

## Homozygous haplotype deficiency in Manech Tête Rousse dairy sheep revealed a nonsense variant in MMUT gene affecting newborn lamb viability

Maxime Ben Braiek, Carole Moreno-Romieux, Céline André, Jean-Michel Astruc, Philippe Bardou, Arnaud Bordes, Frédéric Debat, Francis Fidelle, Itsasne Granado Tajada, Chris Hozé, et al.

### ▶ To cite this version:

Maxime Ben Braiek, Carole Moreno-Romieux, Céline André, Jean-Michel Astruc, Philippe Bardou, et al.. Homozygous haplotype deficiency in Manech Tête Rousse dairy sheep revealed a nonsense variant in MMUT gene affecting newborn lamb viability. 2023. hal-04158620

## HAL Id: hal-04158620 https://hal.inrae.fr/hal-04158620

Preprint submitted on 11 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Homozygous haplotype deficiency in Manech Tête Rousse dairy sheep revealed a nonsense variant in *MMUT* gene

- 3 affecting newborn lamb viability
- 4 Maxime Ben Braiek<sup>1</sup>, Carole Moreno-Romieux<sup>1\$</sup>, Céline André<sup>2</sup>, Jean-Michel Astruc<sup>3</sup>,
- 5 Philippe Bardou<sup>4</sup>, Arnaud Bordes<sup>1</sup>, Frédéric Debat<sup>1</sup>, Francis Fidelle<sup>2</sup>, Itsasne Granado–Tajada<sup>5</sup>,
- 6 Chris Hozé<sup>6</sup>, Florence Plisson-Petit<sup>1</sup>, François Rivemale<sup>1</sup>, Julien Sarry<sup>1</sup>, Némuel Tadi<sup>1</sup>, Florent
- 7 Woloszyn<sup>1</sup> and Stéphane Fabre<sup>1\*</sup>
- 8 <sup>1</sup>GenPhySE, Université de Toulouse, INRAE, ENVT, 31326 Castanet-Tolosan, France
- 9 <sup>2</sup>CDEO, Quartier Ahetzia, 64130 Ordiarp, France
- <sup>10</sup> <sup>3</sup>Institut de l'Elevage, 24 chemin de Borde-Rouge, 31321 Castanet-Tolosan, France
- <sup>4</sup>Sigenae, INRAE, 31326 Castanet-Tolosan, France
- <sup>5</sup>Department of Animal Production, NEIKER-BRTA Basque Institute of Agricultural Research
   and Development, Agrifood Campus of Arkaute s/n, E-01080 Arkaute, Spain
- <sup>6</sup>Eliance, 149 rue de Bercy, 75595 Paris, France
- 15 *\$ in memoriam*
- 16 \*Corresponding author
- 17 Email: stephane.fabre@inrae.fr (SF)
- 18 Maxime Ben Braiek, ORCID iD https://orcid.org/0000-0003-4770-0867
- 19 Itsasne Granado–Tajada, ORCID iD https://orcid.org/0000-0002-6557-1467
- 20 Stéphane Fabre, ORCID iD https://orcid.org/0000-0001-7350-9500

### 21 Abstract

22 Recessive deleterious variants are known to segregate in livestock populations as in human, and 23 some may cause lethality when homozygous. By scanning the genome of 6,845 Manech Tête 24 Rousse dairy sheep using phased 50k SNP genotypes and pedigree data, we searched for 25 deficiency in homozygous haplotype (DHH). Five Manech Tête Rousse deficient homozygous 26 haplotypes (MTRDHH1 to 5) were identified with a homozygous deficiency ranging from 84% 27 to 100%. These haplotypes are located on OAR1 (MTRDHH2 and 3), OAR10 (MTRDHH4), 28 OAR13 (MTRDHH5) and OAR20 (MTRDHH1), and have frequencies ranging from 7.8% to 29 16.6%. When comparing at-risk mating between DHH carriers to safe mating between non-30 carriers, two DHH (MTRDHH1 and 2) showed significant effects on decreasing artificial 31 insemination success and/or increasing stillbirth rate. We particularly investigated the 32 MTRDHH1 haplotype highly increasing stillbirth rate, and we identified a single nucleotide variant (SNV) inducing a premature stop codon (p.Gln409\*) in the MMUT gene 33 34 (methylmalonyl-CoA mutase) by using a whole genome sequencing (WGS) approach. We 35 generated homozygous lambs for the MMUT mutation by oriented mating, and most of them 36 died within the first 24h after birth without any obvious clinical defect. RT-qPCR and western 37 blotting performed on post-mortem liver and kidney biological samples showed a decreased 38 expression of MMUT mRNA in the liver and absence of a full-length MMUT protein in mutated 39 homozygous lambs. In parallel, MTRDHH4 and MTRDHH5 showed partial association with variants in RXFP2 and ASIP genes, respectively, already known to control horned/polled and 40 41 coat color phenotypes in sheep, two morphological traits accounting in the MTR breed standard. 42 Further investigations are needed to identified the supposed recessive deleterious variant hosted by MTRDHH2 and MTRDHH3. Anyway, an appropriate management of these 43 44 haplotypes/variants in the MTR dairy sheep selection program should increase the overall 45 fertility and lamb survival.

## 46 Author Summary

47 In this article, we used reverse genetics screen in ovine using large genotype data available in 48 the framework of genomic selection program in Manech Tête Rousse dairy sheep. We identified 49 five genomic regions with a highly significant deficit in homozygous animal. These regions are 50 thus supposed to host recessive deleterious mutations. In one of these genomic regions, we 51 identified a nonsense mutation in MMUT that alters the functioning of this essential gene of cell 52 metabolism, causing perinatal mortality of homozygous lambs. In this work, we also identified 53 other regions possibly associated with morphological appearance part of the breed standard 54 such as polledness and coat color. Increasing knowledge in these genomic regions will help the 55 future genetic management of the Manech Tête Rousse breed, particularly to reduce lamb 56 mortality.

## 57 Introduction

58 In livestock, genetic selection has largely improved production traits over the past 59 decades, but the last one has seen the emergence of new technological tools allowing to 60 implement the genomic selection that further enhanced the genetic progress [1]. The availability 61 of high-density single nucleotide polymorphism (SNP) chip and the improvement of knowledge 62 of genomes (genome assembly and gene annotations) have allowed to fine-map genomic 63 regions and identify causal variants associated with production traits [1–4]. Despite a successful 64 selection on these traits, undesired decline of fertility was observed [5]. Although the 65 environment explains a large part of performance in ruminants ), genetic studies have made it possible to correct this trend and improve fertility while its heritability is less than 0.05 [6–8]. 66

67 These studies have shown Mendelian monogenic disorders as one of the causes of fertility68 failure.

69 Nowadays two main approaches were broadly developed to identify recessive 70 deleterious variants. The first approach is a "top-down" strategy based on case-control analysis 71 performing genome-wide association [9] when biological samples from affected animals are 72 available. In this method, distinctive phenotypes between non-affected and affected animals are 73 essential. Subsequently, homozygosity mapping approach could be performed to detect 74 homozygous regions in affected animals supposed to host the causal variant, further determined 75 by whole genome sequencing (WGS) data [10,11]. In livestock, Charlier et al. [10] have used 76 this approach for the first time and successfully detected three causal variants in cattle breeds 77 located in ATP2A1, SLC6A5 and ABCA12 genes responsible for "congenital muscular dystony" 78 types 1 (OMIA 001450-9913) and 2 (OMIA 001451-9913) and "ichthyosis fetalis" (OMIA 79 002238-9913), respectively. However, this approach showed some limits when biological samples and descriptive phenotypes are not available. To raise this drawback, a second 80 81 approach called "bottom-up" or reverse genetic screen strategy was developed to specifically 82 identify recessive lethal variants. This strategy, initially developed by VanRaden et al. [12], is 83 based on the exploitation of large number of genotyped animals easily available from genomic 84 selection datasets to detect haplotypes showing deficit in homozygous animals, with a 85 significant deviation from the Hardy-Weinberg equilibrium. Initially, this method was 86 developed to detect embryonic lethal variant but the generalization of the method can also fine-87 map deleterious variants leading to neonatal or juvenile lethality and morphological disorders. 88 This reverse genetic screen has successfully identified numerous deficient homozygous 89 haplotypes in several species, cattle [12–27], pigs [28,29], chicken [30], turkey [31] and horses 90 [32], and the whole genome sequencing that followed has revealed the associated causative 91 variants hosted by these regions [13,14,18,19,21,24–27,33–41]. We recently, and for the first time, validated this approach in sheep with the identification of 8 independent deficient
homozygous haplotypes in the Lacaune dairy breed [42]. Thereafter, focusing on the Lacaune
Deficient Homozygous Haplotype 6 (LDHH6, OMIA 002342-9940), we identified a nonsense
variant in *CCDC65* gene causing juvenile mortality associated with respiratory distress when
homozygous [43].

In the present work, still leveraging genomic selection data, we searched for lethal variants using a reverse genetic screen in Manech Tête Rousse (MTR) dairy sheep. This breed is raised in the French Basque Country and represents the second most important breed by its population size (~450,000 ewes) in France [44]. Genetic variability is well controlled in MTR breed with an increase by +0.4% of inbreeding per generation over 1999-2009. The effective population size ranges from 110 to 200 according to the estimation methods [44–46]. As for other dairy sheep breeds in France, genomic selection in MTR was implemented in 2017 [47].

The aim of this study was to identify deficient homozygous haplotypes by reverse genetic screen using large amount of genotyping data available in MTR dairy sheep, to test the hypothesis of negative impacts on fertility traits in at-risk matings, to propose relevant candidate genes located in these regions that could host recessive deleterious variants. We particularly focus on one region to identify the associated causal variant from WGS data and manage atrisk mating between carriers to determine the associated phenotype.

## 110 **Results**

#### 111 Identification of deficient homozygous haplotypes in Manech Tête Rousse dairy sheep

Using a reverse genetic screen strategy based on 5,271 genotyped animals belonging to trios, we have detected 150 highly significant Homozygous Haplotype Deficiency (HHD) of 20 SNP markers (listed in S1A Table). These 150 HHD were clustered to five independent regions called "Manech Tête Rousse Deficient Homozygous Haplotype" (MTRDHH). Three

116	haplotypes showed a total deficit in homozygous animals (MTRDHH1, 2 and 3), whereas two
117	haplotypes, MTRDHH4 and 5, only showed a partial deficit (84% and 91%) with 1 and 8
118	homozygous animals genotyped while 11 and 49 were expected, respectively (Table 1, Fig 1,
119	S1A Table). The complete description of MTRDHH SNP markers (SNP name, SNP allele and
120	position on sheep reference genomes Oar_v3.1, Oar_Rambouillet_v1.0 and ARS-
121	UI_Ramb_v2.0) is available in S1B Table. The different MTRDHH were located on OAR20
122	(MTRDHH1), OAR1 (MTRDHH2 and 3), OAR10 (MTRDHH4) and OAR13 (MTRDHH5),
123	and their length ranged from 1.1 to 4.6 Mb on Oar_rambouillet_v1.0. The observed frequencies
124	of heterozygous carriers were between 7.8% and 16.6%. MTRDHH2 and MTRDHH3 both
125	located on OAR1 were not in linkage disequilibrium. Consequently, the five MTRDHH
126	identified are likely to harbor five independent variants causal of the observed homozygous
127	deficiency.

Haplotype	OAR	<sup>a</sup> Number	<sup>b</sup> Position	<sup>c</sup> Heterozygous	Number of home		ozygotes	
		of markers	(Mb)	carrier frequency (%)	dExp	°Obs	Deficit (%)	Poisson P-value
MTRDHH1	20	32	23.0-25.0	9.7	13	0	100	2.9×10 <sup>-6</sup>
MTRDHH2	1	66	251.9-256.4	8.7	10	0	100	3.8×10 <sup>-5</sup>
MTRDHH3	1	39	103.8-106.6	7.8	9	0	100	9.6×10 <sup>-5</sup>
MTRDHH4	10	26	30.5-31.5	8.7	11	1	91	$1.5 \times 10^{-4}$
MTRDHH5	13	29	64.3-67.2	16.6	49	8	84	5.3×10 <sup>-13</sup>

128 Table 1. List of Manech Tête Rousse deficient homozygous haplotypes.

<sup>a</sup>MTRDHH haplotypes and SNP markers composition are details in S1 Table

130 <sup>b</sup>Position on ovine genome assembly Oar\_rambouillet\_v1.0

- 131 <sup>c</sup>Frequency of carriers in the entire genotyping population (n=6,845)
- 132 <sup>d</sup>Expected
- 133 <sup>e</sup>Observed

#### 134 Impact of MTRDHH on fertility traits

135 In order to identify a putative lethal effect of the five MTRDHH, two recorded fertility traits

136 were analyzed: artificial insemination success, a proxy for embryonic loss (AIS: 330,844

137 matings) and stillbirth rate associated with perinatal lethality (SBR: 201,637 matings) (Fig 2).

138 The average AIS of the population was 60.9%. When comparing at-risk and safe matings, only

139 MTRDHH2 showed a significant decrease of -3.3% of AIS (P= $3.5 \times 10^{-4}$ ) in at-risk matings.

140 The average SBR of the population was 7.5%. As described in Fig 2, MTRDHH1 and 2 showed 141 a huge increase in SBR with +7.5% ( $P=4.0\times10^{-24}$ ) and +4.3% ( $P=1.3\times10^{-6}$ ) in at-risk matings 142 compared to safe matings, respectively. The three other haplotypes showed no significant 143 impact on the fertility traits studied.

#### 144 **Pleiotropic effects of MTRDHH on milk production traits**

145 Six dairy traits are routinely included in genomic evaluation of the French dairy sheep. Thus, 146 standardized daughter yield deviation (sDYD) of five milk production traits (milk, fat and 147 protein yields, and fat and protein contents) and lactation somatic cell score (a proxy for udder 148 health) were compared between carrier and non-carrier rams for each of the 5 MTRDHH 149 evidenced (Fig 3). Among the five haplotypes, three were associated with significant effect on 150 sDYD. Daughters of MTRDHH2 carrier rams showed a significant increase in milk production (sDYD +0.06, P= $6.5 \times 10^{-3}$ ) but a decrease in protein content (sDYD -0.11, P= $4.7 \times 10^{-4}$ ). For 151 MTRDHH4, there was a significant increase in lactation somatic cell score (sDYD -0.13, 152  $P=2.4\times10^{-3}$ ), and daughters of MTRDHH5 carrier rams showed higher fat yield (sDYD -0.06, 153  $P=9.6\times10^{-3}$ ). Additionally, the total merit genomic index, named ISOLg, was extracted from 154 155 each male lamb of the 2021 genomic selection cohort to estimate the MTRDHH impact on the 156 genetic gain of the selected traits. No significant difference was observed on ISOLg between 157 heterozygous carrier and non-carrier lambs for each MTRDHH (S1 Fig).

#### 158 Evolution of the MTRDHH frequencies in the population

Since the implementation of genomic selection in 2017, all candidate rams coming from elite mating were genotyped on low density SNP chip at 7 days old, representating the genetic diversity disseminated by AI in the selection scheme. As shown in Fig 4, the frequencies of the MTRDHH heterozygous carriers were quite stable during the last five years around 6.8, 7.4, 9.3, 10.2 and 16.1% for MTRDHH1, 2, 3, 4 and 5, respectively. Nevertheless, we can notice a spectacular increase in the frequency of MTRDHH3 from 2.8% to 10.9% when comparing 2017 and 2018.

#### 166 Candidate genes located in MTRHH regions

167 Within the five MTRDHH genomic regions extended by 1Mb on each side, 408 protein coding 168 genes are annotated (S2 Table). When available, information on mouse phenotypes (including 169 lethal phenotypes) and association with mammalian genetic disorders were extracted for each 170 gene using MGI, IMPC, OMIM and OMIA databases. Among the 408 genes, we highlighted 171 64 genes involved in lethal phenotypes in knock-out mice, and 45 genes associated with human 172 genetic disorders. Twenty-three relevant candidate genes were identified by the intersection of 173 both information (Fig 5). In addition, 7 genes are known to be associated with genetic disorders or morphological traits in livestock (GJA5, ITGA10, ADAMTLS4, RXFP2, KIF3B, ASIP and 174 175 CEP250). Overall, these candidate genes are involved in essential functions such as 176 transcription (POLR3GL, SF3B4, PRPF3, ASXL1 and DNMT3B), cell division (POGZ, BRCA2, 177 PIGU and CEP250), basal metabolic processes (MMUT, SLC33A1, TARS2 and AHCY), cell 178 structure and signaling (CD2AP, PKHD1, GJA5, ITGA10, ECM1, GJA5, ITGA10, ECM1, 179 PRUNE1, RXFP2, KL, POFUT1, KIF3B, ASIP and GSS), and DNA/protein binding (TFAP2B 180 and *PEX11B*).

# 181 Known variants in *RXFP2* and *ASIP* genes are partially associated with MTRDHH4 and 182 MTRDHH5

183 The above list of candidate genes particularly highlighted *RXFP2* (within MTRDHH4) and 184 *ASIP* (within MTRDHH5) genes well-known to impact sheep morphological traits such as

8

185 polledness [48] and black coat color [49,50], respectively. Interestingly, horned female and 186 black males are not desired in accordance with the MTR breed standards. Using the 22 WGS 187 of MTR animals, we searched for the 1.8kb insertion in the 3'-UTR of RXFP2 associated with 188 polledness (OMIA 000483-9940, [48]) and the different variants affecting ASIP leading to 189 recessive black coat color (OMIA 000201-9940); OAR13:g.66,475,132\_66,475,136del and 190 g.66,474,980T>A, Oar\_rambouillet\_v1.0) [49,50]. S2 Fig shows the segregation of these 191 variant among the 22 sequenced animals and their relationship with the status of MTRDHH4 192 and 5, but with no obvious association. This being made on a reduced number of animals, we 193 specifically genotyped a larger set of animals (n=714 male lambs born in 2021) for the 1.8kb 194 insertion in RXFP2 and the two variants in ASIP. Using this cohort, we evidenced a partial 195 association between MTRDHH4 and the 1.8kb insertion in the 3'-UTR of RXFP2, none 196 heterozygous MTRDHH4/+ being Del/Del corresponding to the horned phenotype (Fig 6). 197 Concerning ASIP variants, MTRDHH5 was also partially associated with the 5pb deletion. 198 Indeed, 89% of the MTRDHH5 heterozygous carriers were heterozygous for the 5pb deletion 199 and none Ins/Ins animal was MTRDHH5 carrier. In contrast, we identified almost the same 200 proportion of genotypes for g.66,474,980T>A in MTRDHH5 carriers and non-carriers (+/+), 201 suggesting that MTRDHH5 was not linked to this polymorphism (Fig 6C).

#### 202 Identification of a nonsense variant in *MMUT* gene associated with MTRDHH1

With the most important impact on stillbirth rate increased by +7.5% in at-risk matings (Fig 2), we particularly focused on MTRDHH1 that may represent a putative recessive lethal haplotype. In order to identify the MTRDHH1 causal mutation, we have considered biallelic variants (SNP and InDels) for 100 ovine WGS containing the 22 Manech Tête Rousse dairy sheep, and among them two heterozygous carriers of the MTRDHH1 haplotype. Within the MTRDHH1 region extended by 1Mb on each side, 78,019 variants were called with a quality score >30, call rate

>95% and only four candidate variants had a perfect correlation  $(r^2=1)$  between biallelic variant 209 210 genotypes and MTRDHH1 status (Table 2, Fig 7A). Among those candidate variants, we 211 identified two small insertions, one intergenic single nucleotide variant (SNV), and one 212 nonsense (stop-gain) SNV located in the Methylmalonyl-CoA Mutase (MMUT) gene. This latter 213 SNV (NC\_040271.1: g.23,776,347G>A; XM\_004018875.4: c.1225C>T; Fig 7B, C) in MMUT 214 is predicted to create a premature stop codon at position 409 encoded by exon 6 215 (XP\_004018923.1:p.Gln409\*) whereas the full protein length is composed of 750 amino acids 216 (Fig 7D). The variant would disrupt the methylmalonic coenzyme-A mutase domain and would 217 result in the loss of the vitamin B12 binding domain.

218 Table 2. Candidate variants located in MTRDHH1.

Position	Ref/Alt	Quality score	Location Annotation	Functional Consequence <sup>a</sup>
23,436,234	G/GTCACA	385.8	Intergenic	Modifier
23,436,236	T/TTTGTG	385.8	Intergenic	Modifier
23,776,347	G/A	146.0	Exonic, <i>MMUT</i> (c.1225C>T)	High, stop-gain (p.Gln409*)
23,969,676	C/T	370.3	Intergenic	Modifier

<sup>a</sup>Variant annotation and effect predicted by SnpEff [51].

220 In order to validate the association between the MMUT variant and MTRDHH1, we 221 genotyped the cohort of male lambs born in 2021 (n=714) with a specific genotyping test for 222 the MMUT g.23,776,347G>A SNV. The A variant allele frequency was 3.8%. All these animals 223 have a known status at the MTRDHH1 locus and the contingency table indicates a clear association between the MTRDHH1 status and MMUT variant genotypes (Fig 8, Fischer's exact 224 225 test p<0.0001). However, 15 animals showed discrepancy between MMUT and MTRDHH1 226 genotypes, supposed to be in perfect linkage disequilibrium. A specific focus on haplotypes 227 carried by these animals in the MTRDHH1 region from marker 1 (\$75212.1) to marker 32 228 (OAR20\_24583511.1) showed that the 14 animals heterozygous for the variant exhibited 229 shorter recombinant versions of the MTRDHH1 haplotype (S3 Fig). Nonetheless, one animal

230 was heterozygous for MTRDHH1 but did not carry the *MMUT* variant.

#### 231 Occurrence of the *MMUT* SNV in an ovine diversity panel

An ovine diversity panel composed of 25 French sheep breeds, including MTR, and 3 Latxa Spanish sheep breeds related to MTR was used to genotype the *MMUT* g.23,776,347G>A SNV (Table 3). As expected, some MTR animals (n=5) from this panel were evidenced as heterozygous carriers, and the variant was also detected in one animal of the Spanish Latxa Cara Rubia population. All the other animals tested did not carry the polymorphism.

		Geno	type			Gene	otype
Breed	Total	G/G	A/G	Breed	Total	G/G	A/G
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)	30	30		Tarasconnaise (FR)	33	33	
Manech Tête Rousse (FR)	29	24	5	Texel (FR)	27	27	
				Total	851	845	6

Table 3. *MMUT* SNV genotype distribution from a DNA diversity panel of French (FR)
 and Spanish (ES) ovine breeds.

#### 239 Viability of homozygous lambs for the MMUT variant

- 240 To validate the impact of the *MMUT* A variant allele we have manage an oriented mating
- 241 between heterozygous carriers to generate homozygous lambs. Blood samples were collected
- from 181 MTR ewes, daughters of MTRDHH1 carrier sires, in 6 private farms. The MMUT
- 243 SNV specific genotyping identified 82 heterozygous ewes. Among these ewes, 73 were raised

244 in the 6 private farms (Experiment 1) and 9 were moved into an INRAE experimental farm 245 (Experiment 2). All ewes were artificially inseminated with fresh semen from MMUT 246 g.23776347G/A heterozygous rams. Forty-five days after AI, 44 and 5 were diagnosed as 247 pregnant by ultrasonography in Experiment 1 and 2, respectively. This corresponds to an AIS 248 of 59.8% in accordance with the average AIS of 60.9% determined previously in the whole 249 population. In experiment 1, only 37 among the 44 pregnant ewes were monitored after 250 gestation diagnosis and resulted in the birth of 59 lambs (mean prolificacy of 1.6, litter size 251 ranging from 1 to 3) with a gestation length between 139 and 159 days. In experiment 2, the 5 252 pregnant ewes gave birth to 13 lambs (mean prolificacy of 2.6, litter size ranging from 2 to 4) 253 with a gestation length between 151 and 157 days. No abortion during the five months of 254 gestation was observed. Finally, 72 lambs (52% males and 48% females) were born and an ear 255 punch was collected for genotyping the MMUT SNV (Table 4). The distribution of genotypes 256 did not differ between the two experiments (Fisher's exact test, p=0.686). In total, 21 lambs 257 were genotyped homozygous carriers (A/A), 29 heterozygous carriers (A/G) and 21 258 homozygous non-carriers (G/G). All lambs were monitored during the 0-30 days period until 259 weaning. Twenty-five lambs died during this period representing a huge mortality rate of 34.7% 260 (Fig 9). Contingency table between lamb genotypes (A/A, G/A, G/G) and viability (alive or 261 dead) indicated a higher mortality rate for homozygous A/A lambs (Table 4, Fisher's exact test, 262 p<0.001). Indeed, the A/A dead lambs accounted for 78% of the whole lamb mortality. The 263 death of A/A homozygous lambs occurred very soon after birth within the first 24 hours. 264 Clinical examination of dead lambs did not allow us to identify specific symptoms. Two 265 homozygous A/A lambs have passed the weaning age (around 4 weeks). Additionally, in 266 Experiment 2, the 13 lambs were weighted at birth (males:  $4.0 \pm 0.9$  kg, females:  $2.9 \pm 1.4$  kg) 267 and A/A lambs had significantly lower birth weight compared to the other genotypes, regardless 268 gender (Wilcoxon's non parametric test, p=0.019) (Fig 10).

Table 4. Genotyping results of lambs generated in at-risk matings in private and experimental farms according to the sex and *MMUT* SNV genotype.

	g.23,77	g.23,776,347G>A genotype in MMUT			
	G/G	G/A	A/A	-/-	Total
Experiment 1:					
Private farms (n=6)					
Male	9	9	10	1	29
Female	8	16	5		29
Undetermined			1		1
Total	17	25	16	1	59
Experiment 2:					
Experimental farm (n=1)					
Male	4	2	2		8
Female	0	2	3		5
Total	4	4	5		13
All					
Male	13	11	12	1	37
Female	8	18	8	0	34
Undetermined			1		1
Total	$21 (19^{a}/2^{\dagger})$	29 (25 <sup>a</sup> /4 <sup>†</sup> )	$21 (2^{a}/19^{\dagger})$	$1^* (0^a/1^\dagger)$	72

- 271 \*The ear punch was not available for this lamb.
- <sup>a</sup> Number of alive lambs.
- <sup>†</sup>Number of dead lambs.

#### 274 MMUT protein expression and activity

275 Based on the sheep gene atlas (http://biogps.org/sheepatlas/; accessed 17 February 276 2022), MMUT is shown to be highly expressed in kidney and liver [52]. In order to evidence a 277 putative nonsense-mediated mRNA decay (NMD) due to the nonsense variant in MMUT, we 278 evaluated the MMUT mRNA relative expression in kidney and liver by qPCR (Fig 11A). We 279 evidenced a significant reduction of MMUT expression in liver (Wilcoxon's non parametric 280 test, p=0.0015) but not in kidney (Wilcoxon's non-parametric test, p=0.66). We also assessed the MMUT protein expression from liver and kidney protein extracts collected from two 281 282 homozygous A/A and two homozygous G/G dead lambs. As expected, using Western blotting, 283 the wild type protein was expressed in kidney and liver whereas the mutated protein was not detected at least as a full-length form in both tissues (Fig 11B). We also tried to evaluate the accumulation of methylmalonic acid quantified by ELISA in urine and blood of A/A lambs collected post-mortem or soon after birth in alive animals (Fig 12) but no significant difference was observed compared to G/G or G/A lambs (Wilcoxon's non-parametric test, p=0.54 in urine, Wilcoxon's non-parametric test, p=0.34 in plasma).

## 289 **Discussion**

290 Using a reverse genetic screen in MTR population, we successfully identified five genomic 291 regions (named MTRDHH1 to 5) with a significant deficit of homozygous animals ranging 292 from 84 to 100%. Compared to our previous analysis in Lacaune dairy sheep with 8 independent 293 haplotypes, we identified less deficient haplotypes in MTR possibly due to a lower number of 294 genotyped animals (19,102 Lacaune vs 6,845 MTR) [42]. In the MTR population, we estimated 295 the frequencies of MTRDHH heterozygous carriers between 7.8% and 16.6%, and thus allele 296 frequency between 3.9 and 8.3%. This is in line with an allele frequency of 5% expected from 297 the analysis of a population of 6,000 genotyped animals [24].

298 We tested a putative impact of each MTRDHH on production traits to search for a selective 299 advantage at heterozygous state. The positive effects of these DHH on selected traits were quite 300 low and only significant for MTRDHH2 carriers on milk yield and for MTRDHH5 on fat yield. 301 It is thus unlikely that selective advantage explains the observed DHH frequencies, or the rapid increase in MTRDHH3 frequency between 2017 and 2018. However, many examples of 302 303 balancing selection for deleterious alleles have been described in livestock [53]. For example 304 in sheep, a missense variant in FGFR3 was associated with enhanced skeletal growth and meat 305 yield when heterozygous while it induced chondrodysplasia (spider lamb syndrome) when 306 homozygous (OMIA 001703-9940) [54].

The populational analysis of AIS and SBR recorded on more than 300,000 matings, allowed us to do a first sorting of the different MTRDHH based on their supposed deleterious impact on early gestation (AIS), around the time of birth (SBR), or on postnatal viability or morphological phenotypes when AIS and SBR were not altered. Accordingly, we classified the five haplotypes into three groups and we highlighted potential candidate genes according to their implication in lethal phenotypes in mouse and/or more generally involved in inherited mammalian disorders.

314 The first group is composed of MTRDHH3 (OAR1), MTRDHH4 (OAR10) and 315 MTRDHH5 (OAR13) which showed no significant impact on fertility traits. We assume that 316 these haplotypes host deleterious variants leading to postnatal lethality or morphological 317 disorders. Within these three regions, 18 candidate genes are of interest for their implication in 318 neonatal to juvenile lethality or associated with morphological defects counter selected at the 319 time of candidate lamb genotyping. Candidate genes associated with postnatal lethality mainly 320 affect metabolism (TARS2 in MTRDHH3, KL in MTRDHH4) or DNA repair (BRCA2 in 321 MTRDHH4). Some candidate genes not necessarily lead to lethality when altered, but may 322 affect animal welfare with the alteration of vision (PRPF3, ADAMTSL4 in MTRDHH3 and 323 KIF3B in MTRDHH5), neurological disorders (PRUNE1, POGZ in MTRDHH3 and ASXL1, 324 PIGU, AHCY in MTRDHH5) or morphological/stature defects (ITGA10, POLR3GL, ECM1 in 325 MTRDHH2 and DNMT3B, CEP250 in MTRDH5). However, by searching causal variants in 326 the 22 WGS of MTR animals, we failed to detect candidate mutations in the genes listed above. 327 In contrast, we were puzzled by the presence of *RXFP2* (MTRDHH4) and *ASIP* (MTRDHH5) 328 as positional candidate genes already known to host variants controlling the horned/polled 329 phenotype (OMIA 000483-9940) [48] and black coat color (OMIA 000201-9940) [49,50] in 330 sheep, respectively. Interestingly, in MTR selection scheme, horned females and black animals 331 do not fit with breed standard, and thus are not desirable. This may lead excluding these animals

from genotyping which fully express their phenotype at the homozygous state. Accordingly, based on specific SNP markers already present on the LD SNP chip used, we were able to significantly associate MTRDHH5 with the 5pb deletion in *ASIP* leading to black coat color when homozygous. In contrast, the deficit observed for MTRDHH4 is significantly associated with the Ins allele of *RXFP2* leading to the desirable polledness trait. This intriguing observation suggests a hitchhiking phenomenon explaining the deficit [55], with the presence of a deleterious variant in linkage disequilibrium with the Ins allele of *RXFP2*.

339 The second group with MTRDHH2 (OAR1) associated with significant negative effects on 340 AIS and SBR. Then, we hypothesized that the causative variant hosted by MTRDHH2 could 341 induce embryo/fetal losses throughout the gestation period and until birth. In this region, only 342 SLC33A1 (Solute Carrier Family 33 Member 1) gene has both impact on embryonic lethality 343 and decrease survival rate when knocked-out in mouse (MGI:1332247) and thus appears as an 344 obvious candidate gene. In addition, variants in SLC33A1 (OMIM 603690) are known to cause 345 "Congenital cataracts, hearing loss, and neurodegeneration" and "Spastic paraplegia 42" 346 phenotypes [56,57]. A study is ongoing to evidence a candidate causal variant affecting the 347 SLC33A1 gene in linkage disequilibrium with MTRDHH2.

348 MTRDHH1 (OAR20) was the only haplotype with a strong negative impact exclusively on 349 SBR suggesting that this haplotype hosts a lethal variant affecting the perinatal period. Within 350 the MTRDHH1 region, we identified MMUT (Methylmalonyl-CoA Mutase), TFAP2B 351 (Transcription Factor AP-2 Beta) and PKHD1 (PKHD1 Ciliary IPT Domain Containing 352 Fibrocystin/Polyductin) as obvious candidate genes, all resulting in neonatal and/or postnatal 353 lethality when knocked-out in mouse. MMUT and PKHD1 are involved in metabolic disorders 354 such as "Methylmalonic aciduria" (OMIM 251000) and "Polycystic kidney disease" (OMIM 355 263200), respectively. TFAP2B is involved in bone defects and heart failure (OMIM 169100). 356 Using WGS data from 22 MTR animals and among them, two MTRDHH1 heterozygous 357 carriers, we were able to identify four candidate polymorphisms within MTRDHH1, but only 358 one appeared as a strong functional candidate, a SNV (G>A) located in the MMUT gene at 359 position g.23,776,347 on OAR20. This SNV leads to a nonsense variation XM\_004018875.4: 360 c.1225C>T introducing a premature stop-codon (p.Gln409\*). The genotyping of g.23,776,347G>A in 714 animals with a known status at MTRDHH1 indicated an almost 361 362 perfect association. The only 15 discordant animals were largely explained by shorter 363 recombinant version of the MTRDHH1 haplotype (between 2 to 31 markers surrounding the 364 SNV). Only one MTRDHH1 heterozygous carrier did not carry the MMUT variant. This 365 discrepancy could be attributed to errors from SNP array genotyping, phasing and/or 366 imputation. Immunoblotting realized on liver and kidney proteins extracts from homozygous 367 carrier lambs has confirmed the predicted impact of the candidate SNV on the protein. The anti-368 MMUT antibody has revealed a band at 83kDa as expected for the full-length MMUT 369 polypeptide in wild-type extract. In contrast, due to the lack of the antigenic epitope (from aa 370 451 to 750) in the p.Gln409\* truncated form, the MMUT band was not detected in homozygous 371 carriers proving that this SNV has an effect on the functional expression of the MMUT gene, 372 reinforced by a NMD phenomenon detected in liver. In order to confirm the perinatal lethality 373 supposed for MTRDHH1 increasing SBR by 7.5%, we have managed at-risk mating between 374 heterozygous carriers of the p.Gln409\* variant. In this experiment, 84% of homozygous lambs 375 died within the first 24 hours after birth fitting perfectly with the hypothesis and explaining the 376 complete deficit of homozygous carriers of MTRDHH1 in the DHH analysis, due to lethality 377 before the time of genotyping. We also evidenced that homozygous newborn lambs had a lower 378 birth weight, an observation which could be compared to postnatal growth retardation in Mmut 379 knock-out mice (MGI:5527455).

In human, numerous pathogenic variants in *MMUT* cause "Methylmalonic aciduria"
(OMIM 609058, MMA, [58]), an autosomal recessive metabolism disorder. MMUT is part of

382 a metabolic pathway starting from the degradation of amino acids (Valine, Isoleucine, 383 Methionine and Threonine), odd-chain fatty acids, cholesterol and propionic acid to succinyl-384 CoA by three main enzymes: Propionyl-CoA carboxylase (PCC), Methylmalonyl-CoA 385 epimerase (MCE) and Methylmalonyl-CoA mutase (MMUT) [59,60]. The MMUT protein is a 386 mitochondrial enzyme that catalyze the L-methylmalonyl-CoA to succinyl-CoA, an 387 intermediate of Krebs cycle. The isomerization of methylmalonyl-CoA requires 388 adenosylcobalamin (AdoCbl), the cofactor form of vitamin B12 (also known as Cobalamin) 389 [61]. In sheep, the mutant protein (p.409Gln\*) does not carry the B12 binding domain (615-733 390 amino acids), suggesting that the AdoCbl cofactor is unable to act in the conversion of L-391 methylmalonyl-CoA to succinyl-CoA. When the MMUT enzyme is not functional, 392 methylmalonic acid (MLA) accumulates in body fluids, mainly in blood and urine [60,62]. 393 However, we failed in evidencing such MLA accumulation in plasma or urine of homozygous 394 A/A lambs, while this was observed in homozygous knock-out mouse 24h after birth [63]. In 395 our study, many of our urine and blood samples were collected at necropsy only after natural 396 death of the lambs without time control, possibly affecting the results. The methylmalonic 397 aciduria in this sheep genetic model will need further clinical investigation.

398 All the above elements clearly indicate that the deficit of homozygous MTRDHH1 is due 399 to a loss-of-function mutation in the MMUT gene altering an essential metabolic pathway. 400 Many reverse genetic screen approaches have also evidenced the association of DHH with 401 mutations in genes implied in metabolism [64]. Particularly in bovine, many variants were 402 evidenced affecting metabolic processes in several breeds such as Braunvieh (BH24/CPT1C, 403 lipid metabolism) [19], Holstein (HH4/GART, nucleotide metabolism) [13], Montbéliarde 404 (MH1/PFAS, nucleotide metabolism; MH2/SLC37A2, glucose metabolism) [13,40], 405 Normande (NH7/CAD, nucleotide metabolism) [26], and Simmental (SH8/CYP2B6, respiratory 406 chain) [27].

407 Diversity analysis have also revealed the segregation of MMUT SNV variant in Spanish 408 Latxa Cara Rubia (LCR) breed. French MTR and Spanish LCR are very close populations. 409 Since 1970s, many exchanges have occurred between these population across the border, first 410 from Spain to France during the seventies, then the reverse since the nineties [65]. The MMUT 411 SNV evidenced in this study is not shared by other French sheep breeds and more largely by 412 the individuals from the International Sheep Genome Consortium (dataset composed of 453 413 animals from 38 breeds all over the world, https://www.sheephapmap.org/). However, 414 searching this dataset, we found another SNV located in ovine *MMUT* gene (rs1093255812) 415 leading to a premature stop-gain (ENSOART00020022357.1: p.Trp43\*). This variant was 416 identified at heterozygous state in the New Zealand Coopworth breed, and as in MTR, it could 417 have the same deleterious impact on lamb viability.

## 418 **Conclusion**

419 In this study, we firstly identified in the MTR dairy sheep the segregation of five 420 independent haplotypes possibly hosting five recessive deleterious alleles of which at least three 421 have negative impact on fertility traits. Among them, we evidenced the MMUT c.1225T 422 (p.409Gln\*) variant associated with the MTRDHH1 haplotype causing early lamb mortality 423 when homozygous. This could provide an excellent animal model for the study of 424 methylmalonic aciduria occurring in human with the same recessive genetic determinism 425 affecting MMUT. The MTRDHH2 and 3 represent promising haplotypes to discover other 426 recessive lethal mutations in the near future. This reverse genetic study has also allowed us to 427 hypothesize that two of these haplotypes, MTRDHH4 and 5, do not associate with lethal 428 mutations but with mutations possible affecting morphological traits. The causal mutations are 429 still to be identified since the already known mutations in *RXFP2* and *ASIP* do not fit perfectly with the segregation of these two haplotypes. Anyway, an appropriate management of these
haplotypes/variants in the MTR dairy sheep selection program should increase the overall
fertility and lamb survival, and may help for selection of morphological breed standards.

## 433 Materials and methods

#### 434 Animal and genotyping data

The total dataset is composed of 6,845 genotyped Manech Tête Rousse animals (82% males and 18% females) born between 1993 and 2021 (data description in S4 Fig). Genotyping was performed at the Labogena facility (http://www.labogena.fr/) in the framework of the French dairy sheep genomic selection [66]. Both low density 15k SNP chip (SheepLD; n=2,956) and

- 439 medium density 50k SNP chip (Ovine SNP50 BeadChip; n=3,889) were purchased from
- 440 Illumina Inc. (San Diego, USA) (Table 5). Pedigree information was extracted from the official
- 441 national database SIEOL (Système d'Information en Elevage Ovin Laitier, France).

#### 442 **Table 5. Description of genotyped animals**

Year of birth	< 2017	≥ 2017	Total		
Background	Research programs	Genomic selection			
		programs			
Number of animals	2,533 rams	3,077 rams	6,845		
	692 ewes	543 ewes			
SNP chip	MD	LD	MD	LD (n = 2,956)	
	(n=3,225)	(n = 2,956)	(n = 664)	MD (n =3,889)	
Genotyping age	>12 months	~15 days	8-12 months		

443 MD: medium density (50k), LD: low density (15k)

#### 444 Genotype quality control, imputation and phasing

445 Quality control for each SNP was carried out following the French genomic evaluation pipeline 446 based on (*i*) call frequency >97%, (*ii*) minor allele frequency >1%, (*iii*) respect Hardy-Weinberg 447 equilibrium (P>10<sup>-5</sup>). Genotypes were phased and imputed from LD to MD using *FImpute3* 448 [67]. The accuracy of LD to MD imputation in MTR was previously assessed and resulted in a 449 concordance rate per animal of 98.86%, a concordance rate per SNP of 98.95% and a squared 450 Pearson's correlation coefficient of 93.57% between imputed and observed SNP genotypes 451 [68]. After quality control, the 38,523 remaining autosomal SNPs were mapped onto the *Ovis*  *aries* genome assembly Oar v3.1 (current version used in the French genomic evaluation) [69].
These SNP were also located on the genome assembly Oar\_rambouillet\_v1.0
(GCF\_002742125.1). Genomic coordinates of both version of the sheep genome are available
at https://doi.org/10.6084/m9.figshare.8424935.v2 (https://www.sheephapmap.org/).

#### 456 **Detection of homozygous haplotype deficiency**

457 Following the method used previously to detect deficit in homozygous haplotypes in sheep [42], 458 we screened the genome of the 5,271 genotyped animals belonging to trios (77 offspring had 459 both parents genotyped, and 4,799 offspring had both sire and maternal grandsire genotyped). 460 Briefly, the method consists in (i) screening the genome using a sliding window of 20 SNP 461 markers, (ii) selecting all 20 SNP haplotypes with frequency>1% on maternal phase, (iii) 462 comparing the observed number  $N_{Obs}(k)$  to the expected number  $N_{Exp}(k)$  of homozygous 463 offspring for each haplotype k using within-trios transmission probability and further considered haplotypes with P-poisson  $< 1.9 \times 10^{-4}$ , (*iv*) retaining deficit between 75% and 100% 464 defined as  $(N_{Exp}(k) - N_{Obs}(k))/N_{Exp}(k)$ . Finally, consecutive windows with these same 465 466 parameters were clustered to define larger region called "Manech Tête Rousse Deficient 467 Homozygous Haplotype" (MTRDHH). For each MTRDHH, the status (homozygous non-468 carriers, heterozygous and homozygous carriers) was thereafter determined for all 6,845 469 genotyped animals available. Linkage disequilibrium was estimated between two MTRDHH 470 regions located on the same chromosome by the r2 coefficient measure as described [42].

#### 471 Analysis of fertility traits

472 Mating trait records of MTR between 2006 and 2019 were obtained from the national database
473 SIEOL. Only artificial insemination success (AIS) and stillbirth rate (SBR) records from mating
474 where both sire and maternal grand sire were genotyped (i.e. had a known status at each

MTRDHH) were analyzed. AIS was coded "1" for success and "0" for failure based on lambing 475 476 date according to the gestation length starting from the day of AI ( $151 \pm 7$  days; n=330,844 477 mating records). SBR was determined only in the AI success group, and coded "1" if there was 478 at least one stillbirth in the litter or "0" if all lambs were born alive (n=201,637 mating records). 479 We considered "at-risk mating" a mating between a carrier ram and a ewe from a carrier sire. We considered "safe mating" when the other combinations occurred: (i) non-carrier ram  $\times$  ewe 480 481 from a non-carrier sire, (ii) non-carrier ram  $\times$  ewe from a carrier sire, (iii) carrier ram  $\times$  ewe 482 from a non-carrier sire. A logistic threshold binary model with a logit link function was used to 483 compare AIS and SBR between at-risk and safe mating (Ismeans estimate), using the 484 GLIMMIX procedure in the SAS software (version 9.4; SAS Institute Inc., Cary, NC). The 485 fixed effects for AIS and SBR were mating type (safe or at-risk), season of AI (spring or 486 summer), and lactation number (L1, L2, L3 and L4+). For SBR only, prolificacy of the ewe (1, 487 2, 3+ lambs/litter) was added as a fixed effect. For AIS and SBR, the random effect was the 488 interaction herd×year (n=313 herds between 2006 and 2019). Traits were considered to differ significantly when the mating type fixed effect had a P-value lower than  $1.0 \times 10^{-2}$  after 489 490 Bonferroni correction for multiple testing with a level of significance  $\alpha$  at 5%. This threshold 491 was obtained by dividing the level of significance  $\alpha$  by the number of tests corresponding to the 492 number of independent haplotypes (n=5).

#### 493 Analysis of milk parameters and total merit genomic index (ISOLg)

Daughter yield deviations (DYD) for milk traits from genotyped sires with known status at each MTRDHH were computed from official genetic evaluation (GenEval, Jouy-en-Josas, France). The DYD corresponds to the average performance of the daughters of each sire, corrected for environmental effects and the average genetic value of the dam [66]. The six traits studied were milk yield (MY), fat (FC) and protein (PC) contents, fat (FY = MY×FC) and 499 protein (PY = MY $\times$ PC) yields, and lactation somatic cell score (LSCS) as described [42]. To 500 compare all the traits on the same scale, each DYD was divided by its genetic standard deviation 501 to obtain standardized DYD (sDYD). Only genotyped rams with records from at least 20 502 daughters were included in the analysis in order to obtain sufficiently accurate DYD values (n 503 ~2570 rams). Each trait was tested by variance analysis comparing MTRDHH carrier and non-504 carrier rams using the GLM procedure in the SAS software (version 9.4; SAS Institute Inc., 505 Cary, NC). The fixed effects were the genetic status (carrier, non-carrier) and year of birth 506 (2000 to 2016) to correct for annual genetic gain. Traits were considered to differ significantly when the genetic status fixed effect had a P-value lower than  $1.0 \times 10^{-2}$  after Bonferroni 507 508 correction for multiple testing with a level of significance  $\alpha$  at 5%. This threshold was obtained 509 as explained above for fertility traits (n=5 independent haplotypes).

Total merit genomic index in dairy sheep (called "ISOLg", *Index Synthétique des Ovins Laitiers*) was extracted from the official genomic evaluation (GenEval, Jouy-en-Josas, France) for all the 714 genomic candidate lambs born in 2021. ISOLg is determined by a combination of four selected traits: MY, FC, PC and LSCS. For each MTRDHH, ISOLg from heterozygous carrier and non-carrier lambs were compared with a Wilcoxon non-parametric test under the null hypothesis with a risk of  $\alpha$ =5% using "wilcox.test" function in R software (version 4.1.3, R Core Team, 2022).

#### 517 Identification of positional and functional candidate genes

All annotated genes located in the MTRDHH region extended by 1 Mb from each side were extracted from the ovine genome Oar\_rambouillet\_v1.0 (OAR1: NC\_040252.1, OAR10: NC\_040261.1, OAR13: NC\_040264.1 and OAR20: NC\_040271.1) using CLC export annotation function (QIAGEN CLC Main Workbench 7.9). Genes with a known knock-out phenotype in mouse including mortality and aging (embryonic, prenatal, perinatal, neonatal, 523 postnatal, preweaning, premature death and decreased survival rate) or associated with 524 mammalian genetic disorders were sorted using "biomaRt" R package (version 2.52.0, 525 https://doi.org/doi:10.18129/B9.bioc.biomaRt) extracted from Mouse Genome Informatics 526 (MGI, http://www.informatics.jax.org), International Mouse Phenotyping Consortium (IMPC, 527 https://www.mousephenotype.org), Online Mendelian Inheritance in Man (OMIM, 528 https://omim.org) and Online Mendelian Inheritance in Animal (OMIA, https://omia.org) 529 databases (last accession on 27 May 2022). Relevant candidate genes were presented as a 530 heatmap using "pheatmap" R package (version 1.0.12).

#### 531 Whole genome sequencing data

Publicly available data of 100 ovine short-read Illumina HiSeq/NovaSeq whole genome sequences (WGS) from 14 breeds generated in various INRAE and Teagasc research projects were used for variant calling. Among them, 22 WGS were obtained from MTR dairy sheep also genotyped with the MD SNP chip. A description of the different breeds and the accession numbers of sequencing raw data are available in S3 Table.

#### 537 WGS variant calling and filtering

Reads mapping, variant calling and functional annotation were performed using Nextflow v20.10.0 and Sarek v2.6.1 pipelines for the 100 short-read WGS as previously described [43].
Regions of interest were extracted using SnpSift Filter, part of the SnpEff toolbox [51].
Candidate variants were filtered based on the correlation between haplotype status
(homozygous non-carriers, heterozygous and homozygous carriers encoded as 0, 1 and 2, respectively) and allele dosage for bi-allelic variants (also encoded 0, 1 and 2) using geno--r2 command of VCFtools [70].

#### 545 Specific variant genotyping assays

546 Genotyping at the RXFP2 (1.8 kb InDel) and ASIP (5pb InDel) loci were directly 547 obtained from the LD SNP chip based on the specific markers described in S4 Table. The 548 NC\_040264.1:g.66,474,980T>A in the last exon of ASIP and the SNP NC\_040271.1: 549 g.23,776,347G>A in *MMUT* were both genotyped by PACE (PCR allele competitive extension) 550 analysis. Fluorescent PACE analysis was done with 15 ng of purified DNA using the PACE-551 IR 2x Genotyping Master mix (3CR Bioscience) in the presence of 12 µM of a mix of extended 552 allele specific forward primers and 30 µM of common reverse primers in a final volume of 10 553 µL (primer sequences described in S5 Table). The touch-down PCR amplification condition 554 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 555 (dropping 0.8°C per cycle), then 36 cycles of 20 s at 94°C and 60 s at 54°C performed on an 556 ABI9700 thermocycler followed by a final point read of the fluorescence on an ABI 557 QuantStudio 6 real-time PCR system and using the QuantStudio software 1.3 (Applied 558 Biosystems).

The presence of the *RXFP2*, *ASIP* and *MMUT* variants was checked in a DNA set of the 2021 cohort of 714 MTR male lamb candidates for genomic selection. DNA was extracted by Labogena (Jouy-en-Josas, France) on behalf of the MTR breed industry. A DNA diversity panel of 851 animals from 25 French sheep breeds [71] and 3 Spanish sheep breeds was also specifically genotyped for the *MMUT* variant.

#### 564 Generation of homozygous lambs

565 Blood samples (3 ml) were first collected from 181 ewes, daughters of MTRDHH1 566 carrier sires, located in 6 private farms by jugular vein puncture with the Venoject system 567 containing EDTA (Terumo, Tokyo, Japan) and directly stored at -20°C. Among them, 82 ewes 568 were genotyped as heterozygous carriers of the MMUT variant and were selected to be 569 inseminated with MMUT variant heterozygous rams (at-risk matings). Experiment 1 (n=73 570 ewes) was performed in 6 private farms in Pays-Basque (France) and Experiment 2 (n=9 ewes) 571 was performed at the INRAE experimental farm of Langlade under the agreement number 572 E31429001 (Pompertuzat, France). Experimental design is described in Fig 13. An ultrasound 573 diagnosis of gestation was realized between 45 and 60 days after AI. Gestations were followed 574 and each lamb was monitored form birth to weaning. Ear biopsies (1 mm<sup>3</sup>) from the 72 lambs 575 born in both experiments were obtained with a tissue sampling unit (TSU, Allflex Europe, Vitré, 576 France) and directly placed in the TSU storage buffer at 4°C. Ear biopsies were placed twice 577 consecutively in 180 µL of 50 mM NaOH, heated 10 min at 95°C, neutralized with 20 µL of 1 578 M Tris-HCl, and then vortexed during 10s. All neutralized samples were used for direct 579 genotyping without DNA purification as described [72]. In Experiment 2, all lambs were 580 weighted at birth. Biological samples (plasma, urine, liver and kidney) were collected on 581 animals as described in Fig 13 and frozen at -80°C until use.

#### 582 Immunoblot analysis

583 Frozen kidney and liver tissues were crushed in liquid nitrogen using a Mixer Mill during 30 584 seconds at 30 Hz (MM400, Retsch technology), and 12 mg of tissue powder were lysed with 585 500µL of RIPA solution (Ref#R0278, Sigma-Aldrich). The protein extracts were centrifuged 586 for 20 minutes at 16 000g and 4°C, and protein concentration in the supernatant was determined 587 using BCA protein assay Kit (Ref#K812-1000, Biovision). Each protein sample (45 µg) was 588 denatured and reduced in Laemmli buffer (62.5mM TRIS pH 6.8, 2% SDS; 10% glycerol) 589 containing 5% ß-mercaptoethanol before SDS-PAGE on a 4-15% polyacrylamide gel (Bio-590 Rad). Proteins were transferred onto a nitrocellulose membrane blocked with AdvanBlock-591 Chemi blocking solution (Advansta) during 1 hour at room temperature. After washing in PBS-

592 0.1% Tween 20, the membrane was incubated overnight à 4°C with a rabbit polyclonal anti-593 MMUT primary antibody (MUT Rabbit pAb, Ref#A3969, ABclonal), followed (after washing) 594 by 1h with a rabbit polyclonal anti-Actin primary antibody (Ref#A2066, Sigma-Aldrich) both 595 at 1/1000 in blocking solution. Revelation of the primary antibodies was performed by 596 incubation with goat anti-rabbit Horseradish peroxidase-conjugated secondary antibody 597 (Ref#A0545, Sigma-Aldrick) at 1/10000 ratio in blocking solution for 1 hour at room 598 temperature, followed by enhanced chemiluminescence detection (WesternBright Quantum 599 HRP substrate, Advansta) on a ChemiDoc Touch low-light camera (Bio-Rad) in automatic 600 mode.

#### 601 **RNA extraction, reverse transcription and quantitative PCR**

602 Total RNA was extracted from 80 mg of frozen kidney (n=15) and liver tissue powders (n=15) 603 in 1mL Trizol reagent (Invitrogen, #Ref 15596-018) and isolated from Nucleospin® RNA II 604 kit (Macherey-Nagel, #Ref 740955.50) according to the manufacturer's protocol and including 605 DNAseI digestion treatment. The RNAs were quantified by spectrophotometry (NanoDrop® 606 ND-8000 spectrophotometer, ThermoFischer) and stored at -80°C. Reverse transcription was 607 carried out from 1µg of total RNA in solution with anchored oligo(dT) T22V (1µL at 10µM), 608 random oligo-dN9 (1 $\mu$ L at 10 $\mu$ M) and dNTPs (2 $\mu$ L at 10 mM) in a reaction volume of 10  $\mu$ L. 609 This mixture was incubated at 65°C for 5 min in an ABI2700 thermocycler (Applied 610 Biosystems) then ramped down to 4°C. A second reaction mixture (8 µL/reaction) containing 611 the reaction buffer (5µL of First strand Buffer 5X, Invitrogen, France), DTT (Dithiothreitol, 612 1µL at 0.1M), Rnasine (1µL, 40 units/µL, Promega, France) and Superscript II reverse 613 transcriptase (1µL, 200 units/µL, Invitrogen, France) was added to the denatured RNA solution 614 (final volume reaction of 18µL) then incubated for 50 minutes at 42°C and placed for 15 615 minutes at 70°C. The complementary DNA (cDNA) solution obtained was directly diluted at 616 1:5 ratio and stored at -20°C. For each pair of primers, amplification efficiency was evaluated 617 by  $E = e^{-1/\alpha}$  which  $\alpha$  is the slope of a linear curve obtained from cDNA serial dilution (1:5 to 618 1:80) and corresponding Ct (cycle threshold) values. Quantitative PCR (qPCR) was performed 619 using 3µL of cDNA at 1:20 ratio, 5µL of SYBR Green real-time PCR Master Mix 2X (Applied 620 Biosystems) and 2µL of primers at 3µM in a total reaction volume of 10µL on qPCR was 621 realized on a QuantStudio 6 Flex Real-Time PCR system (ThermoFisher). Each sample was 622 tested in duplicate. RNA transcript abundance was quantified using the  $\Delta\Delta Ct$  method corrected 623 by four reference genes (GAPDH, YWHAZ, RPL19 and SDHA) and a calibrator sample. Primers 624 were design using Beacon Designer<sup>™</sup> 8 (Premier Biosoft). The list of qPCR primer sequences, 625 amplification length and amplification efficiency used is available in S5 Table.

#### 626 Methylmalonic acid dosage assay

627 Dosage of methylmalonic acid (MLA) was performed on 11 urine and 15 plasma samples 628 collected on lambs. MLA dosage was performed using Sheep Methylmalonic Acid Elisa kit 629 (MyBioSource, Ref# MBS7266308) following the manufacture's protocol starting with 100μL 630 of samples. The incubation phase with specific antibody was performed overnight at 4°C. The 631 optical density at 450 nm was determined on a GloMax®-Multi Detection System (Promega).

#### 632 Ethic statement

The experimental procedures on animals were approved (approval numbers APAFIS#30615-2021032318054889 v5) 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.

#### 637 Acknowledgments

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 and Soline Szymczak for help with genotyping.

#### 645 Funding

- 646 This research has received funding from the European Union's Horizon 2020 research and
- 647 innovation program under the Grant Agreement No. 772787 (SMARTER) and PRESAGE
- 648 project (CASDAR n°20ART1532777, the responsibility of the French ministry of Agriculture
- and Food cannot be engaged). MB was supported by a Ph.D. grant for the HOMLET program
- 650 co-funded by APIS-GENE and Région Occitanie.

#### 651 **References**

- Georges M, Charlier C, Hayes B. Harnessing genomic information for livestock
   improvement. Nature Reviews Genetics. 2019;20: 135–156. doi:10.1038/s41576-018 0082-2
- Rupp R, Mucha S, Larroque H, McEwan J, Conington J. Genomic application in sheep
  and goat breeding. Animal Frontiers. 2016;6: 39–44. doi:10.2527/af.2016-0006
- 657 3. Eggen A. The development and application of genomic selection as a new breeding
  658 paradigm. Animal Frontiers. 2012;2: 10–15. doi:10.2527/af.2011-0027
- 659 4. Chamberlain AJ, Cheng HH, Giuffra E, Kuehn C, Tuggle CK, Zhou H. Editorial:
  660 Functional Annotation of Animal Genomes. Frontiers in Genetics. 2021;12. Available:
  661 https://www.frontiersin.org/article/10.3389/fgene.2021.768626

662 5. Pryce JE, Royal MD, Garnsworthy PC, Mao IL. Fertility in the high-producing dairy cow. 663 Livestock Production Science. 2004;86: 125-135. doi:10.1016/S0301-6226(03)00145-3 664 6. Berry DP, Wall E, Pryce JE. Genetics and genomics of reproductive performance in dairy and beef cattle. Animal. 2014;8: 105-121. doi:10.1017/S1751731114000743 665 666 7. David I, Astruc JM, Lagriffoul G, Manfredi E, Robert-Granié C, Bodin L. Genetic Correlation Between Female Fertility and Milk Yield in Lacaune Sheep. Journal of Dairy 667 668 Science. 2008;91: 4047-4052. doi:10.3168/jds.2008-1113 669 8. Capitan A, Michot P, Baur A, Saintilan R, Hozé C, Valour D, et al. Genetic tools to 670 improve reproduction traits in dairy cattle. Reprod Fertil Dev. 2015;27: 14-21. doi:10.1071/RD14379 671 672 9. Duchesne A, Grohs C, Michot P, Bertaud M, Boichard D, Floriot S, et al. From 673 phenotype to the causal mutation: recessive defects. INRAE Productions Animales. 674 2016;29: 319-328. doi:10.20870/productions-animales.2016.29.5.3000 675 10. Charlier C, Coppieters W, Rollin F, Desmecht D, Agerholm JS, Cambisano N, et al. 676 Highly effective SNP-based association mapping and management of recessive defects in 677 livestock. Nat Genet. 2008;40: 449-454. doi:10.1038/ng.96 678 11. Lander ES, Botstein D. Homozygosity Mapping: A Way to Map Human Recessive Traits 679 with the DNA of Inbred Children. Science. 1987;236: 1567-1570. 680 doi:10.1126/science.2884728 681 12. VanRaden PM, Olson KM, Null DJ, Hutchison JL. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. Journal of Dairy Science. 2011;94: 682 683 6153-6161. doi:10.3168/jds.2011-4624 684 13. Fritz S, Capitan A, Djari A, Rodriguez SC, Barbat A, Baur A, et al. Detection of 685 Haplotypes Associated with Prenatal Death in Dairy Cattle and Identification of Deleterious Mutations in GART, SHBG and SLC37A2. PLOS ONE. 2013;8: e65550. 686 687 doi:10.1371/journal.pone.0065550 688 14. Pausch H, Schwarzenbacher H, Burgstaller J, Flisikowski K, Wurmser C, Jansen S, et al. 689 Homozygous haplotype deficiency reveals deleterious mutations compromising 690 reproductive and rearing success in cattle. BMC Genomics. 2015;16: 312. 691 doi:10.1186/s12864-015-1483-7 692 15. Jenko J. Analysis of a large dataset reveals haplotypes carrying putatively recessive lethal 693 and semi-lethal alleles with pleiotropic effects on economically important traits in beef 694 cattle. 2019; 14. 695 16. Hoff JL, Decker JE, Schnabel RD, Taylor JF. Candidate lethal haplotypes and causal 696 mutations in Angus cattle. BMC Genomics. 2017;18: 799. doi:10.1186/s12864-017-4196-697 2 17. Cooper TA, Wiggans GR, Null DJ, Hutchison JL, Cole JB. Genomic evaluation, breed 698 699 identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle. Journal of Dairy Science. 2014;97: 3878-3882. doi:10.3168/jds.2013-7427 700

- 18. Null D, Hutchison J, Bickhart D, VanRaden P, Cole J. Discovery of a haplotype affecting
  fertility in Ayrshire dairy dattle and identification of a putative causal variant. Journal of
  Dairy Science. 2017;100(Suppl. 2): 199(abstr. 205).
- 19. Häfliger IM, Seefried FR, Spengeler M, Drögemüller C. Mining massive genomic data of
  two Swiss Braunvieh cattle populations reveals six novel candidate variants that impair
  reproductive success. Genetics Selection Evolution. 2021;53: 95. doi:10.1186/s12711021-00686-3
- 20. Sahana G, Nielsen US, Aamand GP, Lund MS, Guldbrandtsen B. Novel Harmful
  Recessive Haplotypes Identified for Fertility Traits in Nordic Holstein Cattle. Pas MFW
  te., editor. PLoS ONE. 2013;8: e82909. doi:10.1371/journal.pone.0082909
- McClure MC, Bickhart D, Null D, VanRaden P, Xu L, Wiggans G, et al. Bovine Exome
  Sequence Analysis and Targeted SNP Genotyping of Recessive Fertility Defects BH1,
  HH2, and HH3 Reveal a Putative Causative Mutation in SMC2 for HH3. PLOS ONE.
  2014;9: e92769. doi:10.1371/journal.pone.0092769
- 22. Cooper TA, Wiggans GR, VanRaden PM, Hutchison JL, Cole JB, Null DJ. Genomic
  evaluation of Ayrshire dairy cattle and new haplotypes affecting fertility and stillbirth in
  Holstein, Brown Swiss and Ayrshire breeds. 2013; ADSA-ASAS Joint Annual
  Meeting :poster T206.
- 719 23. Fritz S, Hozé C, Rebours E, Barbat A, Bizard M, Chamberlain A, et al. An initiator codon
  720 mutation in SDE2 causes recessive embryonic lethality in Holstein cattle. Journal of Dairy
  721 Science. 2018;101: 6220–6231. doi:10.3168/jds.2017-14119
- 24. Hozé C, Fritz S, Baur A, Grohs C, Danchin-Burge C, Boichard D. Accounting for genetic
  conditions in breeding objectives in dairy cattle. In Proceedings of the 24ème Rencontres
  Autour des Recherches Sur les Ruminants: 5-6 December. Paris; 2018.
- 25. Häfliger IM, Spengeler M, Seefried FR, Drögemüller C. Four novel candidate causal
  variants for deficient homozygous haplotypes in Holstein cattle. Sci Rep. 2022;12: 5435.
  doi:10.1038/s41598-022-09403-6
- 26. Mesbah-Uddin M, Hozé C, Michot P, Barbat A, Lefebvre R, Boussaha M, et al. A
  missense mutation (p.Tyr452Cys) in the CAD gene compromises reproductive success in
  French Normande cattle. Journal of Dairy Science. 2019;102: 6340–6356.
  doi:10.3168/jds.2018-16100
- 732 27. Häfliger IM, Seefried FR, Drögemüller C. Reverse Genetic Screen for Deleterious
  733 Recessive Variants in the Local Simmental Cattle Population of Switzerland. Animals.
  734 2021;11: 3535. doi:10.3390/ani11123535
- 28. Derks MFL, Megens H-J, Bosse M, Lopes MS, Harlizius B, Groenen MAM. A systematic
  survey to identify lethal recessive variation in highly managed pig populations. BMC
  Genomics. 2017;18: 858. doi:10.1186/s12864-017-4278-1
- 29. Derks MFL, Lopes MS, Bosse M, Madsen O, Dibbits B, Harlizius B, et al. Balancing
  selection on a recessive lethal deletion with pleiotropic effects on two neighboring genes
  in the porcine genome. Leeb T, editor. PLoS Genet. 2018;14: e1007661.
  doi:10.1371/journal.pgen.1007661

- 30. Derks MFL, Megens H-J, Bosse M, Visscher J, Peeters K, Bink MCAM, et al. A survey
  of functional genomic variation in domesticated chickens. Genet Sel Evol. 2018;50: 17.
  doi:10.1186/s12711-018-0390-1
- 31. Abdalla EA, Id-Lahoucine S, Cánovas A, Casellas J, Schenkel FS, Wood BJ, et al.
  Discovering lethal alleles across the turkey genome using a transmission ratio distortion approach. Animal Genetics. 2020;51: 876–889. doi:10.1111/age.13003
- Todd ET, Thomson PC, Hamilton NA, Ang RA, Lindgren G, Viklund Å, et al. A
  genome-wide scan for candidate lethal variants in Thoroughbred horses. Sci Rep.
  2020;10: 13153. doi:10.1038/s41598-020-68946-8
- 33. Venhoranta H, Pausch H, Flisikowski K, Wurmser C, Taponen J, Rautala H, et al. In
  frame exon skipping in UBE3B is associated with developmental disorders and increased
  mortality in cattle. BMC Genomics. 2014;15: 890. doi:10.1186/1471-2164-15-890
- 34. Schwarzenbacher H, Burgstaller J, Seefried FR, Wurmser C, Hilbe M, Jung S, et al. A
  missense mutation in TUBD1 is associated with high juvenile mortality in Braunvieh and
  Fleckvieh cattle. BMC Genomics. 2016;17: 400. doi:10.1186/s12864-016-2742-y
- 35. Burgstaller J, Url A, Pausch H, Schwarzenbacher H, Egerbacher M, Wittek T. Clinical and biochemical signs in Fleckvieh cattle with genetically confirmed Fanconi-Bickel syndrome (cattle homozygous for Fleckvieh haplotype 2). Berl Munch Tierarztl Wochenschr. 2016;129: 132–137.
- 36. Adams HA, Sonstegard TS, VanRaden PM, Null DJ, Van Tassell CP, Larkin DM, et al.
  Identification of a nonsense mutation in APAF1 that is likely causal for a decrease in
  reproductive efficiency in Holstein dairy cattle. Journal of Dairy Science. 2016;99: 6693–
  6701. doi:10.3168/jds.2015-10517
- 37. Yang Y, Si J, Lv X, Dai D, Liu L, Tang S, et al. Integrated analysis of whole genome and
  transcriptome sequencing reveals a frameshift mutation associated with recessive
  embryonic lethality in Holstein cattle. Animal Genetics. 2021;n/a. doi:10.1111/age.13160
- 38. Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brøndum RF, et al.
  Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex
  traits in cattle. Nature Genetics. 2014;46: 858–865. doi:10.1038/ng.3034
- 39. Schütz E, Wehrhahn C, Wanjek M, Bortfeld R, Wemheuer WE, Beck J, et al. The
  Holstein Friesian Lethal Haplotype 5 (HH5) Results from a Complete Deletion of TBF1M
  and Cholesterol Deficiency (CDH) from an ERV-(LTR) Insertion into the Coding Region
  of APOB. PLOS ONE. 2016;11: e0154602. doi:10.1371/journal.pone.0154602
- 40. Michot P, Fritz S, Barbat A, Boussaha M, Deloche M-C, Grohs C, et al. A missense
  mutation in PFAS (phosphoribosylformylglycinamidine synthase) is likely causal for
  embryonic lethality associated with the MH1 haplotype in Montbéliarde dairy cattle.
  Journal of Dairy Science. 2017;100: 8176–8187. doi:10.3168/jds.2017-12579
- 41. Sonstegard TS, Cole JB, VanRaden PM, Van Tassell CP, Null DJ, Schroeder SG, et al.
  Identification of a Nonsense Mutation in CWC15 Associated with Decreased
  Reproductive Efficiency in Jersey Cattle. PLOS ONE. 2013;8: 54872.
  doi:10.1371/journal.pone.0054872

- 42. Ben Braiek M, Fabre S, Hozé C, Astruc J-M, Moreno-Romieux C. Identification of
  homozygous haplotypes carrying putative recessive lethal mutations that compromise
  fertility traits in French Lacaune dairy sheep. Genetics Selection Evolution. 2021;53: 41.
  doi:10.1186/s12711-021-00634-1
- 43. Ben Braiek M, Moreno-Romieux C, Allain C, Bardou P, Bordes A, Debat F, et al. A
  Nonsense Variant in CCDC65 Gene Causes Respiratory Failure Associated with
  Increased Lamb Mortality in French Lacaune Dairy Sheep. Genes. 2022;13: 45.
  doi:10.3390/genes13010045
- 44. Buisson D, Astruc J-M, Barillet F. Overview and perspectives of the management of
  genetic diversity in French dairy sheep. INRAE Productions Animales. 2018;31: 1–12.
  doi:10.20870/productions-animales.2018.31.1.2202
- 45. IDELE. VARUME. In: idele.fr [Internet]. 2016 [cited 14 Jun 2023]. Available:
  https://idele.fr/detail-article/indicateurs-de-variabilite-genetique-races-ovines-laitieresedition-2016
- 46. Rodríguez-Ramilo ST, Elsen JM, Legarra A. Inbreeding and effective population size in
  French dairy sheep: Comparison between genomic and pedigree estimates. Journal of
  Dairy Science. 2019;102: 4227–4237. doi:10.3168/jds.2018-15405
- 47. Astruc J-M, Baloche G, Buisson D, Labatut J, Lagriffoul G, Larroque H, et al. Genomic
  selection in French dairy sheep. Inra Prod Anim. 2016;29: 41–55.
- 48. Wiedemar N, Drögemüller C. A 1.8-kb insertion in the 3'-UTR of RXFP2 is associated
  with polledness in sheep. Animal Genetics. 2015;46: 457–461.
  doi:https://doi.org/10.1111/age.12309
- 49. Smit MA, Shay TL, Beever JE, Notter DR, Cockett NE. Identification of an agouti-like
  locus in sheep. Animal Genetics. 2002;33: 383–385. doi:10.1046/j.13652052.2002.00896\_5.x
- 50. Norris BJ, Whan VA. A gene duplication affecting expression of the ovine ASIP gene is
  responsible for white and black sheep. Genome Res. 2008;18: 1282–1293.
  doi:10.1101/gr.072090.107
- 51. Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, et al. Using Drosophila
  melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New
  Program, SnpSift. Frontiers in Genetics. 2012;3: 35. doi:10.3389/fgene.2012.00035
- 52. Clark EL, Bush SJ, McCulloch MEB, Farquhar IL, Young R, Lefevre L, et al. A high
  resolution atlas of gene expression in the domestic sheep (Ovis aries). PLOS Genetics.
  2017;13: e1006997. doi:10.1371/journal.pgen.1006997
- 53. Derks MFL, Steensma M. Review: Balancing Selection for Deleterious Alleles in
  Livestock. Frontiers in Genetics. 2021;12. Available:
  https://www.frontiersin.org/article/10.3389/fgene.2021.761728
- 54. Beever JE, Smit MA, Meyers SN, Hadfield TS, Bottema C, Albretsen J, et al. A singlebase change in the tyrosine kinase II domain of ovine FGFR3 causes hereditary

- chondrodysplasia in sheep. Animal Genetics. 2006;37: 66–71. doi:10.1111/j.13652052.2005.01398.x
- 55. Boichard D, Grohs C, Danchin-Burge C, Capitan A. Genetic defects: definition, origin,
  transmission and evolution, and mode of action. INRAE Productions Animales. 2016;29:
  297–306. doi:10.20870/productions-animales.2016.29.5.2997
- 56. Lin P, Li J, Liu Q, Mao F, Li J, Qiu R, et al. A Missense Mutation in SLC33A1, which
  Encodes the Acetyl-CoA Transporter, Causes Autosomal-Dominant Spastic Paraplegia
  (SPG42). The American Journal of Human Genetics. 2008;83: 752–759.
  doi:10.1016/j.ajhg.2008.11.003
- 57. Huppke P, Brendel C, Kalscheuer V, Korenke GC, Marquardt I, Freisinger P, et al.
  Mutations in SLC33A1 Cause a Lethal Autosomal-Recessive Disorder with Congenital
  Cataracts, Hearing Loss, and Low Serum Copper and Ceruloplasmin. The American
  Journal of Human Genetics. 2012;90: 61–68. doi:10.1016/j.ajhg.2011.11.030
- 58. Acquaviva C, Benoist J-F, Pereira S, Callebaut I, Koskas T, Porquet D, et al. Molecular
  basis of methylmalonyl-CoA mutase apoenzyme defect in 40 European patients affected
  by mut<sup>o</sup> and mut– forms of methylmalonic acidemia: Identification of 29 novel mutations
  in the MUT gene. Human Mutation. 2005;25: 167–176. doi:10.1002/humu.20128
- 839 59. Baumgartner MR, Hörster F, Dionisi-Vici C, Haliloglu G, Karall D, Chapman KA, et al.
  840 Proposed guidelines for the diagnosis and management of methylmalonic and propionic 841 acidemia. Orphanet J Rare Dis. 2014;9: 130. doi:10.1186/s13023-014-0130-8
- 60. Deodato F, Boenzi S, Santorelli FM, Dionisi-Vici C. Methylmalonic and propionic
  aciduria. American Journal of Medical Genetics Part C: Seminars in Medical Genetics.
  2006;142C: 104–112. doi:10.1002/ajmg.c.30090
- 61. González-Montaña J-R, Escalera-Valente F, Alonso AJ, Lomillos JM, Robles R, Alonso
  ME. Relationship between Vitamin B12 and Cobalt Metabolism in Domestic Ruminant:
  An Update. Animals. 2020;10: 1855. doi:10.3390/ani10101855
- Lucienne M, Aguilar-Pimentel JA, Amarie OV, Becker L, Calzada-Wack J, da SilvaButtkus P, et al. In-depth phenotyping reveals common and novel disease symptoms in a
  hemizygous knock-in mouse model (Mut-ko/ki) of mut-type methylmalonic aciduria.
  Biochimica et Biophysica Acta (BBA) Molecular Basis of Disease. 2020;1866: 165622.
  doi:10.1016/j.bbadis.2019.165622
- 853 63. Peters H, Nefedov M, Sarsero J, Pitt J, Fowler KJ, Gazeas S, et al. A Knock-out Mouse
  854 Model for Methylmalonic Aciduria Resulting in Neonatal Lethality\*. Journal of
  855 Biological Chemistry. 2003;278: 52909–52913. doi:10.1074/jbc.M310533200
- 64. Charlier C, Li W, Harland C, Littlejohn M, Coppieters W, Creagh F, et al. NGS-based
  reverse genetic screen for common embryonic lethal mutations compromising fertility in
  livestock. Genome Res. 2016;26: 1333–1341. doi:10.1101/gr.207076.116
- 65. Legarra A, Baloche G, Barillet F, Astruc JM, Soulas C, Aguerre X, et al. Within- and
  across-breed genomic predictions and genomic relationships for Western Pyrenees dairy
  sheep breeds Latxa, Manech, and Basco-Béarnaise. Journal of Dairy Science. 2014;97:
  3200–3212. doi:10.3168/jds.2013-7745

- 66. Astruc J-M, Baloche G, Buisson D, Labatut J, Lagriffoul G, Larroque H, et al. Genomic
  selection in French dairy sheep. Inra Prod Anim. 2016;29: 41–55.
- 865 67. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype
  866 imputation using information from relatives. BMC Genomics. 2014;15: 478.
  867 doi:10.1186/1471-2164-15-478
- 868 68. Larroque H, Astruc J-M. Integration of a low density chip in dairy sheep genomic
  869 selection. 2017 p. 27.
- 69. Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, et al. The sheep genome
  illuminates biology of the rumen and lipid metabolism. Science. 2014;344: 1168–1173.
  doi:10.1126/science.1252806
- 70. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant
  call format and VCFtools. Bioinformatics. 2011;27: 2156–2158.
  doi:10.1093/bioinformatics/btr330
- 876 71. Rochus CM, Tortereau F, Plisson-Petit F, Restoux G, Moreno-Romieux C, Tosser-Klopp
  877 G, et al. Revealing the selection history of adaptive loci using genome-wide scans for
  878 selection: an example from domestic sheep. BMC Genomics. 2018;19: 71.
  879 doi:10.1186/s12864-018-4447-x
- 72. Chantepie L, Bodin L, Sarry J, Woloszyn F, Plisson-Petit F, Ruesche J, et al. GenomeWide Identification of a Regulatory Mutation in BMP15 Controlling Prolificacy in Sheep.
  Front Genet. 2020;11. doi:10.3389/fgene.2020.00585

## 883 Figures



**Fig 1. Manhattan plot of HHD identified in Manech Tête Rousse dairy sheep.** Each point represents one haplotype of 20 markers with a frequency > 1% in the maternal phase. The red line represents the P-value threshold  $(1.9 \times 10^{-4})$  used to consider a haplotype with a significant deficit in homozygotes. Only HHD with a deficit in homozygotes  $\ge 75\%$  (green dots) were selected and resulted in the identification of 150 significant HHD clustered in 5 regions (MTRDHH1 to 5). Genomic coordinates refer to the sheep reference genome Oar\_v3.1.



890 Fig 2. Effects of MTRDHH on artificial insemination success (AIS) and stillbirth rate 891 (SBR) between at-risk and safe matings. For each MTRDHH, the frequency of at-risk 892 matings is shown in parentheses. Significant effects are indicated by the corrected P-value for 893 multiple tests with a threshold set at  $\alpha$ =0.1% (\*\*), 0.01% (\*\*\*). NS, not significant.



894 Fig 3. sDYD relative difference between heterozygous and non-carrier rams for 6 traits 895 under selection. MY: milk yield, FY: fat yield, PY: protein yield, FC: fat content, PC: protein content, LSCS: lactation somatic cell score, sDYD standardized daughter yield deviation (DYD 896 897 divided by genetic standard deviation). sDYD relative difference value is obtained from 898 Ismeans estimate according to mating class. Significant effects are indicated by the corrected P-value for multiple tests with a threshold set at  $\alpha = 5\%$  (\*), 0.1% (\*\*). Error bars indicate 899 900 standard errors. Significant favorable effects of heterozygous are in green while significant 901 unfavorable effects are in red.



Fig 4. Evolution of the MTRDHH heterozygous carrier frequencies between 2017 and
 2021 in Manech Tête Rousse male lambs. For each year the effective refers to all candidates
 genotyped to enter in the genomic selection scheme.



905 Fig 5. Relevant candidate genes located in MTRDHH implicated in lethal mouse 906 phenotype and/or associated with mammalian genetic disorders. Red boxes indicate time 907 of known lethality stages in mouse when the gene is knocked-out from MGI/IMPC databases. 908 When the gene is associated with mammalian genetic disorders, the OMIM and/or OMIA trait 909 phenotypes are described in parentheses. Black boxes indicate a gene implicated in animal 910 genetic disorder (OMIA) but with no lethal phenotype observed in mouse.



911 Fig 6. Association between MTRDHH4-5 and causal variants involved in morphological 912 traits. (A) Association between MTRDHH4 status and 1.8kb insertion in the 3'-UTR of RXFP2 913 associated with polledness. Association between MTRDHH5 status and **(B)** OAR13:g.66,475,132\_66,475,136del and (C) OAR13:g.66,474,980T>A in ASIP associated 914 915 with black coat color. Coordinates refer to sheep genome Oar\_rambouillet\_v1.0.



916 Fig 7. Nonsense variant in MMUT gene within MTRDHH1 genomic region. (A) Scatter 917 plot showing the correlation between MTRDHH1 status (NC\_040271.1, OAR20:22,996,023-24,950,669 extended from each side by 1 Mb) and genotype of variants from 100 whole genome 918 919 sequenced animals. Each dot represents one variant. (B) Position of the MMUT gene within the 920 MTRDHH1 haplotype. Black bars indicate the first and the last markers of the Illumina Ovine 921 SNP50 BeadChip defining the limits of MTRDHH1 (S1B Table). (C) MMUT gene structure 922 (GeneID: 101122731) and localization of the MMUT C>T polymorphism identified in the sixth exon (XM\_004018875.4). (UTR: untranslated region; CDS: coding sequence) (D) MMUT 923 annotations 924 protein (XP\_004018923.1) with Pfam domain (accession number: 925 A0A6P3T7X3\_SHEEP) composed of methylmalonyl-CoA mutase (PF01642) and B12-926 binding (PF02310) domains. The mutation creates a premature stop-gain at amino-acid position 927 409.



- 928 Fig 8. Association of *MMUT* SNV genotypes with MTRDHH1 status. +/+: non-carriers;
- 929 MTRDHH1/+: heterozygous carriers and MTRDHH1/MTRDHH1: homozygous carriers
- 930 (Fisher's exact test, p<0.001, without the homozygous MTRDHH1 carriers).



Fig 9. Time distribution of dead lambs in the pre-weaning period. Bat charts depending on
 *MMUT* SNV genotype and lambing place.



Fig 10. *MMUT* SNV genotype effect on lamb birth weight. (A) Boxplot representation of
birth weight (n=13) according to *MMUT* genotypes (B) distribution of birth weight by sex (M:
Male n=8, F: Female n=5) and genotypes. Affected homozygous lambs are in red (A/A)

936 genotype).



Fig 11. RNA and protein expression of *MMUT* in homozygous A/A lambs. *MMUT* gene
expression (mean ±SEM) at mRNA (A) and protein (B) levels in kidney and liver depending
on *MMUT* SNV genotypes.



- 940 Fig 12. Methylmalonic acid dosage in urine and plasma of A/A lambs. Mean ± SEM of
- 941 Methylmalonic acid (MLA) ELISA quantification in blood plasma and urine collected from
- 942 lambs of different genotypes at the *MMUT* SNV.





Fig 13. Experimental design to generate *MMUT* homozygous variant lambs.

## 944 **Supporting information captions**

945 **S1A Table. Clustering HHD into MTRDHH regions.** Table shows all significant haplotypes 946 of 20 markers (150 HHD with frequency > 1%, P-value  $< 1.9 \times 10^{-4}$  and deficit  $\ge 75\%$ ). As 947 described in the Materials and methods section, the 150 HHD could be clustered into 5 948 MTRDHH regions. **S1B Table. SNPs defining the MTRDHH regions.** Table gives the 949 position of each SNP within MTRDHH regions according to the sheep reference genomes 950 Oar\_v3.1, Oar\_rambouillet\_v1.0 and ARS-UI\_Ramb\_v2.0, and the phased alleles of each 951 deficient haplotype.

S1 Fig. Total merit genomic index (ISOLg) for the 2021 cohort genomic lambs (n=714).
(A) Distribution of ISOLg in MTR dairy sheep. The ISOLg is determined by a combination of
four selected traits: MY, FC, PC and LSCS. Comparison of ISOLg according to each DHH
status, (B) MTRDHH1, (C) MTRDHH2, (D) MTRDHH3, (E) MTRDHH4, (F) MTRDHH5.

956 S2 Table. List of the 408 protein coding genes located in the five MTRDHH extended by 957 1 Mb on each side. Information on mouse phenotypes and association with mammalian 958 disorders were extracted for each gene using several databases: MGI: www.informatics.jax.org; 959 IMPC: https://www.mousephenotype.org); OMIM: Online Mendelian Inheritance in Man 960 (https://omim.org) and OMIA: Online Mendelian Inheritance in Animal (https://omia.org). For 961 mouse phenotypes associated with lethality, these were separated by developmental stages and 962 encoded 0 (no lethality) or 1 if the gene is involved in lethality at the given stage.

963 S2 Fig. Association between MTRDHH4-5 and causal variants involved in morphological 964 traits from the 22 sequenced MTR animals. (A) Association between MTRDHH4 status and 965 1.8kb insertion in the 3'-UTR of RXFP2 associated with polledness. Association between 966 MTRDHH5 **(B)** OAR13:g.66,475,132\_66,475,136del status and and (C) 967 OAR13:g.66,474,980T>A in ASIP associated with black coat color. Coordinates refer to sheep 968 genome Oar rambouillet v1.0.

969 S3 Fig. MTRDHH1 recombinant haplotypes from 15 animals showing mismatch between 970 MTRDHH1 status and *MMUT* SNV genotype. MTRDHH1/+ and +/+ refer to heterozygous 971 and non-carriers of MTRDHH1, respectively. The grey column represents the localization of 972 the *MMUT* SNV (g.23,776,347G>A) within the MTRDHH1 haplotype. For each animal, only 973 the phase supposed to host the *MMUT* SNV variant allele is represented. The blue color 974 indicates the portion of local haplotype matching with the MTRDHH1 haplotype.

975 S4 Fig. Distribution of genotyped animals over time. The bar charts represent the number of
 976 genotyped animals according to sex and year of birth. The genomic selection in MTR dairy
 977 sheep was implemented in 2017.

## 978 S3 Table. EMBL-EBI accession numbers of the 100 whole-genome sequences used in the 979 analysis.

## 980 S4 Table. Markers on SheepLD chip for genotyping at the *RXFP2* (1.8 kb InDel) and *ASIP*981 (5pb InDel) loci.

982 **S5 Table. List of PCR primer sequences.**