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

[2]Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, Ontario, Canada

[3]Embrapa Dairy Cattle, Juiz de Fora, Minas Gerais, Brazil

[4]Departamento de Farmácia, Escola de Farmácia, Universidade Federal de Ouro Preto, Ouro Preto, MG, Brazil

[5]Agência Paulista de Tecnologia dos Agronegócios e diretor do Laboratório de Biotecnologia do Instituto de Zootecnia do Governo do Estado de São Paulo, São Paulo, SP, Brazil

Correspondence
Pablo Augusto de Souza Fonseca, Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, Ontario N1G 2W1, Canada.
Email: pfonseca@uoguelph.ca

Funding information
Brazilian National Research Council; Coordination for the Improvement of Higher Education Personnel; Fundação de Amparo à Pesquisa do Estado de Minas Gerais

Abstract

Individual variation in milk fatty acid (FA) composition has been partially attributed to stearoyl-CoA desaturase 1 (*SCD1*) gene polymorphisms in taurine breeds, but much less is known for Zebu breeds. This study investigated the phenotypic variation in milk FA composition, and the influence of *SCD1* variants on this trait and on milk fat desaturase indices (DI) in Gir cows. The functional impact of *SCD1* variants was predicted using bioinformatics tools. Milk and blood samples were collected from 312 cows distributed in 10 herds from five states of Brazil. *SCD1* variants were identified through target sequencing, and milk FA composition was determined by gas chromatography. Phenotypic variation in milk FA composition fell within the range reported for taurine breeds, with SCD_{18} index showing the lowest variation among the DI. Fourteen *SCD1* variants were identified, six of which not previously described. Regarding the A293V polymorphism, all cows were homozygous for the C allele (coding for alanine), whereas all genotypes were detected for the second SNP affecting the 293 codon (G > A), with compelling evidence for functional effects. Significant associations (based on raw

[#]Principal investigator

Gama and Carvalho should be considered joint senior author.

p-values) were found between this SNP and C12:0, *cis*-9, *trans*-11 CLA and short-chain FA, and between another SNP (rs523411937) and C15:0 and odd-chain linear FA. A new SNP on Chr26:21277069 was associated with *trans*-11 C18:1, *cis*-9, *trans*-11 CLA, C18:3 n-3 and n-3 FA. These findings indicate that *SCD1* polymorphisms also contributes to the phenotypic variation in milk FA composition of Gir cows, with potential use in their breeding programmes.

KEYWORDS

Δ-9 desaturase, dairy fat, genetic variants, human health, lipogenic enzymes, zebu breeds

1 | INTRODUCTION

Milk and dairy products have long been acknowledged as part of a healthy diet in most food-based dietary guidelines (Herforth et al., 2019), as this group of nutrient-dense foods is a rich source of high-quality protein and makes a significant contribution to the requirements for calcium, magnesium, selenium, riboflavin, vitamin B12 and pantothenic acid (Weaver, 2014). However, milk and dairy products are also a significant source of saturated fatty acids (SFA) in the human diet, and the intake of SFA-rich foods has long been linked to an increased risk of cardiovascular diseases (CVD) due to adverse effects on blood lipids, particularly on levels of low-density lipoprotein (LDL) cholesterol (Astrup et al., 2020). As a result, public health authorities from many countries have long advised the consumption of low-fat and fat-free versions instead of regular-fat dairy foods. Nevertheless, emerging evidence indicates that the consumption of regular-fat dairy foods has a neutral or inverse association with risk of atherosclerotic cardiovascular disease, type 2 diabetes and associated risk factors (Astrup, 2014; Dariush Mozaffarian, 2015; Hirahatake et al., 2020; Kratz et al., 2013).

The reported beneficial effects of regular-fat dairy intake on cardiometabolic health may be partially attributed to the presence of numerous health-promoting compounds in dairy fat, such as C4:0 (butyric acid), odd- and branched-chain fatty acids (OBCFA), *trans*-11 C18:1 (vaccenic acid, VA), *cis*-9, *trans*-11 CLA (rumenic acid, RA), *cis*-9 C18:1 (oleic acid) and C18:3 n-3 (α-linolenic acid) (Gómez-Cortés et al., 2018; Kratz et al., 2013). Thus, different strategies have been undertaken to increase the concentrations of these bioactive FA in milk fat while reducing the proportions of hypercholesterolemic SFA, notably C14:0 and C16:0 (Kliem & Shingfield, 2016).

Evidence indicate that diet has a major influence on the milk FA composition of cows, goats and ewes (Shingfield et al., 2013), but genetics has also been shown to play a key role (Samková et al., 2012). Studies with taurine cattle breeds have reported low to moderate heritability values for several milk FA, indicating the possibility of using genetic selection to improve the nutritional quality of milk fat (Arnould & Soyeurt, 2009; Bastin et al., 2012; Bobe et al., 2008; Carvajal et al., 2016; Garnsworthy et al., 2010; Mele et al., 2007; Pegolo et al., 2016; Schennink et al., 2008; Soyeurt et al., 2007; Stoop et al., 2008). Compared with conventional selection programmes based on performance recording and prediction of breeding values, faster progress towards the production of milk and dairy products containing a more desirable FA composition could be achieved through genomic or marker-assisted selection based on the identification of polymorphisms in genes involved in milk fat synthesis. Among the several enzymes regulating the lipid metabolism in the mammary gland, the stearoyl-CoA desaturase 1 enzyme (SCD1; also known as Δ9-desaturase) plays a key role by introducing a double bond at the Δ9-position in a large spectrum of FA (Ntambi & Miyazaki, 2004). The most important SCD1 substrates are C14:0, C16:0, C18:0 and VA, which are converted into *cis*-9 C14:1, *cis*-9 C16:1, *cis*-9 C18:1 and RA, respectively (Corl et al., 2001).

The ratios of products to precursors of SCD1 in milk fat, also known as desaturase indices (DI) (Kelsey et al., 2003), are commonly used as a proxy for SCD1 activity in the mammary gland (Garnsworthy et al., 2010). Most of the oleic acid and RA secreted in cow's milk are derived from endogenous synthesis by SCD1 action in the mammary gland (Bernard et al., 2013), which shows the key role of this enzyme in determining the contents of these health-beneficial FA in milk fat. In taurine breeds, a significant proportion of the variation in milk FA composition has been associated with the occurrence of a non-conservative polymorphism in the SCD1 protein (A293V) in which an alanine is replaced with a valine on the 293rd residue of the amino acid sequence (Conte et al., 2010; Kgwatalala et al., 2007; Li et al., 2016; Mele et al., 2007; Schennink et al., 2008).

In Brazil, cow's milk is derived from taurine cattle (*Bos taurus*), Zebu cattle (*Bos indicus*) and their crosses (Madalena et al., 2012). The two predominant Zebu breeds,

Gir and Guzerá, have been selected over the last 30 years for milk production and solids content through breeding programmes based on progeny tests and/or on multiple ovulation and embryo transfer (MOET) nucleus selection schemes (Bruneli et al., 2021; Peixoto et al., 2010; Santana et al., 2014). The main factors contributing to variation in milk production and composition of these breeds have been investigated (Rosse et al., 2014; Shahlla et al., 2014), but there is limited information regarding the genetic influence on milk FA composition in those breeds. In a pilot study (Freitas et al., 2013; Gama et al., 2013), our research group found a substantial between-cow variation in DI (to a lesser extent for the C18 index) and in the proportions of *cis*-9 C18:1 and RA in milk fat of Gir and Guzerá cows fed on the same diet, indicating a genetic influence on SCD1 activity in both breeds.

To the best of our knowledge, there is a lack of studies looking at variants in the *SCD1* gene and their association with milk FA composition in Zebu breeds. Given the particular relevance of Gir cows and their crosses for milk production in Brazil, we carried out the present study with the following objectives: (1) to estimate the phenotypic variation in milk FA composition of Gir cows from Brazilian herds; (2) to develop a targeted sequencing methodology looking for breed-specific variants in *SCD1* and genotyping by sequencing; (3) to undertake an association study between the *SCD1* variants and proportions of nutritionally relevant FA and FA groups, n-6:n-3 FA ratio and DI in milk fat; (4) to perform bioinformatic analysis on potential new *SCD1* variants, in the search for evidence of functionality. We hypothesized that variants in the *SCD1* gene, some of which possibly not previously reported in the literature, would be significantly associated with DI and proportions of nutritionally relevant FA in milk fat, which may allow for the use of marker-assisted selection in future breeding programmes for Gir cows and its crosses.

2 | MATERIALS AND METHODS

2.1 | Animal welfare statement

All animal procedures were performed in accordance with the guidelines and recommendations of the Animal Care and Ethics Committee of Embrapa Dairy Cattle, under protocol number 11/2015.

2.2 | Sample and data collection

Individual milk and blood samples (one sample per cow) were collected between December 2016 and October 2018 from 312 primiparous and multiparous Gir cows between 12 and 214 days in milk. The cows were distributed in 10 different herds (ranging from nine to 92 samples per herd) located in the Brazilian states of Minas Gerais, São Paulo, Paraíba, and Rio Grande do Norte. Detailed information on the genetic architecture and management practices of the Gir breed in Brazil can be found elsewhere (Milanesi et al., 2022; Peripolli et al., 2018, 2020; Santana et al., 2014). Diets offered to cows during the period of sample collection were recorded and grouped into categories as described in detail later in this section.

Composite milk samples from morning and afternoon milking were collected in 15-mL Falcon tubes and immediately frozen at −20°C until analysis of fatty acid composition. Blood samples were collected in vacutainer tubes containing K_2EDTA and immediately frozen at −20°C until DNA extraction.

2.3 | Analysis of milk fatty acid composition

Milk samples in 15-mL Falcon tubes were thawed at room temperature, and a 1 mL volume was used for lipid extraction according to AOAC Official Method 0.05 (AOAC International, 2012). Following solvent evaporation at 40°C under oxygen-free nitrogen, extracted milk lipids were dissolved in hexane and methyl acetate and transesterified to fatty acid methyl esters (FAME) using freshly prepared methanolic sodium methoxide, as described elsewhere (Baldin et al., 2013). The FAME in a 1.0 µL sample were separated and quantified using a gas chromatograph (model 7820A, Agilent Technologies Inc.) fitted with a flame-ionization detector and equipped with a CP-Sil 88 fused-silica capillary column (100 m × 0.25 mm × 0.2 µm film thickness; Varian Inc). Operating conditions were the same as described by (Cruz-Hernandez et al., 2007). The FAME were identified by comparison of retention times with reference FAME standards (Sigma-Aldrich®; Larodan AB, Stockholm, Sweden; Luta-CLA® 60, BASF); minor *trans/cis*-18:1 isomers and *trans*-9, *cis*-11 CLA were identified according to their order of elution reported under the same analytical conditions (Cruz-Hernandez et al., 2007). Milk FA composition was expressed as a weight percentage of total FA using theoretical response factors (Wolff et al., 1995). SCD1 indices (also referred to as DI) were calculated for four pairs of FA (*cis*-9 14:1/14:0, *cis*-9 16:1/16:0, *cis*-9 18:1/18:0 and *cis*-9, *trans*-11 CLA/*trans*-11 18:1) by expressing each product as a proportion of precursor *plus* product (Kelsey et al., 2003). Also, an overall SCD1 index was calculated using all the selected pairs of products and precursors.

In addition, the n-6:n-3 FA ratio and proportions of specific FA groups (n-3 FA, n-6 FA, long-chain n-3 FA, milk fat inhibitors, odd-chain linear SFA, branched-chain SFA, short-chain SFA and medium-chain SFA) in milk fat were also calculated given their nutritional and physiological relevance for human health and mammary lipid metabolism, respectively (Gómez-Cortés et al., 2018; Koch & Lascano, 2018). A full description of FA groups is provided in Table 1.

2.4 | Definition of diet groups

As diet is the major factor in influencing milk FA composition (Shingfield et al., 2013), the diets provided to cows on different farms were grouped into categories according to types of forages and concentrates. Categories were defined based on results from previous studies showing pronounced differences in milk FA composition of cows fed fresh forage (notably native pastures) as compared to those fed conserved forages (silage/hay), as well as of cows fed lipid-rich concentrates as compared to those fed conventional concentrates (Elgersma, 2015; Hanus et al., 2018; Kliem & Shingfield, 2016; Shingfield et al., 2013). Diets were further grouped based on results from our research team indicating substantial differences between tropical grasses regarding their impact on the milk FA composition of dairy cows (Lopes et al., 2015). For example, cows grazing *Brachiaria* ssp. and sugarcane appear to produce milk with a less desirable FA profile (e.g. lower proportions of RA and C18:3 n-3, and a higher proportion of C16:0) than cows grazing on pastures of *Panicum*, *Pennisetum* and *Cynodum* ssp., although differences among forage species and cultivars within the same genus also seem to exist (Dias et al., 2019). Finally, there is recent evidence that cows consuming cactus *Opuntia* cladodes produce milk with a singular FA composition, notably the low proportions of C18:0 (Gama et al., 2021).

Based on the above, two major groups were formed according to the type of forage offered to the cows. The first group consisted of conserved forages (corn silage and hay) and sugarcane, whereas the other group included tropical pastures. Sugarcane was included in the group of conserved forages due to its low lipid content (<1% DM). The second group was further divided into three subcategories: (1) Pastures of *Brachiaria* ssp. (mostly represented by *B. brizanta* cv. Marandu); (2) Unconventional pastures which included Pangola grass (*Digitaria decumbens*), native pastures (different forage species found in the semi-arid region of Brazil), Buffel grass (*Cenchrus ciliares* L.) and cactus *Opuntia*; and (3) Other common tropical pastures, which included those from the genera *Panicum*, *Pennisetum* and *Cynodon*. For concentrates, three major groups were formed: conventional concentrates, lipid-rich concentrates, and no concentrate feeding. The lipid-rich concentrate group was further divided into two subcategories, one consisting of soybean seed (toasted or extruded) *plus* rumen-protected fat (Megalac or LacFat100), and the other consisting of cottonseed (whole or cake). The categories and subcategories of forages and concentrate feeds used for defining the diet groups are shown in Table S1.

In cases where both corn silage and fresh pasture were offered to cows on the same farm, the diet category was defined based on the season in which milk samples were collected. If samples were taken in spring or summer, the diet was included in the pasture group, whereas the diet was assigned to the corn silage group if samples were taken in autumn or winter.

2.5 | Construction of the targeted sequencing

DNA was extracted from blood samples using a proteinase K/salting out protocol (Miller et al., 1988). A target sequencing strategy was developed with primers designed to cover five regions in the *SCD1* gene. Regions were selected to cover the better-characterized polymorphism in this gene, *SCD1* A293V, and to prospect regulatory polymorphisms. Genotyping by sequencing was done by adapting the Illumina 16S Sample Preparation Guide protocol (Illumina, Inc) using a protocol described elsewhere (Pimentel et al., 2018).

Primer sequences are presented in Table S2 and a detailed description of the target sequencing strategy developed is available under request. PCR reactions were performed using 100 ng of genomic DNA, 1 μM of each primer (5 pmol/μL), 10% DMSO, 20 mM dNTP, 50 mM $MgCl_2$, 2.5 U Taq DNA polymerase (Ludwig Biotecnologia), and ultra-pure water *qsp*. for a 25 μL final volume. PCR amplifications were evaluated using GelRed® stained, 1.5% agarose gels. The target sequencing protocol was composed of the following steps: (1) PCR products from each individual were pooled and purified using the Agencourt AMPure XP purification system (Beckman Coulter); (2) purified pools were quantified in Qubit (Thermo Fisher Scientific Inc) using the dsDNA HS Assay Kit (Invitrogen); (3) the addition of sample identification indexes Nextera XT Index Kit v2 Set A, B, and C (Illumina, Inc) was followed by an additional purification using magnetic beads and a quantification in Qubit (Thermo Fisher Scientific Inc) using the dsDNA HS Assay Kit (Invitrogen); (4) Sequencing was conducted in a MiSeq Sequencer using the MiSeq Reagent Nano Kit v2 in 300 cycles (Illumina, Inc).

TABLE 1 Mean, coefficient of variation and minimum and maximum values for the proportions (g/100 g of total FA) of individual and selected FA groups, n-6:n-3 FA ratio and desaturase (SCD1) indices in milk fat of Gir cows (n = 312).

Fatty acid (FA)	Mean	CV[1]	Median	Min[2]	Max[3]
C4:0	3.542	18.0	3.536	1.915	5.613
C5:0	0.029	45.9	0.026	0.008	0.081
C6:0	2.104	17.2	2.091	0.820	3.307
C7:0	0.021	48.0	0.019	0.005	0.077
C8:0	1.317	18.9	1.320	0.478	2.118
C9:0	0.035	62.7	0.032	0.005	0.315
C10:0	2.440	26.0	2.427	0.741	4.364
cis-9 C10:1	0.257	30.2	0.253	0.091	0.513
C11:0	0.075	45.3	0.071	0.014	0.248
C12:0	2.767	26.7	2.716	0.989	5.040
cis-9 C12:1[5]	0.164	36.3	0.156	0.051	0.457
iso C14:0	0.115	42.7	0.109	0.035	0.320
C14:0	9.139	17.8	9.156	3.691	15.553
iso C15:0	0.233	38.0	0.234	0.088	0.623
anteiso C15:0	0.428	30.0	0.439	0.181	0.885
cis-9 C14:1	1.038	26.7	1.020	0.298	2.388
C15:0	0.962	27.1	0.987	0.414	1.992
iso C16:0	0.265	32.4	0.266	0.082	0.520
C16:0	28.592	14.3	27.740	20.094	42.793
iso C17:0[6]	0.367	24.3	0.350	0.044	0.739
trans-12 C16:1	0.175	23.7	0.166	0.032	0.322
anteiso C17:0[7]	1.917	21.3	1.874	1.129	3.222
C17:0	0.609	27.1	0.579	0.179	1.214
iso C18:0	0.057	44.3	0.051	0.010	0.139
cis-9 C17:1	0.255	39.2	0.227	0.090	0.617
C18:0	9.788	21.0	9.493	4.461	16.626
trans-4 C18:1	0.032	40.2	0.030	0.012	0.084
trans-5 C18:1	0.021	40.8	0.020	0.007	0.050
trans-6 + 7 + 8 C18:1	0.227	38.1	0.208	0.016	0.535
trans-9 C18:1	0.257	36.6	0.225	0.112	0.594
trans-10 C18:1	0.305	35.7	0.270	0.144	0.691
trans-11 C18:1 (VA)	1.465	41.8	1.352	0.447	3.901
trans-12 C18.1	0.348	38.6	0.314	0.134	0.776
trans-13 + 14 C18:1	0.353	44.2	0.332	0.092	0.869
cis-9 C18:1	21.251	22.2	20.530	11.465	33.332
cis-11 C18:1	0.871	25.9	0.834	0.451	1.683
cis-12 C18:1	0.349	52.4	0.292	0.106	1.083
cis-13 C18:1	0.127	42.8	0.115	0.016	0.348
trans-16 C18:1	0.328	33.9	0.322	0.039	0.974
C19:0[8]	0.105	30.9	0.100	0.023	0.246
trans-9, *trans*-12 C18:2	0.020	45.4	0.018	0.004	0.052
cis-9, *trans*-12 C18:2	0.058	45.2	0.056	0.020	0.140
trans-9, *cis*-12 C18:2	0.035	53.1	0.028	0.009	0.094

TABLE 1 (Continued)

Fatty acid (FA)	Mean	CV[1]	Median	Min[2]	Max[3]
C18:2 n-6	1.889	46.5	1.532	0.736	4.223
C20:0	0.142	32.3	0.133	0.065	0.315
C18:3 n-6	0.019	43.3	0.017	0.006	0.055
C18.3 n-3	0.366	33.0	0.363	0.144	0.667
cis-9, *trans*-11 CLA (RA)[9]	0.770	40.0	0.721	0.281	2.522
trans-9, *cis*-11 CLA	0.030	30.3	0.029	0.012	0.062
trans-10, *cis*-12 CLA	0.011	52.5	0.010	0.003	0.042
C21:0	0.028	35.4	0.027	0.005	0.056
C20:2 n-6	0.024	61.4	0.021	0.008	0.213
C22:0	0.056	47.4	0.049	0.012	0.176
C20:3 n-6	0.048	52.0	0.045	0.003	0.353
C20:4 n-6	0.117	27.8	0.112	0.013	0.250
C23:0	0.029	54.5	0.025	0.005	0.112
C20:5 n-3 (EPA)	0.024	39.6	0.022	0.009	0.062
C24:0	0.037	55.8	0.035	0.006	0.141
C22:5 n-3 (DPA)	0.055	34.8	0.052	0.020	0.139
SCD1 indices[10]					
SCD_{14}	0.102	21.2	0.102	0.035	0.168
SCD_{16}	0.064	23.4	0.061	0.035	0.114
SCD_{18}	0.681	8.3	0.679	0.543	0.826
SCD_{CLA}	0.348	12.6	0.347	0.203	0.462
SCD_{total}	0.337	18.6	0.331	0.195	0.496
FA groups					
n-3 FA[11]	0.445	28.7	0.445	0.196	0.798
n-6 FA[12]	2.097	42.8	1.745	0.948	4.427
n-6:n-3 FA ratio[13]	4.785	31.5	4.774	1.741	7.984
Long-chain n-3 FA[14]	0.079	32.1	0.074	0.032	0.188
Milk fat inhibitors[15]	0.347	33.9	0.310	0.174	0.779
Odd-chain linear SFA[16]	1.788	22.4	1.793	1.003	3.178
Branched-chain FA[17]	1.090	31.7	1.098	0.514	2.426
Short-chain SFA[18]	9.403	16.9	9.347	4.392	14.582
Medium-chain SFA[19]	40.498	14.3	39.983	27.318	63.386

Note: [1]CV: coefficient of variation; [2]Min, minimum; [3]Max, maximum; [4]N: Sample size; [5]Co-elutes with C13:0; [6]Co-elutes with *trans*-9 C16:1; [7]Co-elutes with *cis*-9 C16:1; [8]Co-elutes with *cis*-15 C18:1; [9]Contains *trans*-7, *cis*-9 CLA and *trans*-8, *cis*-10 CLA as minor components; [10]Stearoyl-CoA desaturase-1 (SCD1) indices calculated for *cis*-9 14:1/14:0 (SCD_{14}), *cis*-9 16:1/16:0 (SCD_{16}), *cis*-9 18:1/18:0 (SCD_{18}), and *cis*-9, *trans*-11 CLA/*trans*-11 18:1 (SCD_{CLA}) pairs (Kelsey et al., 2003); [11]Sum of all n-3 FA; [12]Sum of all n-6 FA; [13]Ratio of n-6 to n-3 FA; [14]EPA plus DPA; [15]Sum of *trans*-10 C18:1, *trans*-10, *cis*-12 CLA, and *trans*-9, *cis*-12 CLA; [16]Sum of odd-chain linear saturated FA, except C13:0 as it co-eluted with *cis*-9 C12:1; [17]Sum of branched-chain FA, except *iso* C17:0 and *anteiso* C17:0 as they co-eluted with *trans*-9 C16:1 and *cis*-9 C16:1, respectively; [18]Sum of short-chain saturated FA (C4:0 to C10:0); [19]Sum of medium-chain saturated FA (C12:0 to C16:0).

2.6 | Bioinformatic analyses

The fastq files were trimmed and filtered using the Trimmomatic software (Bolger et al., 2014). Bases with Phred scores lower than 20 and/or read lengths shorter than 20 bp were removed. Filtered reads were aligned against a multi-FASTA file containing the amplicon of each region according to the reference ARS-UCD 1.2 (Btau 5.0.1Y), using the software BWA-MEM (Li, 2013). After mapping, the reads were converted from .sam to .bam format, sorted, and indexed using the SAMtools software (Li et al., 2009). We add a group for each individual using tools AddOrReplaceReadGroups from the Picard software (https://broadinstitute.github.io/picard). Reads were indexed and merged with the reads of all individuals, sorted and indexed using the SAMtools software (Li, 2011). SNVs

and INDELs were identified using mpileup and bcftools tools from the SAMtools software, with –d 5,000,000 (Li, 2011). The novel genetic variants identified in the current study were deposited at the European Variant Archive under the project PRJEB53443.

For each variant, allelic and genotypic frequencies and the Hardy–Weinberg Equilibrium (HWE) test were performed using the R (version 4.2.0) package Hardy–Weinberg (https://cran.r-project.org/web/packages/HardyWeinberg/index.html). Variants with MAF >= 0.05 and with a *p*-value >0.05 for HWE test were used in the construction of the haplotypes and in the association studies.

2.7 | Prediction of the functional impacts of the genetic variants

Exonic variants were evaluated in terms of their impacts on the protein sequence and protein domains. Mutation impact on protein sequences was predicted using the Sift (Sim et al., 2012), PolyPhen-2 (Adzhubei et al., 2010), Mutation Taster 2 (Schwarz et al., 2014), and PROVEAN (Choi & Chan, 2015) software. Protein domains, post-synthesis modifications, motifs and active sites were predicted using the neXtProt (Zahn-Zabal et al., 2020), PhosphoSitePlus (Hornbeck et al., 2015), UniProt (UniProt Consortium) and InterPro (Blum et al., 2021). For each program, the default threshold score was adopted and wild-type protein structures available at the Protein Data Bank were used. All the other genetic variants (upstream, exonic and downstream) were included in subsequent bioinformatic analyses. Prediction of functional repercussions of each genetic variant included detection of the creation/abolition of splice sites and branching sites using the Human Splicing Finder 3.0 (HSF); miRNA recognition sites using the TargetScanHuman (Agarwal et al., 2015) and MiRanda software (Betel et al., 2008), filtered with a maximum value of −10 for the free energy connection; and, mutation impact on DNA/RNA sequences using the Mutation Taster 2 (Schwarz et al., 2014). For SNVs in the upstream regions, their effects on promoters were sought using Promoter 2.0 Prediction (Knudsen, 1999). The effects of SNVs on transcription factor binding sites were predicted using GeneQuest version 17.3 (DNASTAR). The effect of SNVs in 3'UTR regions of Poly-A signal was predicted using DNA Functional Site Miner: Poly(A) Signal Miner.

2.8 | Haplotype estimation

The variants maintained in the data set after filtering by the above-mentioned selection criteria were used to estimate the haplotypes in the *SCD1* gene. The R package haplo.stats (Sinnwell & Schaid, 2020) was used to compute the maximum likelihood estimate of haplotype probabilities of each pair of haplotypes composed by the SNPs on *SCD1* gene. The selection of the most probable pair of haplotypes for each individual is based on a progressive insertion algorithm that progressively inserts batches of loci into haplotypes of growing lengths. Subsequently, the package runs the expectation–maximization (EM) algorithms and performs the trimming of pairs of haplotypes per subject when the posterior probability of the pair is below a specified threshold, here defined as 1e-4. The insertion, EM and trimming steps continue until all loci are inserted into the haplotype.

2.9 | Statistical analysis and association studies

Although about 70 individual FA were reported in our GC analysis (Table 1), not all of them were included in the association study. A total of 12 individual FA (C4:0, C12:0, C14:0, C15:0, C16:0, C17:0, C18:0, *cis*-9 C18:1, VA, C18:2 n-6, C18:3 n-3 and RA), 8 FA groups (n-3 FA, n-6 FA, long-chain n-3 FA, milk fat inhibitors, odd-chain linear SFA, branched-chain SFA, short-chain SFA and medium-chain SFA), the n-6:n-3 FA ratio, and four DI (SCD_{14}, SCD_{16}, SCD_{18} and SCD_{CLA}) were selected for the association study due to their nutritional and physiological relevance (Gómez-Cortés et al., 2018; Koch & Lascano, 2018). However, possible biological effects of other milk FA cannot be ruled out.

The genetic association study was performed for each trait in two scenarios: additive (0 vs. 1 vs. 2), where the values correspond to the number of copies of the less frequent allele in the genotype and using the haplotypic classes defined using the *SCD1* alleles maintained after QC. In both scenarios, a linear mixed-effects model was fitted using the lmer4 package in R software v. 4.0.2 (R Core Team, 2021), where the estimates were chosen to optimize the restricted maximum likelihood (REML). The following model was assumed in all the tested scenarios:

$$y_{ijklm} = m + Farm_i + Diet_j + Age_k + Marker_l + DIM_m + S_n + Farm_i \times Diet_j + e_{ijklm}$$

where; y_{ijklm} is the phenotypic observation of the *n*th cow, m is the overall mean, $Farm_i$ is the fixed effect of *i*th Farm (10 levels); $Diet_j$ is the fixed effect of the *j*th diet classification (7 levels); Age_k is the fixed effect of the *k*th cow age, in months; and $Marker_l$ is the allele substitution effect of the *l*th genotypic/haplotypic class; DIM_m is the fixed effect of the *m*th cow days in milk (ranging from 12 to 214 *DIM*, with a mean ± SD of 66.08 ± 28.43); S_n is the random effect

of the *m*th Sire of the cows used in the study (73 levels); $Farm_i \times Diet_j$ is the nested effect between the *i*th Farm and the *j*th diet classification; and e_{ijklm} is the residual error. The proportions of individual FA and FA groups (expressed as g/100 g of total FA) were converted using a squared-root transformation in order to better fit the residuals in a normal distribution. The significantly associated markers were identified using a 5% false discovery rate (FDR).

3 | RESULTS

3.1 | Individual variation in milk FA composition

Phenotypic variation in individual FA, FA groups, n-6:n-3 FA ratio, and DI in milk fat of Gir cows are presented in Table 1. Between-cow variation in the proportion (g/100 g of total FA) of individual FA was the lowest for C16:0 (CV = 14.3%) and the highest for C9:0 (CV = 62.7%). Regarding the DI, the lowest variation was observed for SCD_{18} (CV = 8.3%) followed by SCD_{CLA} (CV = 12.6%), whereas the highest variations were observed for SCD_{14} and SCD_{16} (CV = 21.2 and 23.4%, respectively). The variation in total desaturation index (SCD_{total}) was slightly greater than the average of individual desaturase pairs (CV = 18.6 vs. 16.4%, respectively). For FA groups, medium-chain and short-chain SFA had the lowest variations (CV = 14.3 and 16.9%, respectively), whereas the highest variation was observed for total n-6 FA (CV = 42.8%). The variation observed for the n-6:n-3 FA ratio (CV = 31.5%) was in line that observed for total n-6 FA and total n-3 FA (42.8 and 28.7%, respectively).

3.2 | Development of a targeted sequencing methodology in the search for breed-specific variants in SCD1

A total of 287 Gir cows were successfully genotyped using the *targeted sequencing* tool developed in this study. Five *SCD1* regions were amplified and sequenced. Primers were chosen to cover the *SCD1* polymorphism A293V, in exon 5 and parts of the upstream, *SCD1* Intron 1, and 3′UTR regions of the gene. In 1086 bp sequenced, a total of 14 genetic variants were detected, six of which reported for the first time in the present study (Table 2). Interestingly, one variant (chr26:21272274) was predicted to result in an amino acid substitution (N244D) in the SCD1 protein, as described in more detail below.

Average in-extension coverage and average in-depth coverage were estimated in comparison to the *Bos taurus* reference genome. The criterion adopted in the variant calling was the presence, in a specific nucleotide position, of a different allele from that present in the *Bos taurus* reference genome (ARS-UCD 1.2), in at least one individual.

TABLE 2 Genetic variants identified in the bovine *SCD1* gene.

Gene region	Genomic position (ARS-UCD1.2)	Rs	Number of alleles	Number of allele 1	Number of allele 2	MAF	Location /functional classification
SCD1 Upstream	Chr26:21263567	New	542	541	1	0.002	*upstream*
	Chr26:21263591	New	542	542	0	0	*upstream*
	Chr26:21263602	New	542	541	1	0.002	*upstream*
SCD1 Intron 1	Chr26:21264191	rs475591968	562	562	0	0	intron
SCD1 Exon 5	Chr26:21272246	rs41255691	556	555	1	0.002	*synonymous*
	Chr26:21272274	New	556	555	1	0.002	*missense* (*N244D*)
	Chr26:21272422	rs41255693[a]	562	562	0	0	*missense* (*A293V*)
	Chr26:21272423	rs208932125[a]	562	417	145	0.258	*splicing region*
SCD1 3′UTR	Chr26:21277069	New	474	335	139	0.293	3′UTR
	Chr26:21277094	rs523411937	474	428	46	0.097	3′UTR
	Chr26:21277134	rs452574907	476	475	1	0.002	3′UTR
	Chr26:21277160	New	476	475	1	0.002	3′UTR
	Chr26:21277188	rs437125709	476	475	1	0.002	3′UTR
	Chr26:21277196	rs211383702	474	473	1	0.002	3′UTR

[a]The most studied *SCD1* polymorphism (Ala293Val) is caused by a non-synonym C > T substitution at the second base of the 293th codon, whereas the second SNP (rs208932125) found in the 293th codon does not affect the protein sequence, but the splicing at the 3′ region of *SCD1* exon 5; MAF: minor allele frequency.

3.3 | Allelic distribution of SCD1 genetic variants

In total, 11 genetic variants were identified within the genetic coordinates of the *SCD1* gene, and three were located in the upstream region of *SCD1*. The MAF for all the variants is shown in Table 2. Two variants were maintained after the Hardy–Weinberg Equilibrium test and the filtering by MAF > 0.05: rs208932125 (MAF = 0.258) and rs523411937 (MAF = 0.097). These two variants were further included in the association study. A high linkage disequilibrium was observed between these two variants as evidenced by D′ and r^2 values of 0.97 and 0.54, respectively. Therefore, association was tested using these SNPs separately and in haplotype. Additionally, the new variant mapped on the 3′-UTR region of *SCD1* (Chr26:21277069) identified here passed the MAF threshold (0.293). However, this variant was not under HWE and it was not included in the haplotype estimations due to the potential impact over the linkage disequilibrium estimations. Nevertheless, the effect of the genotypes of Chr26:21277069 over traits evaluated here were tested individually.

3.4 | Bioinformatic analysis of the main genetic variants

The evidence for functional impacts of the two genetic variants analysed in the association study (rs208932125 and rs523411937) and the new missense variant (*SCD1* N244D) are discussed below.

The rs208932125 is a silent mutation in the splicing region of exon 5. The most studied bovine *SCD1* variant, A293V, is a non-synonymous SNP (rs41255693) found in exon 5 (chr26:21272422) in which a C > T substitution occurs at the second base of the 293 codon, resulting in alanine being replaced with a valine in the amino acid sequence. A second SNP (rs208932125) has also been reported in exon 5 (chr26:21272423), in which a synonym G > A substitution occurs at the third base of the 293th codon (Figure 1). Regardless of the third position, all codons having G in the first position and C in the second encode an alanine. Therefore, all individuals in this sample were homozygous for the alanine allele. As rs208932125 was associated with the fatty acid profile in the present study, we sought evidence for additional functional effects of this SNP. Using Human Splicing Finder, it was predicted that rs208932125 abolishes exonic splicing enhancers (ESE-EIE, ESE-ASFB and ESE-SRp55), and also affects an exonic splicing silencer (ESS-Sironi Motifs, ESS-Fas and ESS-hnRNPA1). This finding offers an additional explanation for the impact of *SCD1* A293V polymorphism on mammary lipid metabolism and milk FA composition, other than the amino acid substitution.

3.5 | The functional impact of the new protein variant detected in this study

The new variant mapped at Chr26:21272274 is a missense variant (cDNA:1123/c.730A > G p.N244D) and is located in *SCD1* exon 5, having effects at RNA and protein levels. Using I-Mutant, this variant is predicted to destabilize the protein (at 37.5°C and pH 7, the amino acid substitution induces a −0.99 Free Energy change). This amino acid substitution occurs in an alpha-helix, corresponding to a transmembrane region spanning residues 242–263. Both mCSM and DUET predicted that N244D destabilizes the protein 3D structure. It was predicted that N244D

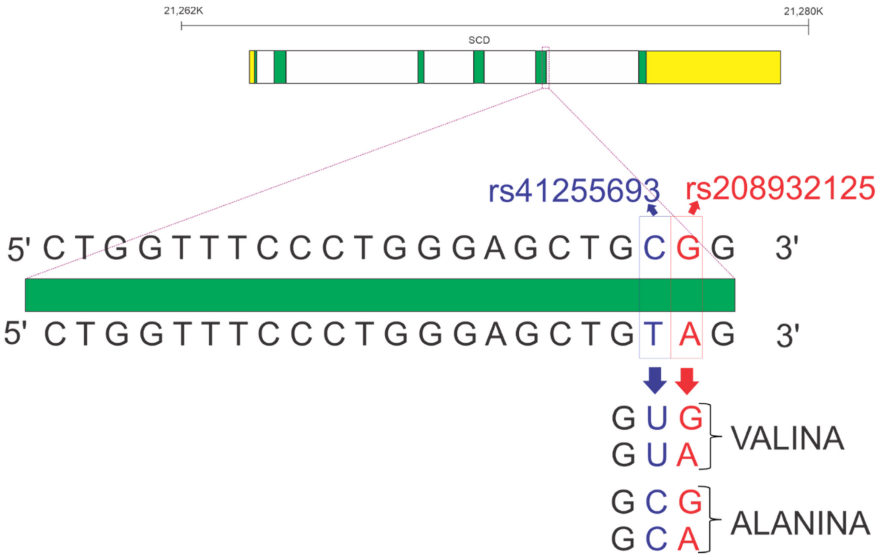

FIGURE 1 Schematic representation of the two SNPs found in exon 5 of the *SCD1* gene. The first variant (rs41255693), commonly referred to as A293V, is a non-synonym SNP caused by a C > T substitution at the second base of the 293th codon, in which an alanine is replaced with a valine. The second SNP (rs208932125) results from a G > A substitution at the third base of the codon, with no change in the amino acid sequence. [Colour figure can be viewed at wileyonlinelibrary.com]

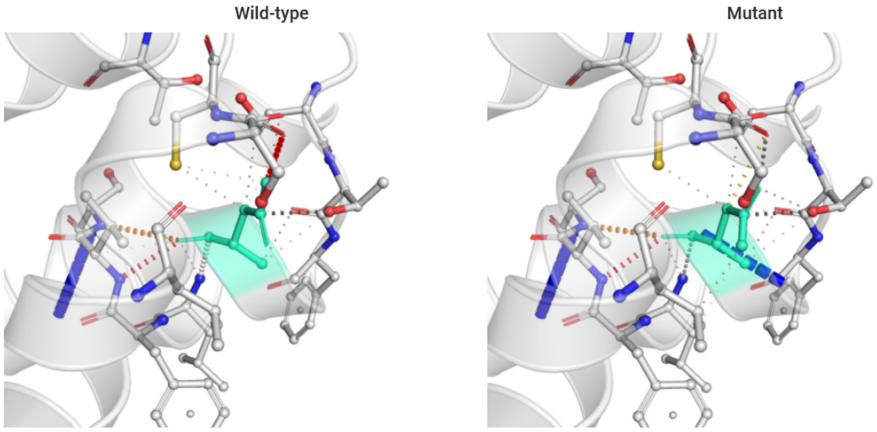

FIGURE 2 SCD1 N244D Molecular dynamics predicted using DynaMut using the 4YMK PDB model. N244D significantly affects the molecular interaction established within the protein. In blue, it shows the halogen bonds; in yellow, it shows the ionic interactions; in red, it shows hydrogen bonds; in orange, it shows weak hydrogen bonds. [Colour figure can be viewed at wileyonlinelibrary.com]

significantly affects the molecular interactions established within the protein (using mCSM-PPI2 and DynaMut programs) (Figure 2). Comparing the molecular dynamics of normal and mutated protein, it is possible to observe the gain of a halogen bond and of an ionic interaction, and the loss of a hydrogen bond. However, the three-dimensional model of bovine SCD1 has not been included in PDB and, therefore, the molecular dynamics was inferred using as models the human and mouse SCD1 structures. As both models present non-conserved residues close to the N244, any conclusions regarding molecular interactions should be made with care. This variant was identified in heterozygosis in only one cow; consequently, it was not included in the association studies because these variants need further validation using a larger sample size. No evidence of possible functional impacts was detected for the other variants detected in the present study (Table 2).

3.6 | Association test for SCD1 genetic variants

In the association study, no markers or haplotypes were significantly associated with milk FA composition after FDR adjustment. However, significant associations based on raw *p*-values (i.e. raw *p*-value <0.05) were identified for some individual FA (C12:0, C14:0, C15:0, VA, RA and C18:3 n-3), three FA groups (short-chain SFA, odd-chain linear SFA, and n-3 FA), and SCD_{CLA} (Table 3). This significance threshold was defined due to the limited sample size, which may have precluded the detection of significant associations after adjustment for FDR. However, the actual relevance of any association found in the present study should be confirmed in future association studies targeting at the detected *SCD1* variants. Specifically, rs208932125 was associated with C12:0, RA and short-chain SFA, with cows carrying the AA genotype showing lower proportions of C12:0 and short-chain SFA, but higher RA than cows with the GG genotype. On the other hand, rs523411937 was associated with C15:0 and odd-chain linear SFA, with cows carrying the CT genotype having a higher proportion of these FA in milk fat than those with the TT genotype. Regarding the *SCD1* 3′UTR Chr26:21277069 variant for which only two genotypes (GG and AG) were detected in the population, AG cows had higher proportions of VA and RA in milk fat, but lower proportions of C18:3 n-3 and n-3 FA than GG cows. Associations were also observed between three haplotypes and some individual FA, FA groups and SCD1 indices. Compared to AT_AT, the most frequent diplotype in the population ($n=153$), we found that AT_GT cows (the second most frequent diplotype, $n=84$) had lower SCD_{CLA} index. Some associations were also observed for diplotypes identified in only a few cows. Compared to the AT_AT diplotype, GT_GC cows had higher proportions of C12:0, C14:0, and short-chain FA in milk fat, whereas GT_GT cows had lower RA. Because only one cow was identified with the AT_AC diplotype, no statistical comparison was made between this diplotype and the reference group (AT_AT). Milk FA proportions and SCD1 indices for all genotypes and haplotypes that were in EHW, as well as for the 3′UTR Chr26:21277069 variant, are shown in Table S3, whereas the respective *p*-values and FDR of the association tests are presented in Table S4.

4 | DISCUSSION

To the best of our knowledge, this is the first study to estimate the phenotypic variation in milk FA composition of Gir cows, a Zebu breed that makes a significant contribution to milk production in Brazil, India and other regions of the world (Santana et al., 2014). We further investigated the association of polymorphisms in the *SCD1* gene, which encodes a key enzyme in mammary lipid metabolism by introducing a *cis* double bond at the Δ9-position in several FA, with milk FA composition. This trait is of great interest on a nutritional point of view as dairy fat contains

TABLE 3 Associations of SNP and haplotypes of *SCD1* with proportions (mean ± CV) of individual fatty acids (FA) and FA groups (g/100 g of total FA), desaturase indices (DI) and n-6:n-3 FA ratio in milk fat of Gir cows[a].

Variant	FA	Genotypic means (±CV)	*p*-value (FDR)[b]
rs208932125	C12:0	AA: 2.749 (±26.85); GA: 2.801 (±27.12); GG: 3.004 (±27.13)	GA: 0.338 (0.884); **GG: 0.029 (0.884)**
rs208932125	*cis*-9, *trans*-11 CLA (RA)	AA: 0.771 (±37.014); GA: 0.766 (±41.781); GG: 0.648 (±39.543)	GA: 0.656 (0.935); **GG: 0.008 (0.884)**
rs208932125	Short-chain SFA	AA: 9.447 (±16.775); GA: 9.348 (±16.521); GG: 9.859 (±20.57)	GA: 0.578 (0.924); **GG: 0.011 (0.884)**
rs523411937	C15:0	TT: 0.943 (±28.394); CT: 1.059 (±22.981); CC: 1.074 (±35.823)	**CT: 0.023 (0.884)**; CC: 0.688 (0.945)
rs523411937	Odd-chain linear SFA	TT: 1.752 (±22.553); CT: 1.939 (±20.446); CC: 1.994 (±17.524)	**CT: 0.041 (0.884)**; CC: 0.597 (0.924)
SCD1 3'UTR Chr26:21277069	*trans*-11 C18:1 (VA)	GG: 1.296 (40.108); AG: 1.455 (40.993)	GG: **0.018 (0.783)**
SCD1 3'UTR Chr26:21277069	C18.3 n-3	GG: 0.357 (31.401); AG: 0.354 (33.76)	**GG: 0.036 (0.783)**
SCD1 3'UTR Chr26:21277069	*cis*-9, *trans*-11 CLA (RA)	GG: 0.700 (35.431); AG: 0.761 (39.414)	**GG: 0.026 (0.783)**
SCD1 3'UTR Chr26:21277069	n-3 FA	GG: 0.434 (27.152); AG: 0.431 (29.025)	**GG: 0.028 (0.783)**
Haplotypes *SCD1*	C12:0	AT_AT: 2.751 (26.555); AT_GT: 2.793 (27.659); GC_AT: 2.824 (26.092); GT_GC: 3.241 (27.962); GT_GT: 2.747 (26.818); GC_GC: 2.935 (34.307); AT_AC: 2.418 (−)	AT_GT: 0.54 (0.906); GC_AT: 0.3 (0.884); **GT_GC: 0.041 (0.884)**; GT_GT: 0.366 (0.884); GC_GC: 0.458 (0.894)
Haplotypes *SCD1*	C14:0	AT_AT: 9.125 (17.96); AT_GT: 9.145 (18.046); GC_AT: 9.282 (15.781); GT_GC: 10.28 (22.05); GT_GT: 8.817 (19.92); GC_GC: 9.634 (23.377); AT_AC: 9.017 (−)	AT_GT: 0.414 (0.884); GC_AT: 0.691 (0.945); **GT_GC: 0.042 (0.884)**; GT_GT: 0.926 (0.995); GC_GC: 0.33 (0.884)
Haplotypes *SCD1*	*cis*-9, *trans*-11 CLA (RA)	AT_AT: 0.774 (36.474); AT_GT: 0.781 (43.22); GC_AT: 0.727 (37.121); GT_GC: 0.544 (37.018); GT_GT: 0.742 (37.117); GC_GC: 0.726 (55.233); AT_AC: 0.482 (−)	AT_GT: 0.556 (0.906); GC_AT: 0.899 (0.995); GT_GC: 0.088 (0.884); **GT_GT: 0.041 (0.884)**; GC_GC: 0.388 (0.884)
Haplotypes *SCD1*	SCD_{CLA}	AT_AT: 0.351 (12.815); AT_GT: 0.336 (12.381); GC_AT: 0.361 (12.57); GT_GC: 0.336 (7.178); GT_GT: 0.308 (5.069); GC_GC: 0.353 (11.755); AT_AC: 0.357 (−)	**AT_GT: 0.013 (0.884)**; GC_AT: 0.222 (0.884); GT_GC: 0.131 (0.884); GT_GT: 0.104 (0.884); GC_GC: 0.62 (0.931)
Haplotypes *SCD1*	Short-chain SFA	AT_AT: 9.442 (16.559); AT_GT: 9.46 (16.796); GC_AT: 9.055 (15.487); GT_GC: 10.072 (27.017); GT_GT: 9.551 (15.586); GC_GC: 9.993 (16.397); AT_AC: 11.386 (−)	AT_GT: 0.52 (0.906); GC_AT: 0.857 (0.995); **GT_GC: 0.03 (0.884)**; GT_GT: 0.149 (0.884); GC_GC: 0.548 (0.906)

Note: Bold values represent *p*-value < 0.05.

[a]Only SNP and haplotypes significantly associated with milk FA proportions and SCD1 indices are reported. Complete data are presented in Tables S3 and S4.

[b]Raw *p*-values and false discovery rates (FDR, between parentheses) adjustment were estimated using one of the allelic or haplotypic categories as reference class. Therefore, for each *loci*, the first allelic or haplotypic class do not have an assigned *p*-value (FDR).

a number of health-promoting FA as well as potentially harmful FA (e.g. C16:0) which contents can be affected to a significant extent by diet and genetics (Shingfield et al., 2013). The between-cow variation in milk FA composition was estimated from 312 cows distributed in 10 herds located in five states of Brazil and fed on different

diets. Therefore, the phenotypic variations reported in our study are likely to be representative of the population. Overall, the variations observed for the proportions of individual FA, FA groups, and DI fell within the range reported in the literature for taurine breeds (Garnsworthy et al., 2010; Mele et al., 2007; Pegolo et al., 2016; Schennink et al., 2008; Stoop et al., 2008). Interestingly, a lower between-cow variation (CV = 8.3%) was observed for the SCD_{18} index when compared to the other DI reported in our study. The second largest variation (CV = 12.6%) was observed for the SCD_{CLA} index, whereas the largest variations were observed for the SCD_{14} and SCD_{16} indices. Similar findings were reported in studies where a large number of cows were used to estimate heritability of DI in taurine breeds (Garnsworthy et al., 2010; Mele et al., 2007; Pegolo et al., 2016; Schennink et al., 2008), as well as in studies where DI were calculated from cows fed on the same diet (Kelsey et al., 2003). As diet is the major factor influencing milk FA composition (Kliem & Shingfield, 2016), which is supported by the significant effect of diet groups in our study (data not shown), the study of Kelsey et al. (2003) suggests that genetic variation in SCD_{18} index may also be lower than for the other DI. The SCD1 indices are calculated as the ratio of products to substrates of SCD1 and are commonly used as a proxy for SCD1 enzyme activity in the mammary gland. The *cis*-9 C14:1/C14:0 ratio has been suggested as the best indicator for the SCD1 activity because C14:0 in milk fat derives almost exclusively from de novo synthesis in the mammary gland, and therefore, almost all the *cis*-9 C14:1 is likely to be synthesized by SCD1 (Bernard et al., 2013), whereas the other precursors and products of SCD1 are derived partially, and to varying extents, from the circulatory system (Garnsworthy et al., 2010). Accordingly, studies with goats showed that the milk *cis*-9 C14:1/C14:0 ratio gave the best estimation for the response of mammary SCD1 activity (Bernard et al., 2013), although no positive correlation was observed between the C14 index and *SCD1* gene expression in mammary gland from dairy cows (Rezamand et al., 2014), which may indicate the existence of post-translational mechanisms in the regulation of SCD1 activity. The smaller variation in the SCD_{18} index as compared to the other DI observed in the present study and in previous reports with taurine breeds may be explained, at least partly, by the pivotal role of *cis*-9 C18:1 (oleic acid) in maintaining milk fat fluidity (Jensen, 2002). Indeed, it is well established that C18:0 is the preferred substrate of SCD1 (Ntambi & Miyazaki, 2004), and that endogenous synthesis of *cis*-9-18:1 via the action of SCD1 on 18:0 in the mammary gland is an important point of regulation in milk TG synthesis and maintenance of milk fat fluidity (Chilliard et al., 2007; Gama et al., 2008). Thus, it could be speculated that SCD1 activity on C18:0 is more tightly regulated than for other SCD1 substrates so that the *cis*-9 C18:1/C18:0 ratio in milk fat is kept within a narrower range, similar to what is observed for milk lactose content as compared to milk protein and fat contents.

One of the FA groups included in our study due to its physiological relevance was the 'milk fat inhibitors' group, which comprises three intermediates of rumen biohydrogenation that have been shown to inhibit the milk fat synthesis in the mammary gland of cows (Koch & Lascano, 2018), or that are strongly associated with milk fat depression (MFD), such as *trans*-10 C18:1 (Bauman et al., 2011). A considerable phenotypic variation (CV = 33.9%) was observed for this group, which was expected given the very different diets fed to cows across the farms and even within the same farm in cases where milk samples were collected in different seasons of the year. However, part of the phenotypic variation observed in the proportion of milk fat inhibitors may also be attributed to cow individuality, as genetic variation in susceptibility to MFD has been shown to exist, indicating that selection for reduced susceptibility to MFD is possible (Calus et al., 2005). This is consistent with evidence that individual cows have different susceptibilities to *trans*-10 shift (Gama et al., 2021), which has also been reported in studies with ewes (Santos-Silva et al., 2016).

In addition to estimating the phenotypic variance in milk FA composition and DI in Gir cows, we investigated the existence of genetic variants in the *SCD1* that could be associated with those traits. As shown in Table 2, 14 variants were identified in different regions of the *SCD1* gene, with six of them being reported for the first time in the present study. Interestingly, one new variant (chr26:21272274) was predicted to result in an amino acid substitution (N244D) in the SCD1 protein. From the 14 detected genetic variants, three were found to be fixed in the population (i.e. all cows were homozygous for a given allele): a new variant on Chr26:21263591, rs475591968, and rs41255693. The latter is the most studied *SCD1* polymorphism in cattle breeds, a non-synonym SNP (C > T substitution) causing the substitution of valine for alanine at position 293 (A293V), which has explained a significant proportion of genetic variation in milk FA composition in taurine breeds (Conte et al., 2010; Kgwatalala et al., 2007; Li et al., 2016; Mele et al., 2007; Schennink et al., 2008). In addition to the A293V polymorphism, a second SNP (rs208932125) has also been detected in the 293th codon of the *SCD1* gene (Figure 1). Interestingly, only the allele C of the first SNP was detected in the present study, so all the cows had an alanine rather than a valine in the amino acid sequence of the SCD1 protein. On the other hand, the two alleles (A and G) were detected for the second SNP (Table 2), and the resulting genotypes were significantly associated (based on raw *p*-values) with C12:0, RA, and

short-chain SFA, with cows carrying the AA genotype showing lower proportions of C12:0 and short-chain SFA, but higher RA than cows with the GG genotype (Table 3). The increased proportion of RA in milk fat from AA cows is consistent with the trend (raw p-value = 0.063) towards higher SCD_{CLA} index observed in these cows as compared to GG cows. However, the lack of effect on the SCD_{14} index, which is considered the best proxy for SCD1 activity in mammary gland, suggests an altered substrate specificity of SCD1 enzyme in AA cows. These results indicate that selecting cows carrying the AA genotype is expected to improve the nutritional quality of milk fat due to the numerous health benefits attributed to RA, although some positive effects have also been reported for short-chain FA and C12:0 (Gómez-Cortés et al., 2018), which concentrations in milk fat are expected to decrease in AA cows. Interestingly, when we sought evidence for possible functional effects of rs208932125 using Human Splicing Finder, it was predicted that this SNP abolishes exonic splicing enhancers (ESE-EIE, ESE-ASFB, and ESE-SRp55), and also affects an exonic splicing silencer (ESS-Sironi Motifs, ESS-Fas, and ESS-hnRNPA1). These findings offer further explanation for the well-described effect of *SCD1* A293V polymorphism on milk FA composition. In addition to rs208932125, associations were observed between rs523411937, a SNP found in the *SCD1* 3′UTR region, and the proportions of C15:0 and odd-chain SFA in milk fat, with cows carrying the CT genotype having higher proportions of these FA in milk fat than those with the TT genotype. Interestingly, the C15:0 has been shown to possess activities and efficacy that parallel associated health benefits in humans (Venn-Watson et al., 2020), leading the authors to propose that C15:0 could be an essential fatty acid. Therefore, an increased proportion of C15:0 in milk fat from cows carrying the CT genotype is desirable from a nutritional standpoint. The possibility of increasing the contents of C15:0 in milk through selective breeding is supported by recent evidence of a heritable subset of the core rumen microbiome (Wallace et al., 2019), since the odd-chain FA found in milk fat are largely derived from cell membranes of rumen bacteria (Vlaeminck et al., 2006). Another new variant detected in the *SCD1* 3′UTR region (Chr26:21277069) was associated with proportions of important milk FA such as *trans*-11 C18:1, *cis*-9, *trans*-11 CLA, C18:3 n-3 and n-3 FA. It should be noted that only two genotypes (GG and AG) were found for this variant in our study, which may reflect the pressure of selection for certain milk fatty acids and deserves further investigation. Associations were also observed for three diplotypes identified in our study (Table 3). As two of them (GT_GC and GT_GT) were found in a small number of cows (6 and 5, respectively), the associations observed with C12:0, C14:0, short-chain SFA and RA should be interpreted with caution despite the nutritional relevance of these FA. The lower SCD_{CLA} index found in milk from cows having the second most frequent diplotype (AT_GT, $n = 84$) when compared to AT_AT cows was not accompanied by a reduction in RA, suggesting that changes in the SCD_{CLA} index may not fully reflect the activity of SCD1 on RA synthesis in the mammary gland of cows.

5 | CONCLUSIONS

To our knowledge, this is the first study to report the phenotypic variation in milk FA composition and SCD1 indices in Gir cows, which was found to be within the range reported for taurine breeds. In particular, a lower variation was observed for the SCD_{18} index as compared to the other DI. This result was very consistent across different studies, suggesting that SCD1 activity on C18:0 may be more tightly regulated than for other SCD1 substrates in mammary gland of dairy cows. A total of 14 variants of the *SCD1* gene were identified, six of which not previously described in the literature. Interestingly, the most studied non-synonym SNP in the *SCD1* gene (A293V), which has been shown to explain a significant proportion of the variation in milk FA composition of taurine breeds, was monomorphic in our study, with all the cows being homozygous for the C allele (that codes for an alanine). However, all the genotypes (AA, AG, and GG) of the second SNP affecting the 293th codon were found in the population. Although this polymorphism does not change the amino acid sequence of the SCD1 protein, we found compelling evidence for functional effects of this variant. Some significant associations (based on raw p-values) with milk FA composition were found for rs208932125 and for two SNP located in the 3′UTR region of the *SCD1* gene (rs523411937 and a new variant on Chr26:21277069), which may be worth exploring in future studies. Interestingly, a new SCD1 protein variant (N244D) was identified, which was predicted to affect the protein function. These findings indicate that *SCD1* variants other than the classic non-synonym A293V polymorphism contribute to variation in milk FA composition of Gir cows, with potential to be used in their breeding programmes, although the magnitude of milk FA changes is of limited extent when compared to that achieved by dietary means. Supportive validation from studies with larger sample sizes is necessary to confirm the putative associations and to better estimate the effects of the novel *SCD1* variants on milk FA composition of Gir cows.

AUTHOR CONTRIBUTIONS

Marco Antonio Sundfeld Gama, Maria Gabriela Campolina Diniz Peixoto, Maria Raquel Santos Carvalho,

Pablo Augusto de Souza Fonseca contributed to *study design*. Maria Gabriela Campolina Diniz Peixoto, Anibal Vercesi Filho, Frank Angelo Tomita Bruneli, Marco Antonio Sundfeld Gama contributed to *sample and data collection*. Carolina Guimarães Ramos Matosinho, Izinara Cruz Rosse, Thalia Zózimo, Maria Raquel Santos Carvalho contributed to *targeted sequencing design, execution and bioinformatic analysis*. Pablo Augusto de Souza Fonseca, Maria Gabriela Campolina Diniz Peixoto, Frank Angelo Tomita Bruneli, Marco Antonio Sundfeld Gama, Maria Raquel Santos Carvalho, Fernando Cesar Ferraz Lopes contributed to d*ata analysis*. Marco Antonio Sundfeld Gama, Maria Raquel Santos Carvalho, Pablo Augusto de Souza Fonseca, Carolina Guimarães Ramos Matosinho, Thalia Zózimo contributed to w*riting of manuscript*.

ACKNOWLEDGEMENTS

We would like to thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for funding this study (project CVZ APQ-02003-15). MRSC was awarded a FAPEMIG grant (APQ-01377-17) and fellowships from the Brazilian National Research Council (312068/2015-8 and 312405/2018-9). CGRM and TZ were awarded fellowships from the Coordination for the Improvement of Higher Education Personnel (CAPES). We also thank the farmers and Glaucyana Gouvêa dos Santos for their contribution to this study.

CONFLICT OF INTEREST STATEMENT

None.

DATA AVAILABILITY STATEMENT

All data used in this paper are listed in the Tables or Supplementary material. Additional data are available under request. The novel genetic variants identified in the current study were deposited at the European Variant Archive under the project PRJEB53443.

ORCID

Pablo Augusto de Souza Fonseca https://orcid.org/0000-0002-6917-7475

REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*, 248–249.

Agarwal, V., Bell, G. W., Nam, J. W., & Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife*, *4*, e05005.

AOAC International. (2012). *Official methods of analysis of AOAC international* (19th ed.). AOAC Official Methods of Analysis.

Arnould, V. M. R., & Soyeurt, H. (2009). Genetic variability of milk fatty acids. *Journal of Applied Genetics*, *50*, 29–39.

Astrup, A. (2014). A changing view on saturated fatty acids and dairy: From enemy to friend. *American Journal of Clinical Nutrition*, *100*, 1407–1408.

Astrup, A., Magkos, F., Bier, D. M., Brenna, J. T., de Oliveira Otto, M. C., Hill, J. O., King, J. C., Mente, A., Ordovas, J. M., Volek, J. S., et al. (2020). Saturated fats and health: A reassessment and proposal for food-based recommendations: JACC State-of-the-Art Review. *Journal of the American College of Cardiology*, *76*(7), 844–857.

Baldin, M., Gama, M. A. S., Dresch, R., Harvatine, K. J., & Oliveira, D. E. (2013). A rumen unprotected conjugated linoleic acid supplement inhibits milk fat synthesis and improves energy balance in lactating goats. *Journal of Animal Science*, *91*, 3305–3314.

Bastin, C., Berry, D. P., Soyeurt, H., & Gengler, N. (2012). Genetic correlations of days open with production traits and contents in milk of major fatty acids predicted by mid-infrared spectrometry. *Journal of Dairy Science*, *95*, 6113–6121.

Bauman, D. E., Harvatine, K. J., & Lock, A. L. (2011). Nutrigenomics, rumen-derived bioactive fatty acids, and the regulation of milk fat synthesis. *Annual Review of Nutrition*, *31*, 299–319.

Bernard, L., Leroux, C., & Chilliard, Y. (2013). Expression and nutritional regulation of stearoyl-CoA desaturase genes in the ruminant mammary gland: Relationship with milk fatty acid composition. In *Stearoyl-CoA desaturase genes in lipid metabolism*. Springer New York.

Betel, D., Wilson, M., Gabow, A., Marks, D. S., & Sander, C. (2008). The microRNA.Org resource: Targets and expression. *Nucleic Acids Research*, *36*, D149–D153.

Blum, M., Chang, H. Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G. A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D. H., Letunic, I., Marchler-Bauer, A., … Finn, R. D. (2021). The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research*, *49*, D344–D354.

Bobe, G., Minick Bormann, J. A., Lindberg, G. L., Freeman, A. E., & Beitz, D. C. (2008). Short communication: Estimates of genetic variation of milk fatty acids in US Holstein cows. *Journal of Dairy Science*, *91*, 1209–1213.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120.

Bruneli, F. A. T., Carvalho, M. R. S., Lôbo, R. B., Verneque, R. D. S., Zadra, L. E. F., Penna, V. M., Arbex, W. A., Pereira, R. J., Santana Júnior, M. L., & Peixoto, M. (2021). National Breeding Program of Guzerá cattle for Milk: Progeny testing, National Zootechnical Archive and MOET nucleus results. Embrapa Gado de Leite-Documentos (INFOTECA-E).

Calus, M. P. L., Carrick, M. J., Veerkamp, R. F., & Goddard, M. E. (2005). Estimation of genetic parameters for milk fat depression in dairy cattle. *Journal of Dairy Science*, *88*, 1166–1177.

Carvajal, A. M., Huircan, P., Dezamour, J. M., Subiabre, I., Kerr, B., Morales, R., & Ungerfeld, E. M. (2016). Milk fatty acid profile is modulated by DGAT1 and SCD1 genotypes in dairy cattle on pasture and strategic supplementation. *Genetics and Molecular Research*, *15(2)*, 1–12.

Chilliard, Y., Glasser, F., Ferlay, A., Bernard, L., Rouel, J., & Doreau, M. (2007). Diet, rumen biohydrogenation and nutritional quality of cow and goat milk fat. *European Journal of Lipid Science and Technology*, *109*, 828–855.

Choi, Y., & Chan, A. P. (2015). PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, *31*, 2745–2747.

Conte, G., Mele, M., Chessa, S., Castiglioni, B., Serra, A., Pagnacco, G., & Secchiari, P. (2010). Diacylglycerol acyltransferase 1, stearoyl-CoA desaturase 1, and sterol regulatory element binding protein 1 gene polymorphisms and milk fatty acid composition in Italian Brown cattle. *Journal of Dairy Science*, *93*, 753–763.

Corl, B. A., Baumgard, L. H., Dwyer, D. A., Griinari, J. M., Phillips, B. S., & Bauman, D. E. (2001). The role of Δ9-desaturase in the production of cis-9, trans-11 CLA. In. *Journal of Nutritional Biochemistry*, *12*(11), 622–630.

Cruz-Hernandez, C., Kramer, J. K. G., Kennelly, J. J., Glimm, D. R., Sorensen, B. M., Okine, E. K., Goonewardene, L. A., & Weselake, R. J. (2007). Evaluating the conjugated linoleic acid and Trans 18:1 isomers in milk fat of dairy cows fed increasing amounts of sunflower oil and a constant level of fish oil. *Journal of Dairy Science*, *90*, 3786–3801.

Dariush Mozaffarian, M. D. (2015). Dietary and policy priorities for cardiovascular disease, diabetes, and obesity – A comprehensive review. *Circulation*, *133*(2), 187–225.

Dias, K. M., da Gama, M. A. S., Schmitt, D., & Sbrissia, A. F. (2019). Milk fatty acid composition of unsupplemented dairy cows grazing on a tropical pasture. *Revista Brasileira de Zootecnia*, *48*, 1–8.

Elgersma, A. (2015). Grazing increases the unsaturated fatty acid concentration of milk from grass-fed cows: A review of the contributing factors, challenges and future perspectives. *European Journal of Lipid Science and Technology*, *117*, 1345–1369.

Freitas, A., Souza, G., Camargo, G., Peixoto, G., & Mrs, C. T. H. (2013). Characterization of stearoil-CoA desaturase gene in Gir and Guzerábreed. In *Simposio Brasileiro de Melhoramento Animal, X*, Uberaba.

Gama, M. A. S., de Paula, T. A., Véras, A. S. C., Guido, S. I., Borges, C. A. V., Antoniassi, R., Lopes, F. C. F., Neves, M. L. M. W., & Ferreira, M. d. A. (2021). Partially replacing sorghum silage with cactus (*Opuntia stricta*) cladodes in a soybean oil-supplemented diet markedly increases trans-11 18:1, cis-9, trans-11 CLA and 18:2 n-6 contents in cow milk. *Journal of Animal Physiology and Animal Nutrition*, *105*(2), 232–246.

Gama, M. A. S., Garnsworthy, P. C., Griinari, J. M., Leme, P. R., Rodrigues, P. H. M., Souza, L. W. O., & Lanna, D. P. D. (2008). *Diet-induced milk fat depression: Association with changes in milk fatty acid composition and fluidity of milk fat*. Livest Sci.

Gama, M. A. S., Lopes, F. C. F., Vercesi Filho, A. E., & others. (2013). Variação individual nas relações produto: substrato da enzima estearoil-CoA desaturase (SCD) e nos teores dos ácidos rumênico (CLA cis-9 trans-11) e oleico no leite de vacas Gir e Guzerá. *Simpósio Brasileiro de Melhoramento Animal*, *10*, 1–3.

Garnsworthy, P. C., Feng, S., Lock, A. L., & Royal, M. D. (2010). Short communication: Heritability of milk fatty acid composition and stearoyl-CoA desaturase indices in dairy cows. *Journal of Dairy Science*, *93*, 1743–1748.

Gómez-Cortés, P., Juárez, M., & de la Fuente, M. A. (2018). *Milk fatty acids and potential health benefits: An updated vision* (Vol. *81*, pp. 1–9). Trends in Food Science & Technology.

Hanus, O., Samkova, E., Křížova, L., Hasoňova, L., & Kala, R. (2018). Role of fatty acids in milk fat and the influence of selected factors on their variability—A review. *Molecules*, *23*(7), 1636.

Herforth, A., Arimond, M., Álvarez-Sánchez, C., Coates, J., Christianson, K., & Muehlhoff, E. (2019). A global review of food-based dietary guidelines. *Advances in Nutrition*, *10*(4), 590–605.

Hirahatake, K. M., Astrup, A., Hill, J. O., Slavin, J. L., Allison, D. B., & Maki, K. C. (2020). Potential Cardiometabolic health benefits of full-fat dairy: The evidence base. *Advances in Nutrition*, *11*(3), 533–547.

Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., & Skrzypek, E. (2015). PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Research*, *43*(D1), D512–D520.

Jensen, R. G. (2002). The composition of bovine milk lipids: January 1995 to December 2000. *Journal of Dairy Science*, *85*, 295–350.

Kelsey, J. A., Corl, B. A., Collier, R. J., & Bauman, D. E. (2003). The effect of breed, parity, and stage of lactation on conjugated linoleic acid (CLA) in milk fat from dairy cows. *Journal of Dairy Science*, *86*, 2588–2597.

Kgwatalala, P. M., Ibeagha-Awemu, E. M., Hayes, J. F., & Zhao, X. (2007). Single nucleotide polymorphisms in the open reading frame of the stearoyl-CoA desaturase gene and resulting genetic variants in Canadian Holstein and Jersey cows. *DNA Sequence – Journal of DNA Sequencing and Mapping*, *18*, 357–362.

Kliem, K. E., & Shingfield, K. J. (2016). Manipulation of milk fatty acid composition in lactating cows: Opportunities and challenges. *European Journal of Lipid Science and Technology*, *118*, 1661–1683.

Knudsen, S. (1999). Promoter2.0: For the recognition of PolII promoter sequences. *Bioinformatics*, *15*, 356–361.

Koch, L. E., & Lascano, G. J. (2018). Milk fat depression: Etiology, theories, and soluble carbohydrate interactions. *Journal of Animal Research and Nutrition*, *3*(2), 1–21.

Kratz, M., Baars, T., & Guyenet, S. (2013). The relationship between high-fat dairy consumption and obesity, cardiovascular, and metabolic disease. *European Journal of Nutrition*, *52*, 1–24.

Li, C., Sun, D., Zhang, S., Liu, L., Alim, M. A., & Zhang, Q. (2016). A post-GWAS confirming the SCD gene associated with milk medium- and long-chain unsaturated fatty acids in Chinese Holstein population. *Animal Genetics*, *47*, 483–490.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*, 2987–2993.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv Preprint *ArXiv:1303.3997*.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079.

Lopes, F. C. F., Silva, B. C. d. M., de Almeida, M. M., & da Gama, M. A. S. (2015). Lácteos naturalmente enriquecidos com ácidos graxos benéficos à saúde. *Embrapa Gado de Leite-Cap{\'\i}tulo Em Livro Cient{\'\i}fico (ALICE)*.

Madalena, F. E., Peixoto, M. G. C. D., & Gibson, J. (2012). Dairy cattle genetics and its applications in Brazil. *Livestock Research for Rural Development*, *24*(6), 1–49.

Mele, M., Conte, G., Castiglioni, B., Chessa, S., Macciotta, N. P. P., Serra, A., Buccioni, A., Pagnacco, G., & Secchiari, P. (2007). Stearoyl-coenzyme a desaturase gene polymorphism and milk fatty acid composition in Italian holsteins. *Journal of Dairy Science*, *90*, 4458–4465.

Milanesi, M., Obrien, A. M., Utsunomiya, A. T., Feres, L. F., Sonstegard, T. S., & Garcia, J. F. (2022). Genomic breed composition of pure registered Brazilian Gir. *Tropical Animal Health and Production*, *54*, 1–7.

Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, *16*, 1215.

Ntambi, J. M., & Miyazaki, M. (2004). Regulation of stearoyl-CoA desaturases and role in metabolism. *Progress in Lipid Research*, *43*, 91–104.

Pegolo, S., Cecchinato, A., Casellas, J., Conte, G., Mele, M., Schiavon, S., & Bittante, G. (2016). Genetic and environmental relationships of detailed milk fatty acids profile determined by gas chromatography in Brown Swiss cows. *Journal of Dairy Science*, *99*(2), 1315–1330.

Peixoto, M. G. C. D., Poggian, C. F., Verneque, R. S., Egito, A. A., Carvalho, M. R. S., Penna, V. M., Bergmann, J. A. G., Viccini, L. F., & Machado, M. A. (2010). Genetic basis and inbreeding in the Brazilian Guzerat (Bos indicus) subpopulation selected for milk production. *Livestock Science*, *131*, 168–174.

Peripolli, E., Reimer, C., Ha, N. T., Geibel, J., MacHado, M. A., Panetto, J. C. D. C., do Egito, A. A., Baldi, F., Simianer, H., & da Silva, M. V. G. B. (2020). Genome-wide detection of signatures of selection in indicine and Brazilian locally adapted taurine cattle breeds using whole-genome re-sequencing data. *BMC Genomics*, *21*, 624.

Peripolli, E., Stafuzza, N. B., Munari, D. P., Lima, A. L. F., Irgang, R., Machado, M. A., Panetto, J. C. d. C., Ventura, R. V., Baldi, F., & da Silva, M. V. G. B. (2018). Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (Bos indicus) dairy cattle. *BMC Genomics*, *19*, 34.

Pimentel, J. S. M., Carmo, A. O., Rosse, I. C., Martins, A. P. V., Ludwig, S., Facchin, S., Pereira, A. H., Brandão-Dias, P. F. P., Abreu, N. L., & Kalapothakis, E. (2018). High-throughput sequencing strategy for microsatellite genotyping using neotropical fish as a model. *Frontiers in Genetics*, *9*, 73.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.r-project.org

Rezamand, P., Watts, J. S., Yavah, K. M., Mosley, E. E., Ma, L., Corl, B. A., & Mcguire, M. A. (2014). Relationship between stearoyl-CoA desaturase 1 gene expression, relative protein abundance, and its fatty acid products in bovine tissues. *Journal of Dairy Research*, *81*, 333–339.

Rosse, I. D. C., Steinberg, R. D. S., Coimbra, R. S., Peixoto, M. G. C. D., Verneque, R. S., Machado, M. A., Fonseca, C. G., & Carvalho, M. R. S. (2014). Novel SNPs and INDEL polymorphisms in the 3′UTR of DGAT1 gene: In silico analyses and a possible association. *Molecular Biology Reports*, *41*, 4555–4563.

Samková, E., Špička, J., Pešek, M., Pelikánová, T., & Hanuš, O. (2012). Animal factors affecting fatty acid composition of cow milk fat: A review. *South African Journal of Animal Sciences*, *42*(2), 83–100.

Santana, M. L., Pereira, R. J., Bignardi, A. B., el Faro, L., Tonhati, H., & Albuquerque, L. G. (2014). History, structure, and genetic diversity of Brazilian Gir cattle. *Livestock Science*, *163*, 26–33.

Santos-Silva, J., Dentinho, M. T., Francisco, A., Portugal, A. P., Belo, A. T., Martins, A. P. L., Alves, S. P., & Bessa, R. J. B. (2016). Replacing cereals with dehydrated citrus pulp in a soybean oil supplemented diet increases vaccenic and rumenic acids in ewe milk. *Journal of Dairy Science*, *99*, 1173–1182.

Schennink, A., Heck, J. M. L., Bovenhuis, H., Visker, M. H. P. W., Van Valenberg, H. J. F., & Van Arendonk, J. A. M. (2008). Milk fatty acid unsaturation: Genetic parameters and effects of stearoyl-CoA desaturase (SCD1) and acyl CoA: Diacylglycerol acyltransferase 1 (DGAT1). *Journal of Dairy Science*, *91*, 2135–2143.

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). Mutationtaster2: Mutation prediction for the deep-sequencing age. *Nature Methods*, *11*(4), 361–362.

Shahlla, N. M., Obaid, U., & Riazuddin, S. (2014). Genetic polymorphism of milk protein variants and their association studies with milk yield in Sahiwal cattle. *African Journal of Biotechnology*, *13*, 555–565.

Shingfield, K. J., Bonnet, M., & Scollan, N. D. (2013). Recent developments in altering the fatty acid composition of ruminant-derived foods. *Animal*, *7*(s1), 132–162.

Sim, N. L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: Predicting effects of amino acid substitutions on proteins. *Nucleic Acids Research*, *40*, W452–W457.

Sinnwell, J. P., & Schaid, D. J. (2020). Haplo.Stats: Statistical analysis of haplotypes with traits and covariates when linkage phase is ambiguous. R Package Version 1.8.6. Available at: https://cran.r-project.org/package=haplo.stats

Soyeurt, H., Gillon, A., Vanderick, S., Mayeres, P., Bertozzi, C., & Gengler, N. (2007). Estimation of heritability and genetic correlations for the major fatty acids in bovine milk. *Journal of Dairy Science*, *90*, 4435–4442.

Stoop, W. M., Van Arendonk, J. A. M., Heck, J. M. L., Van Valenberg, H. J. F., & Bovenhuis, H. (2008). Genetic parameters for major milk fatty acids and milk production traits of dutch Holstein-Friesians. *Journal of Dairy Science*, *91*, 385–394.

Venn-Watson, S., Lumpkin, R., & Dennis, E. A. (2020). Efficacy of dietary odd-chain saturated fatty acid pentadecanoic acid parallels broad associated health benefits in humans: Could it be essential?. *Scientific Reports*, *10*(1), 1–14.

Vlaeminck, B., Fievez, V., Cabrita, A. R. J., Fonseca, A. J. M., & Dewhurst, R. J. (2006). Factors affecting odd- and branched-chain fatty acids in milk: A review. *Animal Feed Science and Technology*, *131*(3-4), 389–417.

Wallace, R. J., Sasson, G., Garnsworthy, P. C., Tapio, I., Gregson, E., Bani, P., Huhtanen, P., Bayat, A. R., Strozzi, F., Biscarini, F., & Snelling, T. J. (2019). A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. *Science Advances*, *5*(7), eaav8391.

Weaver, C. M. (2014). How sound is the science behind the dietary recommendations for dairy? *American Journal of Clinical Nutrition*, *99*, 1217S–1222S.

Wolff, R. L., Bayard, C. C., & Fabien, R. J. (1995). Evaluation of sequential methods for the determination of butterfat fatty acid composition with emphasis on trans-18:1 acids. Application to

the study of seasonal variations in french butters. *Journal of the American Oil Chemists' Society*, *72*(12), 1471–1483.

Zahn-Zabal, M., Michel, P. A., Gateau, A., Nikitin, F., Schaeffer, M., Audot, E., Gaudet, P., Duek, P. D., Teixeira, D., De Laval, V. R., et al. (2020). The neXtProt knowledgebase in 2020: Data, tools and usability improvements. *Nucleic Acids Research*, *48*(D1), D328–D334.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Matosinho, C. G. R., Fonseca, P. A. d. S., Peixoto, M. G. C. D., Rosse, I. C., Lopes, F. C. F., Zózimo, T., Filho, A. E. V., Bruneli, F. Â. T., Carvalho, M. R. S., & Gama, M. A. S. (2023). Phenotypic variation in milk fatty acid composition and its association with stearoyl-CoA desaturase 1 (*SCD1*) gene polymorphisms in Gir cows. *Journal of Animal Breeding and Genetics*, *140*, 532–548. https://doi.org/10.1111/jbg.12777