A single base pair duplication in the SLC33A1 gene is associated with fetal losses and neonatal lethality in Manech Tête Rousse dairy sheep

Maxime Ben Braiek, Soline Szymczak, Céline André, Philippe Bardou, Francis Fidelle, Itsasne Granado-Tajada, Florence Plisson-Petit, Julien Sarry, Florent Woloszyn, Carole Moreno-Romieux, et al.

Received: 6 March 2024 | Accepted: 8 June 2024

DOI: 10.1111/age.13459

RESEARCH ARTICLE

ANIMAL GENETICS WILEY

A single base pair duplication in the *SLC33A1* gene is associated with fetal losses and neonatal lethality in Manech Tête Rousse dairy sheep

Maxime Ben Braiek[1] | **Soline Szymczak**[1] | **Céline André**[2] | **Philippe Bardou**[3] | **Francis Fidelle**[2] | **Itsasne Granado-Tajada**[4] | **Florence Plisson-Petit**[1] | **Julien Sarry**[1] | **Florent Woloszyn**[1] | **Carole Moreno-Romieux**[1,†] | **Stéphane Fabre**[1]

[1]GenPhySE, Université de Toulouse, INRAE, ENVT, Castanet-Tolosan, France

[2]CDEO, Ordiarp, France

[3]Sigenae, INRAE, Castanet-Tolosan, France

[4]Department of Animal Production, NEIKER-BRTA Basque Institute of Agricultural Research and Development, Arkaute, Spain

Correspondence
Stéphane Fabre, GenPhySE, Université de Toulouse, INRAE, ENVT, F-31326, Castanet-Tolosan, France.
Email: stephane.fabre@inrae.fr

Funding information
Ministère de l'Agriculture et de l'Alimentation, Grant/Award Number: PRESAGE/CASDAR n°20ART1532777; APIS-GENE; Région Occitanie Pyrénées-Méditerranée; Horizon 2020 Framework Programme, Grant/Award Number: 772787

Abstract

We recently discovered that the Manech Tête Rousse (MTR) deficient homozygous haplotype 2 (MTRDHH2) probably carries a recessive lethal mutation in sheep. In this study, we fine-mapped this region through whole-genome sequencing of five MTRDHH2 heterozygous carriers and 95 non-carriers from various ovine breeds. We identified a single base pair duplication within the *SLC33A1* gene, leading to a frameshift mutation and a premature stop codon (p.Arg246Alafs*3). SLC33A1 encodes a transmembrane transporter of acetyl-coenzyme A that is crucial for cellular metabolism. To investigate the lethality of this mutation in homozygous MTR sheep, we performed at-risk matings using artificial insemination (AI) between heterozygous *SLC33A1* variant carriers (*SLC33A1_dupG*). Pregnancy was confirmed 15 days post-AI using a blood test measuring interferon Tau-stimulated MX1 gene expression. Ultrasonography between 45 and 60 days post-AI revealed a 12% reduction in AI success compared with safe matings, indicating embryonic/fetal loss. This was supported by the *MX1* differential expression test suggesting fetal losses between 15 and 60 days of gestation. We also observed a 34.7% pre-weaning mortality rate in 49 lambs born from at-risk matings. Homozygous *SLC33A1_dupG* lambs accounted for 47% of this mortality, with deaths occurring mostly within the first 5 days without visible clinical signs. Therefore, appropriate management of *SLC33A1_dupG* with an allele frequency of 0.04 in the MTR selection scheme would help increase overall fertility and lamb survival.

KEYWORDS
abortion, acetyl-CoA transporter, homozygous haplotype deficiency, lethal variant, MTRDHH2, neonatal mortality, *Ovis aries*, *SLC33A1*

[†]In memoriam.

INTRODUCTION

Most individuals carry roughly 100 loss-of-function alleles in their genome, and among them one to five are predicted to be deleterious (Georges et al., 2019; MacArthur et al., 2012). These variants can cause severe defects, leading to lethality (from the embryonic stage to adulthood) or morphological disorders in the homozygous state. Therefore, the identification of the causal variants is essential for managing these defects.

Recent advances in the use of massive genomic data have made it possible to discover the causative variants for many genetic disorders. The human genome was first sequenced in 2001, opening up new possibilities for the fine-mapping of causal variants (Lander et al., 2001). As of 7 March 2023, the OMIM database (https://www.omim.org), which lists human mendelian disorders, identified 7342 entries for which the molecular basis is known. The emergence of next-generation sequencing for livestock has also enabled access to the entire genomes of many individuals and a comprehensive number of genetic markers, such as single nucleotide polymorphisms (SNPs), representative of the individual genome variability (Eggen, 2012; Rupp et al., 2016).

When biological samples and phenotype descriptions of affected cases are available, fine mapping of causative variants can be performed using genome-wide association studies with a case–control approach (Uffelmann et al., 2021) or homozygosity mapping (Charlier et al., 2008; Lander & Botstein, 1987), followed up with whole-genome sequencing (WGS) analyses. In contrast, the reverse genetic screen strategy uses a method that was initially developed to identify lethal variants without any available phenotype (Fritz et al., 2013; VanRaden et al., 2011). This method uses high-throughput SNP genotyping data to detect deficits in homozygous animals for some specific haplotypes, which leads to a significant deviation from the Hardy–Weinberg equilibrium of genotypes. The transmission probability within trios is used to calculate this deviation (VanRaden et al., 2011). The routine use of SNP genotyping in genomic selection in livestock has made it possible to use this method in many species, including cattle (Fritz et al., 2013; VanRaden et al., 2011), beef (Jenko et al., 2019), pigs (Derks et al., 2017), chicken (Derks et al., 2018), turkey (Abdalla et al., 2020) and horses (Todd et al., 2020). We also recently discovered numerous independent deficient homozygous haplotypes (DHHs) in Lacaune and Manech Tête Rousse (MTR) dairy sheep breeds (Ben Braiek et al., 2021, 2024). Utilizing WGS data and controlled matings between DHH carriers, we identified two gene variants responsible for severe health issues in sheep. One variant, situated in the *CCDC65* gene (LDHH6, OMIA 002342–9940), causes respiratory problems leading to death in young Lacaune sheep (Ben Braiek et al., 2022). The other variant, located in the *MMUT* gene (MTRDHH1), disrupts propionic acid metabolism, causing death shortly after birth in MTR sheep (Ben Braiek et al., 2024). Both DHHs were initially linked to a significant increase in stillbirth rates in at-risk matings, as identified through population data analysis. Furthermore, several other identified DHHs, LDHH1-2-8-9 in Lacaune and MTRDHH2 in MTR, appear to have reduced the success rate of artificial insemination (AI) by 3%. This suggests the potential presence of lethal embryonic variants within these DHHs (Ben Braiek et al., 2021, 2024).

Embryonic and early fetal losses are challenging to detect, yet significantly impact the economic success of sheep breeders (Diskin & Morris, 2008; Dixon et al., 2007). While predominantly caused by environmental factors, lethal genetic variants can also contribute to these losses (VanRaden et al., 2011). Ultrasonography can assess fertilization failure and embryonic losses in sheep at 45–60 days post-fertilization. However, the majority of embryonic losses occur between fertilization (day 0) and conceptus implantation (days 12–16) (Bindon, 1971; Spencer et al., 2008; Wilmut et al., 1986). Implantation involves the conceptus binding to the endometrium and the secretion of interferon tau, the key pregnancy recognition signal (Bazer, 2013). In response, interferon-stimulated genes (ISG) are expressed both in the endometrium and in circulating blood immune cells (Ott & Gifford, 2010). Specific ISGs like *MX1* (Myxovirus-influenza virus resistance 1), *STAT1* (Signal transducer and activator of transcription 1) and *CXCL10* (Chemokine C-X-C motif ligand 10) have shown promise in predicting gestation status in sheep based on blood samples taken at days 14–15 (Mauffré et al., 2016).

The MTRDHH2-deficient haplotype in MTR dairy sheep reduces both AI success by 3.0% and the stillbirth rate by 4.3%. This suggests that the haplotype might harbor a lethal variant affecting homozygous fetuses throughout gestation (Ben Braiek et al., 2024). Located on ovine chromosome 1 (NC_040252.1, OAR1:251.9–256.4 Mb on the Rambouillet_v1.0 sheep genome), the haplotype has an estimated carrier frequency of 8.7% among heterozygotes. We previously identified a potential functional candidate gene, *SLC33A1*, within the MTRDHH2 region (Ben Braiek et al., 2024). This study aimed to identify the causative variant associated with MTRDHH2 and validate its potential embryonic lethal effect. We achieved this by conducting at-risk matings and monitoring early gestation at day 15 using ISG molecular diagnosis and at days 45–60 using ultrasonography.

MATERIALS AND METHODS

Sequencing data, WGS variant calling and annotation

The list of the 100 publicly available ovine WGSs (short read Illumina HiSeq/Nova Seq) from 14 different breeds generated in various INRAE and Teagasc research projects is detailed in Table S1 (project and sample accession

numbers). Among them, 22 WGSs were from Manech Tête Rousse dairy sheep, which were also genotyped with the Illumina OvineSNP50 Beadchip as part of a dairy sheep genomic selection program (Astruc et al., 2016). The 22 animals had a known status at the MTRDHH2 haplotype encoded 0 for non-carrier ($n=17$), 1 for heterozygous carrier ($n=5$) and 2 for homozygous carrier ($n=0$) (Ben Braiek et al., 2024). Read mapping, WGS variant calling and annotation were performed using NEXTFLOW v20.10.0 and SAREK PIPELINE v2.6.1 as already described (Ben Braiek et al., 2024).

Identification of candidate causal variants

All SNPs and InDel variants located within the MTRDHH2 haplotype region extended by 1 Mb from each side were extracted from OAR1 (Oar_rambouillet_v1.0; NC_040252.1:250858291–257412373pb) using SnpSift Filter, part of the SnpEff toolbox (Cingolani et al., 2012). Correlation analysis was performed between the MTRDHH2 haplotype status (0, 1 and 2) and allele dosage for bi-allelic variants (also encoded 0, 1 or 2) using the geno-r2 command of VCFTOOLS (Danecek et al., 2011). Only variants with a perfect correlation with the haplotype status ($r^2=1$) were retained. The selected variants were checked manually using the INTEGRATIVE GENOMICS VIEWER (Thorvaldsdóttir et al., 2013) using the BAM files from the 22 WGS.

Biological samples

The experimental design is described in Figure 1. Jugular blood samples from 419 MTR dairy ewes (all daughters of MTRDHH2 carrier sires) located on six private farms were collected by a habilitated veterinarian (Venoject system containing EDTA, Terumo, Tokyo, Japan) and stored at −20°C either for further genotyping ($n=419$) or for RNA extraction ($n=137$). Ear biopsies (1 mm^3) from the 49 newborn lambs were obtained with a tissue sampling unit (Allflex Europe, Vitré, France) and directly placed in the tissue sampling unit storage buffer at 4°C. Ear biopsies were treated with consecutive action of NaOH and Tris–HCl for subsequent genotyping as previously described (Ben Braiek et al., 2024).

DNA sets from MTR genomic lambs ($n=714$) with a known status at MTRDHH2 (Ben Braiek et al., 2024), Manech-related Latxa Spanish sheep ($n=100$) and a diversity panel from 25 French sheep breeds ($n=749$) (Rochus et al., 2018) were also used for single marker genotyping.

SLC33A1 specific genotyping assay

Among the selected variants located in MTRDHH2, a PCR allele competitive extension (PACE) genotyping assay was developed for the variant NC_040252.1:g.252649023dupG located in the *SLC33A1* gene. Fluorescent PACE analysis was done with 15 ng of purified DNA (from DNA panels) using the PACE-IR 2× Genotyping Master mix (3CR Bioscience) in the presence of 12 µM of a mix of extended allele specific forward primers (Table S2) in a final volume of 10 µL. The touch-down PCR amplification condition was 15 min at 94°C for the hot-start activation, 10 cycles of 20 s at 94°C, 54–62°C for 60 s (dropping 0.8°C per cycle), then 36 cycles of 20 s at 94°C and 60 s at 54°C performed on an ABI9700 thermocycler followed by a final point

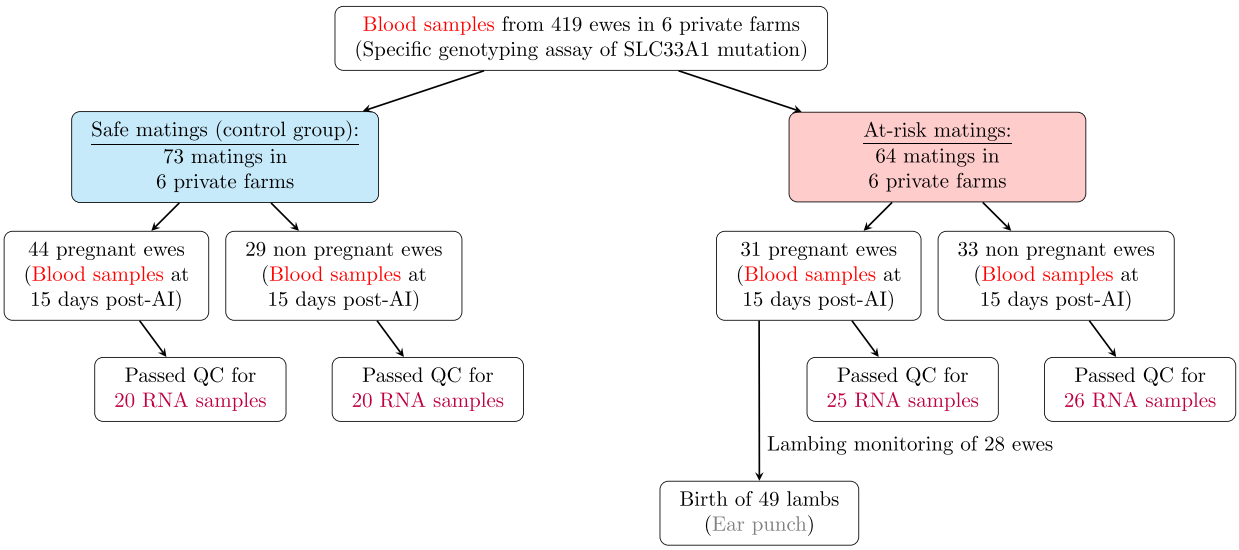

FIGURE 1 Experimental design. Two groups of ewes were artificially inseminated (AI) with ram fresh semen: 73 ewes in safe matings and 64 ewes in at-risk matings. Safe matings refer to matings between non-carrier ewes and rams, while at-risk matings refer to matings between heterozygous carrier ewes and rams. QC stands for quality and quantity control.

read of the fluorescence on an ABI QuantStudio 6 real-time PCR system and using the QUANTSTUDIO software 1.3 (Applied Biosystems). For the genotyping of crude biological samples (whole blood or neutralized NaOH treatment solution of ear biopsies), a preliminary Terra PCR Direct Polymerase mix amplification (Takara Bio, Kusatsu, Japan) using 1 μL of crude sample was used for direct genotyping without DNA purification. The amplification primers (Table S2) were designed using PRIMER3PLUS software (Untergasser et al., 2007). This preliminary PCR was performed on an ABI2720 thermocycler (Applied Biosystems, Waltham, MA, USA) with the following conditions: 5 min at 94°C, 35 cycles of 30 s at 94°C, 30 s at 58°C and 30 s at 72°C, followed by 5 min final extension at 72°C. Then, 1 μL of the PCR product was used for subsequent PACE genotyping.

Programmed mating

Of the 419 daughters of MTRDHH2 carrier sires, 137 ewes were retained for programmed mating based on their *SLC33A1* genotyping results. Two groups of mating were created: safe mating ($n=73$ ewes) between non-carrier ewes and rams, and at-risk mating ($n=64$ ewes) between heterozygous carriers. All ewes were artificially inseminated with fresh semen. A jugular blood sample was collected from each mated ewe 15 days after AI for further gestation molecular diagnosis testing. An ultrasound diagnosis of gestation was performed between 45 and 60 days after AI. Gestations were monitored, and each lamb was monitored from birth to weaning age.

Molecular diagnosis of gestation

RNA extraction, reverse transcription and real-time PCR

Total RNAs were extracted from blood of 137 ewes with the Nucleospin® RNA Blood Kit (Macherey-Nagel, reference 740200.50) according to the manufacturer's protocol starting with 800 μL of whole blood with a DNAseI digestion treatment to eliminate contaminating genomic DNA. RNAs was quantified by spectrophotometry (NanoDrop® ND-8000 spectrophotometer, ThermoFischer) and stored at −80°C. After quality and quantity control, 91 RNA samples ($n=40$ ewes in the safe mating control group and $n=51$ ewes in the at-risk mating group) were kept for reverse transcription. Reverse transcription was carried out from 500 ng of total RNA in solution with anchored oligo(dT) T22V (1 μL at 100 μM), random oligo-dN9 (1 μL at 100 μM) and dNTPs (2 μL at 10 mM) in a reaction volume of 54 μL. This mixture was incubated at 65°C for 5 min in an ABI2700 thermocycler (Applied Biosystems) then ramped down to 4°C. A second reaction mixture (18.5 μL/reaction) containing the reaction buffer (14 μL of First strand Buffer 5×, Invitrogen, France), dithiothreitol (3 μL at 0.1 M), Rnasine (0.5 μL, 40 units/μL, Promega, France) and Superscript II reverse transcriptase (1 μL, 200 units/μL, Invitrogen, France) was added to the denatured RNA solution (final volume reaction of 72.5 μL) then incubated for 50 min at 42°C and placed for 15 min at 70°C. The complementary DNA (cDNA) solution obtained was directly diluted at a 1:2 ratio and stored at −20°C. Quantitative PCR (qPCR) was performed using 3 μL of cDNA, 5 μL of SYBR Green real-time PCR Master Mix 2× (Applied Biosystems) and 2 μL of primers at 3 μM in a total reaction volume of 10 μL. Each sample was tested in quadruplicate. The qPCR was realized on a QuantStudio 6 Flex Real-Time PCR system (ThermoFisher). The amplification efficiency for each primer pair was evaluated using the equation $E = e^{-1/\alpha}$ where α is the slope of a linear curve obtained from cDNA serial dilution and corresponding *Ct* (cycle threshold) values. RNA transcript abundance was quantified using the delta *Ct* (ΔCt) method corrected by two reference genes (*GAPDH*, *YWHAZ*). Primers were designed using BEACON DESIGNER 8 (Premier Biosoft). The qPCR primer sequences, amplification lengths and amplification efficiencies are available in Table S2.

Statistical analyses

The gestation status of the ewes (pregnant/non-pregnant) was first determined by ultrasound diagnosis and then confirmed at the time of lambing. The relative expressions of *MX1* and *STAT1* (Interferon Stimulated Genes, ISG) were compared between pregnant and non-pregnant ewes using the Wilcoxon non-parametric test.

To evaluate the gestation diagnosis molecular test (GDMT) at day 15, we analyzed ISG expression levels. Prediction accuracy was assessed using receiver operating characteristic (ROC) curve analysis with easyROC software (Goksuluk et al., 2016). We established a predictive model using expression data from the safe mating control group (training data). This model was then applied to expression data from the at-risk mating group, assumed to be affected by the *SLC33A1* lethal embryonic variant (testing data). For the training data, we calculated the area under the curve (AUC) and compared it with the expected null hypothesis value of 0.5. The ROC01 cut-off method [which minimizes the distance between the ROC curve and the point (100-Sp = 0, Se = 100)] was used to maximize sensitivity (Se), which is the ability of the GDMT to correctly detect pregnant ewes by ultrasonography. This cut-off value was used to classified the ISG relative expression in four categories: true positive (TP, i.e. GDMT+ and pregnant), false positive (FP, i.e. GDMT+ and non-pregnant), true negative (TN, i.e. GDMT- and non-pregnant) and false negative (FN, i.e. GDMT− and pregnant). The test estimators

were generalized using the prevalence (Pr), which is the probability of being pregnant at days 45–60, corresponding to the population mean of 60.3% in the Manech Tête Rousse overall population (Ben Braiek et al., 2024). The positive predictive value (PPV), i.e. the number of ewes with a positive GDMT and pregnant at days 45–60 among the number of ewes with positive GDMT, and negative predictive value (NPV), i.e. the number of ewes with a negative GDMT and non-pregnant at days 45–60 among the number of ewes with negative GDMT, was determined by

$$\mathrm{PPV} = \frac{\mathrm{Se} \times \mathrm{Pr}}{\mathrm{Se} \times \mathrm{Pr} + (1 - \mathrm{Sp}) \times (1 - \mathrm{Pr})}$$

and

$$\mathrm{NPV} = \frac{\mathrm{Sp} \times (1 - \mathrm{Pr})}{\mathrm{Sp} \times (1 - \mathrm{Pr}) + (1 - \mathrm{Se}) \times \mathrm{Pr}}$$

For testing data, the cut-off value of the training model was used to classify the ewes of the at-risk mating group in the TP, FP, TN and FN categories (observed numbers). The model parameters (Se, Sp, Pr, PPV and NPV) were used to estimate the expected numbers in TP, FP, TN and FN categories. Chi-squared tests were performed between the observed and expected number of ewes with positive and negative GDMT, respectively.

RESULTS

Identification of a single base pair duplication in the *SLC33A1* gene

To identify the causative variant located in the MTRDHH2 haplotype, we scanned the genomes of 100 ovine WGSs, focusing on polymorphisms located in the MTRDHH2 region extended by 1 Mb from each side (OAR1:250858291–257 412 373 bp). In this 6.5 Mb region, we detected 111 984 polymorphisms (variant call rate ≥ 95%, quality score > 30). Among the WGS animals, 22 were from the Manech Tête Rousse breed and five of them were heterozygous carriers of MTRDHH2.

We performed Pearson correlation analysis to compare biallelic variant status (SNPs and InDels) with the MTRDHH2 haplotype status (encoded as 0, 1, and 2). This analysis revealed 189 polymorphisms perfectly correlated with MTRDHH2 status (Figure 2a). Of these, only one variant (NC_040252.1, OAR1:g.252649023dupG; Table S3) was located within a coding sequence and predicted to significantly disrupt the solute carrier family 33 member 1 (*SLC33A1*) gene (XM_012100950.3, c.735dupG). This variant, designated *SLC33A1_dupG*, involves a single guanine duplication within the first exon of the gene (Figure 2b,c). This duplication is predicted to introduce a premature stop codon three amino acids downstream (XP_011956340.1, p.Arg246Alafs*3). The resulting protein, if translated, would be truncated to 248 amino acids, compared with the 550 amino acids of the full-length protein (Figure 2d). The wild-type SLC33A1 protein has nine transmembrane domains (UniProt: A0A6P3TI15_SHEEP), whereas the mutant form would only contain 3.

Variant association with MTRDHH2 and diversity analysis

To validate the association between the putative causal variant in the *SLC33A1* gene and MTRDHH2, we genotyped 714 Manech Tête Rousse lambs from the 2021 genomic cohort with known MTRDHH2 status (Ben Braiek et al., 2024). The variant allele had a frequency of 4%. A contingency table analysis revealed a strong association between the *SLC33A1_dupG* genotype and MTRDHH2 haplotype status (Table 1, Fisher's exact test, $p < 0.0001$, excluding homozygous carriers). Interestingly, all MTRDHH2 heterozygous carriers were also heterozygous for the *SLC33A1_dupG* variant. However, 14 of the 56 heterozygous animals for *SLC33A1_dupG* were not MTRDHH2 heterozygous carriers. Further analysis of the 66 SNPs within the MTRDHH2 haplotype revealed shorter recombinant versions (from 5 to 65 SNPs), which could potentially explain this discrepancy (Figure S1).

We investigated the segregation of the *SLC33A1_dupG* variant across a panel of 28 French and Spanish sheep breeds. The variant was detected in the heterozygous state in three French MTR animals and one Spanish Latxa Cara Rubia animal (Table 2).

Association of *SLC33A1*_dupG with decreased AI success and lamb mortality

To investigate the *SLC33A1_dupG* variant's effect on offspring viability, we generated groups for controlled mating. We genotyped 419 MTR ewes and selected 137 ewes to constitute two mating groups. The safe mating group was composed of 73 non-carrier ewes mated with non-carrier rams through AI. The at-risk mating group was composed of 64 ewes heterozygous for *SLC33A1_dupG* mated with heterozygous rams through AI. Gestation was monitored through blood sampling (day 15), ultrasound (days 45–60) and lambing records (151 ± 7 days post-AI). As expected, the safe mating group achieved an AI success rate (confirmed by lambing) of 60.3%, consistent with the overall MTR population (Ben Braiek et al., 2024). However, the at-risk group with *SLC33A1_dupG* carriers showed a reduced success rate of 48.4% (Figure 3). While statistically non-significant ($p = 0.17$), this 12% reduction holds relevance in sheep breeding.

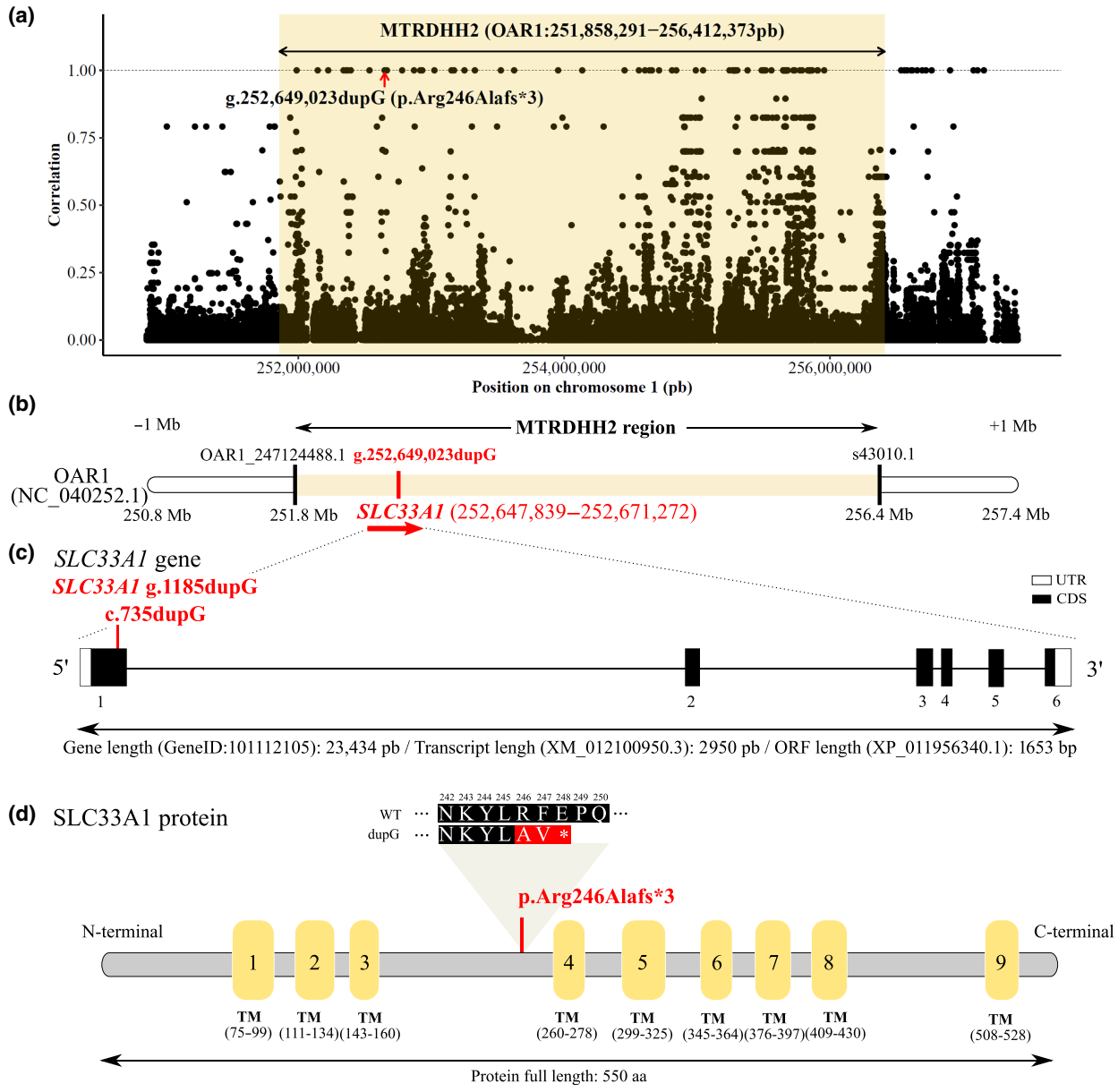

FIGURE 2 Identification of a single base pair duplication in the *SLC33A1* gene. (a) A scatter showing the correlation between MTRDHH2 status (NC_040252.1, OAR1:251858291–256412373pb extended from each side by 1 Mb) and the genotype of variants from 100 whole genome sequenced animals. Each dot represents one variant. (b) The position of the *SLC33A1* gene within the MTRDHH2 haplotype. Black bars indicate the first and the last markers of the Illumina Ovine SNP50 BeadChip, which define the limits of MTRDHH2 (Ben Braiek et al., 2024). (c) The structure of the *SLC33A1* gene (GeneID: 101112105) and the localization of the c.735dupG variant identified in the first exon (XM_012100950.3) (UTR, untranslated region; CDS, coding sequence). (d) The SLC33A1 protein (XP_011956340.1) with UniProt domain annotations (accession number: A0A6P3TI15_SHEEP). The protein is composed of 9 transmembrane domains (TM). The single base pair duplication creates a premature stop codon at amino-acid position 248.

Among the 31 pregnant ewes in the at-risk mating group, 28 were closely monitored throughout gestation, birth and the first month post-lambing. No abortions occurred, and lambs were born at a mean gestation of 151 ± 7 days. Genotyping of 49 lambs (18 females, 30 males, one undetermined) revealed nine homozygous for the *SLC33A1_dupG* variant (Figure 4a). A contingency table analysis (Table 3) confirmed significantly reduced viability of homozygous lambs compared to heterozygous ones (Fisher's exact test, p-value <0.001). Homozygous *SLC33A1_dupG* lambs constituted 47% of the mortality in the at-risk group. Most deaths occurred within the first 5 days post-lambing (Figure 4b). Two stillborn lambs exhibited developmental arrest characteristic of mid-gestation (one genotyped, Figure 4c). The majority of deceased homozygous lambs lacked visible morphological defects (Figure 4d). One exception died between 5 and 30 days of age, presenting with leg weakness, stiffness and spinal deformity leading to mobility issues (Figure 4e).

Genotype	+/+	MTRDHH2/+	MTRDHH2/MTRDHH2	Total
WT/WT	658	0	0	658
WT/dupG	14	42	0	56
dupG/dupG	0	0	0	0
Total	672	42	0	714

Abbreviations: +/+, non-carriers; MTRDHH2/+, heterozygous carriers; MTRDHH2/MTRDHH2, homozygous carriers (Fisher's exact test, $p<0.001$, without the homozygous MTRDHH2 carriers).

TABLE 1 Contingency table between MTRDHH2 status and *SLC33A1_dupG* genotype.

TABLE 2 *SLC33A1_dupG* genotype distribution from a DNA diversity panel of French (FR) and Spanish (ES) ovine breeds.

		Genotype				Genotype	
Breed	Total	WT/WT	WT/dupG	Breed	Total	WT/WT	WT/dupG
Berrichon du Cher (FR)	30	30		Martinik (FR)	23	23	
Blanche du Massif Central (FR)	31	31		Merinos d'Arles (FR)	27	27	
Causse du Lot (FR)	32	32		Mourerous (FR)	26	26	
Charmoise (FR)	31	31		Mouton Vendéen (FR)	30	30	
Charollais (FR)	30	30		Noir du Velay (FR)	28	28	
Corse (FR)	30	30		Préalpes du sud (FR)	27	27	
Ile de France (FR)	28	28		Rava (FR)	29	29	
Lacaune (Meat) (FR)	45	45		Romane (FR)	30	30	
Lacaune (Milk) (FR)	40	40		Romanov (FR)	26	26	
Latxa Cara Negra Euskadi (ES)	30	30		Rouge de l'Ouest (FR)	30	30	
Latxa Cara Negra Navarra (ES)	40	40		Roussin (FR)	30	30	
Latxa Cara Rubia (ES)	30	29	1	Suffolk (FR)	29	29	
Limousine (FR)	29	29		Tarasconnaise (FR)	33	33	
Manech Tête Rousse (FR)	28	25	3	Texel (FR)	27	27	
				Total	849	845	4

Molecular diagnosis of gestation at day 15 to assess for fetal losses

The reduced AI success rate observed in matings between *SLC33A1_dupG* heterozygous carriers suggested potential loss of homozygous embryos during early gestation. To explore this possibility, we adapted a molecular test (GDMT) to predict gestation status at this crucial stage. The adapted test relies on blood mRNA levels of two interferon-stimulated genes, *MX1* and *STAT1* (Mauffré et al., 2016). RNA blood samples from 91 ewes ($n=40$ in the safe mating group, $n=51$ in the at-risk mating group) were analyzed by RT-qPCR (Figure 1). The GDMT was assessed at day 15 to predict the gestation status (pregnant or not pregnant) at days 45–60. As shown in Figure 5a, the relative expression of *MX1* was significantly higher in pregnant ewes (84.3%) than in non-pregnant ewes (67.7%) (p-value=0.043, Wilcoxon test). The same suggestive observation could be made for *STAT1* expression, although the difference was not significant (p-value=0.076, Wilcoxon test). Thus, we only used the relative expression of *MX1* for the GDMT, and we tested the reliability of the diagnostic using a ROC curve on data from the safe mating group. The ROC plot showed an AUC of 0.687, which was significantly different from 0.5 (p-value=0.033). The cut-off value, as the decision threshold, was set at 63%, and this was associated with a sensitivity of 75% and a specificity of 60% (Table 4, Figure 5c,d). The prevalence (Pr) of AI success was set at 60.3% (the average of MTR population), which helped us to calculate the PPV and NPV. The PPV was 74% and the NPV was 61%. The cut-off value was used to classify ewes in four categories (TP, FP, TN and FN) based on their GDMT results (+ or −) and ultrasound diagnostics at days 45–60 (Figure 5e). In the safe mating control group, positive GDMT was able to detect 65% of pregnant ewes (TP/(TP+FP)). However, no significant difference in *MX1* expression was observed between pregnant and non-pregnant ewes in the at-risk mating group (p-value=0.27, Wilcoxon test, Figure 5f). Therefore, the molecular diagnostic parameters (cut-off=63%, Pr=60.3%, PPV=74% and NPV=61%) were applied to the *MX1* expression data from the at-risk mating group. In this group, the comparison of the observed and expected number of ewes with a positive GDMT (*MX1* relative expression >63%) showed a trend toward fewer

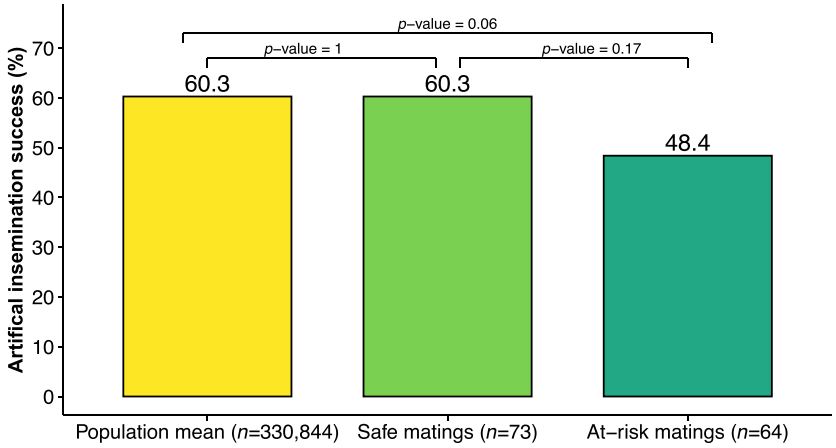

FIGURE 3 Artificial insemination success in MTR. Artificial insemination success in the MTR population is based on lambing date, which is calculated from the day of AI and takes into account the gestation length of 151 ± 7 days. This success rate was established using population mating records from 330 844 animals (Ben Braiek et al., 2024). Safe and at-risk matings between *SLC33A1*_dupG carriers were carried out in 6 private farms (n = 137 ewes). Gestation was diagnosed by ultrasound between days 45 and 60. In these experimental groups, artificial insemination success could be corrected by lambing record. p-values were obtained using Fisher's exact test.

pregnant ewes (Chi-squared test, p = 0.055). Specifically, 16 pregnant ewes were expected with a positive GDMT, but only 12 were observed. No significant difference was found between the observed and expected number of pregnant ewes with a negative GDMT (chi-squared test, p = 0.44). These results suggest that the molecular diagnosis probably evidenced fetal losses in the at-risk mating group between 15 days and 60 days of gestation. Four ewes were supposed to host homozygous fetuses for the *SLC33A1_dupG* variant.

DISCUSSION

In a recent study, we reported five haplotypes associated with a deficit in homozygous animals in MTR dairy sheep. One of these haplotypes, MTRDHH1, harbors a recessive lethal variant in the *MMUT* gene, leading to an increased stillbirth rate (Ben Braiek et al., 2024). This present study focuses on MTRDHH2, the second most impactful haplotype, which negatively affects both AI success (−3.0%) and stillbirth rate (+4.3%). These findings suggest that MTRDHH2 might also harbor a recessive lethal variant. Using WGS data, we mapped the candidate causative mutation to a single-base pair duplication (NC_040252.1:g.252649023dupG) in the *SLC33A1* gene. This gene was previously identified as a functional candidate gene for MTRDHH2 (Ben Braiek et al., 2024).

SLC33A1 (also known as AT-1, acetyl-CoA transporter 1) is an essential protein that transports acetyl-CoA through the endoplasmic reticulum (ER) membrane (Jonas et al., 2010). It is important for the *N*-lysine acetylation of ER proteins and regulating the degradation of protein aggregates by autophagy. Malfunction of SLC33A1 leads to the accumulation of protein aggregates within the ER, resulting in cell death (Peng et al., 2014). The sheep gene atlas indicates widespread expression of *SLC33A1* (http://biogps.org/sheepatlas/; accessed 18 July 2022) (Clark et al., 2017). We found that the *SLC33A1_dupG* variant introduces a premature stop codon, resulting in a shortened protein of 248 amino acids (XP_011956340.1:p.Arg246Alafs*3), compared with the full-length 550 amino acids. This variant causes the loss of six out of nine transmembrane domains, effectively creating a natural knockout in sheep. Notably, *Slc33a1* knockout in mice leads to embryonic lethality and reduced survival (MGI:1332247, http://www.informatics.jax.org). Zebrafish with reduced SLC33A1 show defective axon outgrowth affecting BMP signaling (Liu et al., 2017). In humans, mutations in this gene have been linked to various conditions, including congenital cataracts, hearing loss, neurodegeneration and autosomal dominant spastic paraplegia (Huppke et al., 2012; Lin et al., 2008; Liu et al., 2017; Mao et al., 2015; OMIM 612539, OMIM 614482, https://omim.org). These findings align with the decreased AI success in *SLC33A1_dupG* heterozygous crosses, increased mortality of homozygous *dupG*/*dupG* lambs and the observation of a *dupG*/*dupG* lamb with locomotor issues resembling spastic paraplegia. While mutations in *SLC33A1* have been associated with low serum copper and ceruloplasmin in humans (Huppke et al., 2012), we did not perform biochemical analyses.

The generation of homozygous *SLC33A1 dupG*/*dupG* lambs corroborated the association between MTRDHH2 and a higher rate of stillbirth. Heterozygous animals were unaffected. Notably, one homozygous *SLC33A1 dupG*/*dupG* lamb survived for 30 days without any apparent clinical signs. However, this lamb was subsequently slaughtered for commercial use, hindering further investigation and data collection. This single case suggests incomplete penetrance of the mutation. At gestation, we observed a 12% reduction in AI success in at-risk matings compared with safe matings, suggesting potential embryonic losses.

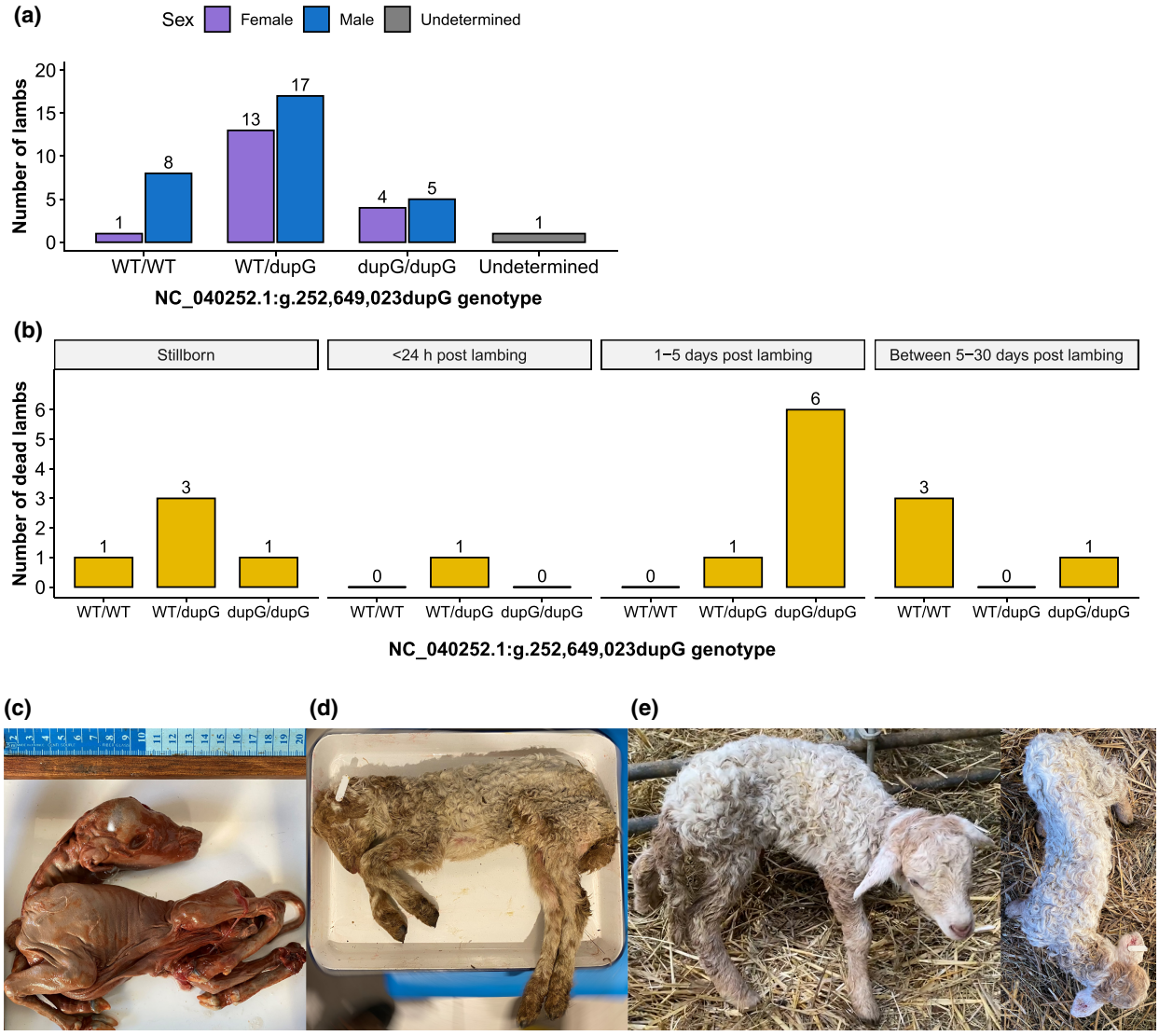

FIGURE 4 Impact of the *SLC33A1_dupG* variant on lamb survival. (a) Distribution of the 49 lambs born from 28 pregnant ewes according to the *SLC33A1*_dupG genotype (one stillborn lamb not genotyped). (b) Distribution of 17 dead lambs in the pre-weaning period depending on *SLC33A1*_dupG genotype. (c) Stillborn lamb. (d) Dead lamb in the first 5 days post-lambing (1.750 kg). (e) Alive lamb after 5 days post-lambing (dead before weaning) with morphologic and locomotor defects.

TABLE 3 Contingency table between lamb *SLC33A1_dupG* genotypes and viability.

Viability	WT/WT	WT/dupG	dupG/dupG	−/−	Total
Alive	5	25	1	0	31
Dead	4	5	8	1	18
Total	9	30	9	1	49

Note: −/−, undermined genotype owing to absence of biological samples for a stillborn lamb. This animal is not taken into account for mortality rate calculation.

To investigate this, we employed a molecular gestation diagnosis 15 days post-IA (embryo implantation) and compared it with the standard ultrasound diagnosis (45–60 days post-IA). This molecular test, utilizing the expression of the interferon-tau stimulated gene *MX1* in blood cells, predicted 16 gestations, while only 12 resulted in lambing. This indicates probable fetal losses between 15 and 60 days of gestation in four heterozygous ewes, probably owing to the lethality of the homozygous dupG/dupG genotype. The role of SLC33A1 in acetyl-CoA metabolism may explain the observed defects in embryogenesis. Acetyl-CoA is crucial for cell growth, proliferation and apoptosis, suggesting that a metabolic deficiency caused by the mutation leads to embryo failure (Tsuchiya et al., 2014).

Several methods exist for early pregnancy detection, including transrectal ultrasonography (Rickard et al., 2017) and circulating biomarker analysis (e.g. progesterone, protein B PSP-B, or pregnancy-associated glycoproteins; Karen et al., 2003). However, these methods are generally reliable only from day 28 of gestation, following the expected return to estrus (Mauffré

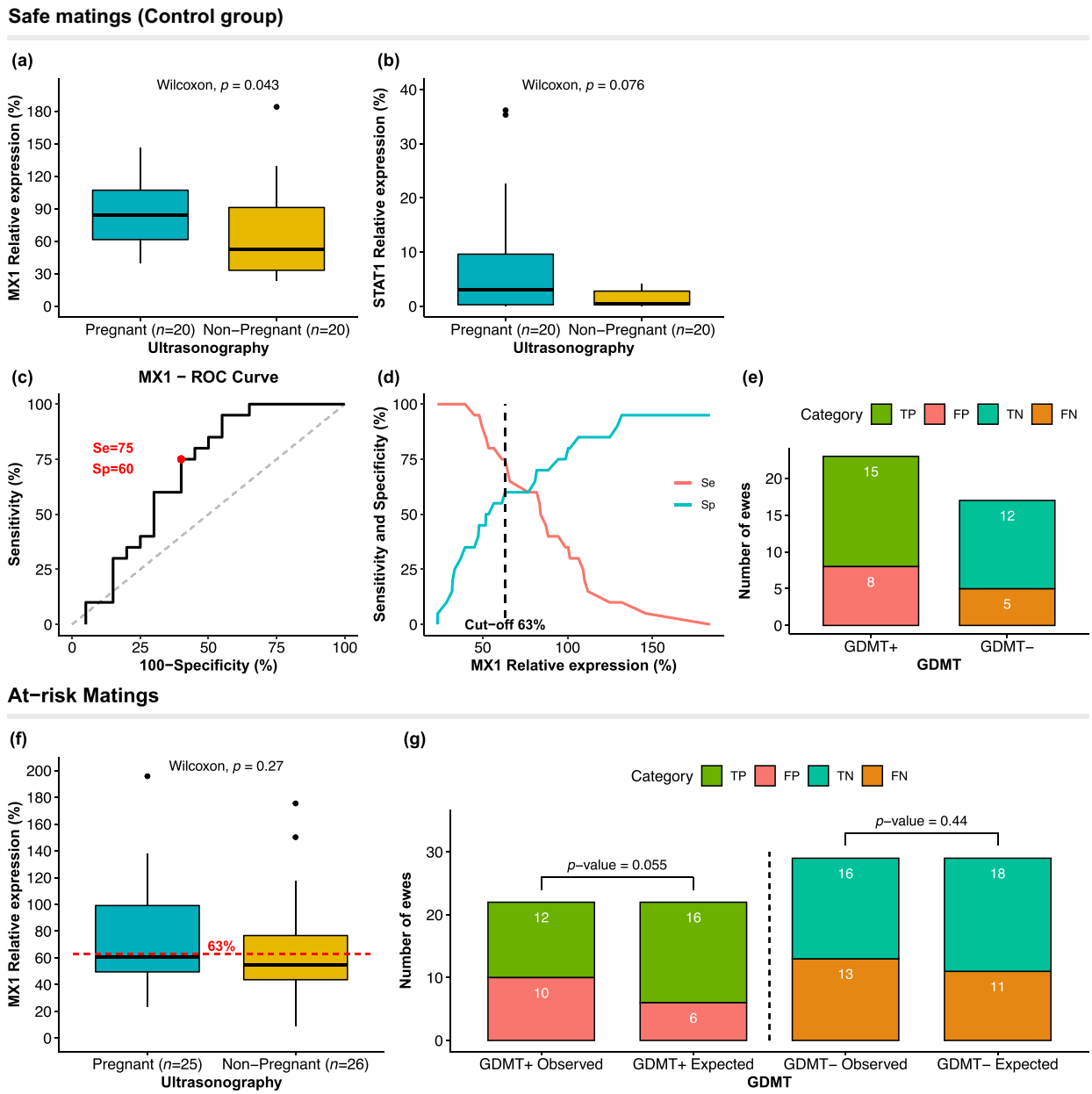

FIGURE 5 Gestation diagnosis molecular test (GDMT). Relative expression of *MX1* (a) and *STAT1* (b) mRNA at 15 days post-IA in pregnant and non-pregnant ewes of the safe matings group, as determined by ultrasonography (control group). (c) ROC curve based on sensitivity (Se) and specificity (Sp) of GDMT using *MX1* relative expression in the control group. (d) Determination of *MX1* relative expression cut-off value (63%) associated with Se = 75% and Sp = 60% in the control group using the ROC01 method. (e) Distribution of pregnant and non-pregnant ewes according to GDMT based on *MX1* relative expression in the control group with TP (true positive), FP (false positive), TN (true negative), and FN (false negative) indicated. (f) *MX1* relative expression in pregnant and non-pregnant ewes in at-risk matings. (g) Distribution of observed and expected pregnant and non-pregnant ewes according to GDMT based on *MX1* relative expression with a cut-off value of 63% in at-risk matings. The expected numbers were calculated based on prevalence (Pr = 60.3%), positive predictive value (PPV = 74%) and negative predictive value (NPV = 61%). Differences between observed and expected numbers from GDMT (+ and −) were assessed by a Chi-squared test.

et al., 2016). The GDMT enabled us to identify early fetal losses, providing a valuable tool for predicting gestation status at day 15. Nonetheless, its current application in farm settings is constrained. In our study, only 66% of RNA samples were usable. The primary issue was RNA degradation owing to inadequate storage of blood samples collected on farms. To mitigate this, samples should be immediately frozen in a suitable buffer that inactivates RNases.

The AUC of a ROC curve effectively assesses the diagnostic power of a test (Janssens & Martens, 2020). Our study yielded an AUC of 0.687, considered acceptable (Swets, 1988). To determine the optimal cut-off value for relative *MX1* expression, we employed the ROC01 method rather than the commonly used Youden index (Perkins & Schisterman, 2006). The ROC01 method prioritizes sensitivity, crucial for reliably identifying ewes with positive GDMT results who are pregnant at days

Gene	AUC ± se	*p*-Value	Se (%)	Sp (%)	Cut-off (%)[a]	Pr (%)	PPV (%)[b]	NPV (%)[b]
MX1	0.687 ± 0.088	0.033	75	60	63	60.3	74	61

Abbreviations: AUC, area under the curve; se, standard error; *MX1*, Myxovirus (influenza virus) resistance 1; NPV, negative predictive value; Pr, prevalence, e.g. artificial insemination success in Manech Tête Rousse dairy sheep; PPV, positive predictive value; Se, sensitivity; Sp, specificity.

[a]ROC01 method (minimizes distance between ROC curve and point (100 − Sp = 0, Se = 100)) was used for optimal cut-off.

[b]PPV and NPV were calculated as described in Section 2 taking into account Se, Sp, Pr.

TABLE 4 Receiver operating characteristic (ROC) curve analysis parameter for the gestation molecular diagnosis test based on *MX1* mRNA level expression.

45–60. Initially, the GDMT incorporated expression analysis of two ISGs, *MX1* and *STAT1*. However, only relative *MX1* mRNA expression differed significantly between pregnant and non-pregnant ewes in the control group. Interferon-tau, produced primarily by the conceptus between days 14 and 17 post-AI, probably drives this difference. While our blood samples were collected at day 15, *STAT1* mRNA expression peaks at days 16–17 (Mauffré et al., 2016). Additionally, *MX1* mRNA overexpression can persist until day 21 in peripheral blood mononuclear cells (Yankey et al., 2001). Consequently, day 17 may be a more optimal blood sampling time for future analyses.

The SLC superfamily, comprising over 400 members, is linked to numerous Mendelian disorders in humans, primarily recessive neurometabolic diseases (Schaller & Lauschke, 2019). Similarly, domestic animals, particularly dogs and cattle, exhibit a growing list of causative variants in SLC genes (omia.org). For example, a missense mutation (rs438228855) in the *SLC35A3* gene identified through reverse genetic screening is responsible for Complex Vertebral Malformation syndrome in cattle. While live homozygous calves with Complex Vertebral Malformation have been observed, many succumb during gestation (Thomsen et al., 2006; VanRaden et al., 2011). In Montbéliarde cattle, a homozygous haplotype (MH2) associated with decreased calving rate was found to carry a stop codon mutation (p.R12*) in the *SLC37A2* gene (Fritz et al., 2013). This gene is also suspected to harbor mutations causing craniomandibular osteopathy in dogs (Hytönen et al., 2016; Letko et al., 2020). Interestingly, the *SLC45A2* gene influences coat color across diverse species: horses (Mariat et al., 2003), chickens and quails (Gunnarsson et al., 2007), medaka fish (Fukamachi et al., 2008), tigers (Xu et al., 2013), gorillas (Prado-Martinez et al., 2013), dogs (Winkler et al., 2014) and cattle (Rothammer et al., 2017). In sheep, a single reported variant, a 1 bp deletion in *SLC13A1*, causes chondrodysplasia (Zhao et al., 2012). Whole-genome sequencing analysis, similar to the approach used in humans (Charlier et al., 2016), could identify loss-of-function mutations in *SLC* genes of livestock. Ensembl variation data for 453 sheep from 38 breeds (https://www.sheephapmap.org/) revealed three potential loss-of-function variants in *SLC33A1*: a stop codon mutation (p.Trp546*) in Composite and D'Man breeds, and two missense mutations (p.Asp172Tyr; p.Tyr380Cys) in Norduz sheep. These variants probably disrupt SLC33A1 function, warranting further investigation into their potential impact on stillbirth rates, similar to the observations with the *SLC33A1_dupG* variant in MTR sheep.

CONCLUSION

Reverse genetic strategies offer a powerful tool to identify embryonic lethal variants and uncover genetic diseases potentially misattributed to environmental factors in lambing deaths. In this study, we identified a single base pair duplication in the *SLC33A1* gene harbored by the homozygous deficient haplotype MTRDHH2 in MTR dairy sheep. SLC33A1 encodes an essential acetyl-CoA transporter, crucial for cellular metabolism. This explains the mutation's impact on various developmental stages, from fetal to perinatal. The link between this mutation and decreased AI success and lamb mortality highlights the importance of managing the *SLC33A1_dupG* variant in MTR dairy sheep breeding to improve fertility rates and lamb survival.

AUTHOR CONTRIBUTIONS

Maxime Ben Braiek: Data curation; formal analysis; investigation; methodology; software; validation; visualization; writing – original draft; writing – review and editing. **Soline Szymczak:** Formal analysis; investigation; methodology. **Céline André:** Investigation; resources. **Philippe Bardou:** Formal analysis; methodology; software. **Francis Fidelle:** Investigation; resources. **Itsasne Granado-Tajada:** Investigation; resources. **Florence Plisson-Petit:** Investigation; resources. **Julien Sarry:** Formal analysis; investigation; resources. **Florent Woloszyn:** Investigation; resources. **Stéphane Fabre:** Conceptualization; formal analysis; funding acquisition; investigation; project administration; supervision; writing – review and editing.

ACKNOWLEDGEMENTS

We are grateful to the genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, https://doi.org/10.15454/1.5572369328961167E12) for providing help and/or computing and/or storage resources. The authors acknowledge the breeding confederations

CDEO (Centre Départemental de l'Elevage Ovin) and CONFELAC (Confederación de Asociaciones de Criadores de Ovino de Razas Latxa y Carranzana) for providing access to private genomic data and/or biological samples. We thank also the breeders involved in the project.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The WGS data used in this study are publicly available, EMBL-EBI accession numbers are described in Table S1.

ETHICS STATEMENT

The experimental procedures on animals (blood sampling, ear biopsies) were approved by the French Ministry of Teaching and Scientific Research and local ethical committee C2EA-115 (Science and Animal Health) in accordance with the European Union Directive 2010/63/EU on the protection of animals used for scientific purposes (approval numbers APAFIS#30615–2021 032 318 0 54 889 v5).

ORCID

Maxime Ben Braiek https://orcid.org/0000-0003-4770-0867
Itsasne Granado-Tajada https://orcid.org/0000-0002-6557-1467
Stéphane Fabre https://orcid.org/0000-0001-7350-9500

REFERENCES

Abdalla, E.A., Id-Lahoucine, S., Cánovas, A., Casellas, J., Schenkel, F.S., Wood, B.J. et al. (2020) Discovering lethal alleles across the Turkey genome using a transmission ratio distortion approach. *Animal Genetics*, 51, 876–889. Available from: https://doi.org/10.1111/age.13003

Astruc, J.-M., Baloche, G., Buisson, D., Labatut, J., Lagriffoul, G., Larroque, H. et al. (2016) Genomic selection in French dairy sheep. *INRAE Production Animal*, 29, 41–55.

Bazer, F.W. (2013) Pregnancy recognition signaling mechanisms in ruminants and pigs. *Journal of Animal Science and Biotechnology*, 4, 23. Available from: https://doi.org/10.1186/2049-1891-4-23

Ben Braiek, M., Fabre, S., Hozé, C., Astruc, J.-M. & Moreno-Romieux, C. (2021) Identification of homozygous haplotypes carrying putative recessive lethal mutations that compromise fertility traits in French Lacaune dairy sheep. *Genetics, Selection, Evolution*, 53, 41. Available from: https://doi.org/10.1186/s12711-021-00634-1

Ben Braiek, M., Moreno-Romieux, C., Allain, C., Bardou, P., Bordes, A., Debat, F. et al. (2022) A nonsense variant in CCDC65 gene causes respiratory failure associated with increased lamb mortality in French Lacaune dairy sheep. *Genes*, 13, 45. Available from: https://doi.org/10.3390/genes13010045

Ben Braiek, M., Moreno-Romieux, C., André, C., Astruc, J.-M., Bardou, P., Bordes, A. et al. (2024) Searching for homozygous haplotype deficiency in Manech Tête Rousse dairy sheep revealed a nonsense variant in the MMUT gene affecting newborn lamb viability. *Genetics, Selection, Evolution*, 56, 16. Available from: https://doi.org/10.1186/s12711-024-00886-7

Bindon, B.M. (1971) Systematic study of preimplantation stages of pregnancy in the sheep. *Australian Journal of Biological Sciences*, 24, 131–148. Available from: https://doi.org/10.1071/bi9710131

Charlier, C., Coppieters, W., Rollin, F., Desmecht, D., Agerholm, J.S., Cambisano, N. et al. (2008) Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nature Genetics*, 40, 449–454. Available from: https://doi.org/10.1038/ng.96

Charlier, C., Li, W., Harland, C., Littlejohn, M., Coppieters, W., Creagh, F. et al. (2016) NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. *Genome Research*, 26, 1333–1341. Available from: https://doi.org/10.1101/gr.207076.116

Cingolani, P., Patel, V.M., Coon, M., Nguyen, T., Land, S.J., Ruden, D.M. et al. (2012) Using drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program. *SnpSift Frontiers in Genetics*, 3, 35. Available from: https://doi.org/10.3389/fgene.2012.00035

Clark, E.L., Bush, S.J., McCulloch, M.E.B., Farquhar, I.L., Young, R., Lefevre, L. et al. (2017) A high resolution atlas of gene expression in the domestic sheep (*Ovis aries*). *PLoS Genetics*, 13, e1006997. Available from: https://doi.org/10.1371/journal.pgen.1006997

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A. et al. (2011) The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158. Available from: https://doi.org/10.1093/bioinformatics/btr330

Derks, M.F.L., Megens, H.-J., Bosse, M., Lopes, M.S., Harlizius, B. & Groenen, M.A.M. (2017) A systematic survey to identify lethal recessive variation in highly managed pig populations. *BMC Genomics*, 18, 858. Available from: https://doi.org/10.1186/s12864-017-4278-1

Derks, M.F.L., Megens, H.-J., Bosse, M., Visscher, J., Peeters, K., Bink, M.C.A.M. et al. (2018) A survey of functional genomic variation in domesticated chickens. *Genetics, Selection, Evolution*, 50, 17. Available from: https://doi.org/10.1186/s12711-018-0390-1

Diskin, M. & Morris, D. (2008) Embryonic and early Foetal losses in cattle and other ruminants. *Reproduction in Domestic Animals*, 43, 260–267. Available from: https://doi.org/10.1111/j.1439-0531.2008.01171.x

Dixon, A.B., Knights, M., Winkler, J.L., Marsh, D.J., Pate, J.L., Wilson, M.E. et al. (2007) Patterns of late embryonic and fetal mortality and association with several factors in sheep1. *Journal of Animal Science*, 85, 1274–1284. Available from: https://doi.org/10.2527/jas.2006-129

Eggen, A. (2012) The development and application of genomic selection as a new breeding paradigm. *Animal Frontiers*, 2, 10–15. Available from: https://doi.org/10.2527/af.2011-0027

Fritz, S., Capitan, A., Djari, A., Rodriguez, S.C., Barbat, A., Baur, A. et al. (2013) Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. *PLoS One*, 8, e65550. Available from: https://doi.org/10.1371/journal.pone.0065550

Fukamachi, S., Kinoshita, M., Tsujimura, T., Shimada, A., Oda, S., Shima, A. et al. (2008) Rescue from Oculocutaneous albinism type 4 using Medaka slc45a2 cDNA driven by its own promoter. *Genetics*, 178, 761–769. Available from: https://doi.org/10.1534/genetics.107.073387

Georges, M., Charlier, C. & Hayes, B. (2019) Harnessing genomic information for livestock improvement. *Nature Reviews. Genetics*, 20, 135–156. Available from: https://doi.org/10.1038/s41576-018-0082-2

Goksuluk, D., Korkmaz, S., Zararsiz, G. & Karaagaoglu, A. (2016) easyROC: an interactive web-tool for ROC curve analysis using R language environment. *R Journal*, 8, 213. Available from: https://doi.org/10.32614/RJ-2016-042

Gunnarsson, U., Hellström, A.R., Tixier-Boichard, M., Minvielle, F., Bed'hom, B., Ito, S. et al. (2007) Mutations in SLC45A2 cause plumage color variation in chicken and Japanese quail. *Genetics*,

175, 867–877. Available from: https://doi.org/10.1534/genetics.106.063107

Huppke, P., Brendel, C., Kalscheuer, V., Korenke, G.C., Marquardt, I., Freisinger, P. et al. (2012) Mutations in SLC33A1 cause a lethal autosomal-recessive disorder with congenital cataracts, hearing loss, and low serum copper and Ceruloplasmin. *American Journal of Human Genetics*, 90, 61–68. Available from: https://doi.org/10.1016/j.ajhg.2011.11.030

Hytönen, M.K., Arumilli, M., Lappalainen, A.K., Owczarek-Lipska, M., Jagannathan, V., Hundi, S. et al. (2016) Molecular characterization of three canine models of human rare bone diseases: Caffey, van den Ende-Gupta, and Raine syndromes. *PLoS Genetics*, 12, e1006037. Available from: https://doi.org/10.1371/journal.pgen.1006037

Janssens, A.C.J.W. & Martens, F.K. (2020) Reflection on modern methods: revisiting the area under the ROC curve. *International Journal of Epidemiology*, 49, 1397–1403. Available from: https://doi.org/10.1093/ije/dyz274

Jenko, J., McClure, M.C., Matthews, D., McClure, J., Johnsson, M., Gorjanc, G. et al. (2019) Analysis of a large dataset reveals haplotypes carrying putatively recessive lethal and semi-lethal alleles with pleiotropic effects on economically important traits in beef cattle. *Genetics Selection Evolution*, 51, 9. Available from: https://doi.org/10.1186/s12711-019-0452-z

Jonas, M.C., Pehar, M. & Puglielli, L. (2010) AT-1 is the ER membrane acetyl-CoA transporter and is essential for cell viability. *Journal of Cell Science*, 123, 3378–3388. Available from: https://doi.org/10.1242/jcs.068841

Karen, A., Beckers, J.-F., Sulon, J., de Sousa, N.M., Szabados, K., Reczigel, J. et al. (2003) Early pregnancy diagnosis in sheep by progesterone and pregnancy-associated glycoprotein tests. *Theriogenology*, 59, 1941–1948. Available from: https://doi.org/10.1016/S0093-691X(02)01289-X

Lander, E.S. & Botstein, D. (1987) Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science*, 236, 1567–1570. Available from: https://doi.org/10.1126/science.2884728

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J. et al. (2001) Initial sequencing and analysis of the human genome. *Nature*, 409, 860–921. Available from: https://doi.org/10.1038/35057062

Letko, A., Leuthard, F., Jagannathan, V., Corlazzoli, D., Matiasek, K., Schweizer, D. et al. (2020) Whole genome sequencing indicates heterogeneity of Hyperostotic disorders in dogs. *Genes*, 11, 163. Available from: https://doi.org/10.3390/genes11020163

Lin, P., Li, J., Liu, Q., Mao, F., Li, J., Qiu, R. et al. (2008) A missense mutation in SLC33A1, which encodes the acetyl-CoA transporter, causes autosomal-dominant spastic paraplegia (SPG42). *American Journal of Human Genetics*, 83, 752–759. Available from: https://doi.org/10.1016/j.ajhg.2008.11.003

Liu, P., Jiang, B., Ma, J., Lin, P., Zhang, Y., Shao, C. et al. (2017) S113R mutation in SLC33A1 leads to neurodegeneration and augmented BMP signaling in a mouse model. *Disease Models & Mechanisms*, 10, 53–62. Available from: https://doi.org/10.1242/dmm.026880

MacArthur, D.G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K. et al. (2012) A systematic survey of loss-of-function variants in human protein-coding genes. *Science*, 335, 823–828. Available from: https://doi.org/10.1126/science.1215040

Mao, F., Li, Z., Zhao, B., Lin, P., Liu, P., Zhai, M. et al. (2015) Identification and functional analysis of a SLC33A1: c.339T>G (p.Ser113Arg) variant in the original SPG42 family. *Human Mutation*, 36, 240–249. Available from: https://doi.org/10.1002/humu.22732

Mariat, D., Taourit, S. & Guérin, G. (2003) A mutation in the MATP gene causes the cream coat colour in the horse. *Genetics, Selection, Evolution*, 35, 119–133. Available from: https://doi.org/10.1186/1297-9686-35-1-119

Mauffré, V., Grimard, B., Eozenou, C., Inghels, S., Silva, L., Giraud-Delville, C. et al. (2016) Interferon stimulated genes as peripheral diagnostic markers of early pregnancy in sheep: a critical assessment. *Animal*, 10, 1856–1863. Available from: https://doi.org/10.1017/S175173111600077X

Ott, T.L. & Gifford, C.A. (2010) Effects of early conceptus signals on circulating immune cells: lessons from domestic ruminants. *American Journal of Reproductive Immunology*, 64, 245–254. Available from: https://doi.org/10.1111/j.1600-0897.2010.00912.x

Peng, Y., Li, M., Clarkson, B.D., Pehar, M., Lao, P.J., Hillmer, A.T. et al. (2014) Deficient import of acetyl-CoA into the ER lumen causes neurodegeneration and propensity to infections, inflammation, and cancer. *The Journal of Neuroscience*, 34, 6772–6789. Available from: https://doi.org/10.1523/JNEUROSCI.0077-14.2014

Perkins, N.J. & Schisterman, E.F. (2006) The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *American Journal of Epidemiology*, 163(7), 670–675. https://doi.org/10.1093/aje/kwj063

Prado-Martinez, J., Hernando-Herraez, I., Lorente-Galdos, B., Dabad, M., Ramirez, O., Baeza-Delgado, C. et al. (2013) The genome sequencing of an albino Western lowland gorilla reveals inbreeding in the wild. *BMC Genomics*, 14, 363. Available from: https://doi.org/10.1186/1471-2164-14-363

Rickard, J.P., Ryan, G., Hall, E., de Graaf, S.P. & Hermes, R. (2017) Using transrectal ultrasound to examine the effect of exogenous progesterone on early embryonic loss in sheep. *PLoS One*, 12, e0183659. Available from: https://doi.org/10.1371/journal.pone.0183659

Rochus, C.M., Tortereau, F., Plisson-Petit, F., Restoux, G., Moreno-Romieux, C., Tosser-Klopp, G. et al. (2018) Revealing the selection history of adaptive loci using genome-wide scans for selection: an example from domestic sheep. *BMC Genomics*, 19, 71. Available from: https://doi.org/10.1186/s12864-018-4447-x

Rothammer, S., Kunz, E., Seichter, D., Krebs, S., Wassertheurer, M., Fries, R. et al. (2017) Detection of two non-synonymous SNPs in SLC45A2 on BTA20 as candidate causal mutations for oculocutaneous albinism in Braunvieh cattle. *Genetics, Selection, Evolution*, 49, 73. Available from: https://doi.org/10.1186/s12711-017-0349-7

Rupp, R., Mucha, S., Larroque, H., McEwan, J. & Conington, J. (2016) Genomic application in sheep and goat breeding. *Animal Frontiers*, 6, 39–44. Available from: https://doi.org/10.2527/af.2016-0006

Schaller, L. & Lauschke, V.M. (2019) The genetic landscape of the human solute carrier (SLC) transporter superfamily. *Human Genetics*, 138, 1359–1377. Available from: https://doi.org/10.1007/s00439-019-02081-x

Spencer, T.E., Sandra, O. & Wolf, E. (2008) Genes involved in conceptus–endometrial interactions in ruminants: insights from reductionism and thoughts on holistic approaches. *Reproduction*, 135, 165–179. Available from: https://doi.org/10.1530/REP-07-0327

Swets, J.A. (1988) Measuring the accuracy of diagnostic systems. *Science*, 240, 1285–1293. Available from: https://doi.org/10.1126/science.3287615

Thomsen, B., Horn, P., Panitz, F., Bendixen, E., Petersen, A.H., Holm, L.-E. et al. (2006) A missense mutation in the bovine SLC35A3 gene, encoding a UDP-N-acetylglucosamine transporter, causes complex vertebral malformation. *Genome Research*, 16, 97–105. Available from: https://doi.org/10.1101/gr.3690506

Thorvaldsdóttir, H., Robinson, J.T. & Mesirov, J.P. (2013) Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14, 178–192. Available from: https://doi.org/10.1093/bib/bbs017

Todd, E.T., Thomson, P.C., Hamilton, N.A., Ang, R.A., Lindgren, G., Viklund, Å. et al. (2020) A genome-wide scan for candidate lethal variants in thoroughbred horses. *Scientific Reports*, 10, 13153. Available from: https://doi.org/10.1038/s41598-020-68946-8

Tsuchiya, Y., Pham, U., Hu, W., Ohnuma, S. & Gout, I. (2014) Changes in acetyl CoA levels during the early embryonic development of Xenopus laevis. *PLoS One*, 9, e97693. Available from: https://doi.org/10.1371/journal.pone.0097693

Uffelmann, E., Huang, Q.Q., Munung, N.S., de Vries, J., Okada, Y., Martin, A.R. et al. (2021) Genome-wide association studies. *Nature Reviews Methods Primers*, 1, 1–21. Available from: https://doi.org/10.1038/s43586-021-00056-9

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. & Leunissen, J.A.M. (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*, 35, W71–W74. Available from: https://doi.org/10.1093/nar/gkm306

VanRaden, P.M., Olson, K.M., Null, D.J. & Hutchison, J.L. (2011) Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *Journal of Dairy Science*, 94, 6153–6161. Available from: https://doi.org/10.3168/jds.2011-4624

Wilmut, I., Sales, D.I. & Ashworth, C.J. (1986) Maternal and embryonic factors associated with prenatal loss in mammals. *Reproduction*, 76, 851–864. Available from: https://doi.org/10.1530/jrf.0.0760851

Winkler, P.A., Gornik, K.R., Ramsey, D.T., Dubielzig, R.R., Venta, P.J., Petersen-Jones, S.M. et al. (2014) A partial gene deletion of SLC45A2 causes Oculocutaneous albinism in Doberman pinscher dogs. *PLoS One*, 9, e92127. Available from: https://doi.org/10.1371/journal.pone.0092127

Xu, X., Dong, G.-X., Hu, X.-S., Miao, L., Zhang, X.-L., Zhang, D.-L. et al. (2013) The genetic basis of white tigers. *Current Biology*, 23, 1031–1035. Available from: https://doi.org/10.1016/j.cub.2013.04.054

Yankey, S.J., Hicks, B.A., Carnahan, K.G., Assiri, A.M., Sinor, S.J., Kodali, K., Stellflug, J.N. & Ott, T.L. (2001) Expression of the antiviral protein Mx in peripheral blood mononuclear cells of pregnant and bred, non-pregnant ewes. *Journal of Endocrinology*, 170, R7–11. https://doi.org/10.1677/joe.0.170r007

Zhao, X., Onteru, S.K., Piripi, S., Thompson, K.G., Blair, H.T., Garrick, D.J. et al. (2012) In a shake of a lamb's tail: using genomics to unravel a cause of chondrodysplasia in Texel sheep. *Animal Genetics*, 43, 9–18. Available from: https://doi.org/10.1111/j.1365-2052.2011.02304.x

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: Ben Braiek, M., Szymczak, S., André, C., Bardou, P., Fidelle, F., Granado-Tajada, I. et al. (2024) A single base pair duplication in the *SLC33A1* gene is associated with fetal losses and neonatal lethality in Manech Tête Rousse dairy sheep. *Animal Genetics*, 55, 644–657. Available from: https://doi.org/10.1111/age.13459