A missense mutation in PFAS (phosphoribosylformylglycinamidine synthase) is likely causal for embryonic lethality associated with the MH1 haplotype in Montbéliarde dairy cattle

Pauline Michot,*† Sébastien Fritz,*† Anne Barbat,* Mekki Boussaha,* Marie-Christine Deloche,*† Cécile Grohs,* Chris Hoze,*† Laureène Le Berre,†‡ Daniel Le Bourhis,‡‡ Olivier Desnoes,‡‡ Pascal Salvetti,‡‡ Laurent Schibler,†‡ Didier Boichard,* and Aurélien Capitan†‡

*UMR GABI, L’Institut National de la Recherche Agronomique, AgroParisTech, Université Paris Saclay, 78350 Jouy en Josas, France
†Allice, 75595 Paris, France
‡Allice, Station de Phénotypage, 37380 Nouzilly, France

ABSTRACT

A candidate mutation in the sex hormone binding globulin gene was proposed in 2013 to be responsible for the MH1 recessive embryonic lethal locus segregating in the Montbéliarde breed. In this follow-up study, we excluded this candidate variant because healthy homozygous carriers were observed in large-scale genotyping data generated in the framework of the genomic selection program. We fine mapped the MH1 locus in a 702-kb interval and analyzed genome sequence data from the 1,000 bull genomes project and 54 Montbéliarde bulls (including 14 carriers and 40 noncarriers). We report the identification of a strong candidate mutation in the gene encoding phosphoribosylformylglycinamidine synthase (*PFAS*), a protein involved in de novo purine synthesis. This mutation, located in a class I glutamine amidotransferase–like domain, results in the substitution of an arginine residue that is entirely conserved among eukaryotes by a cysteine (p.R1205C). No homozygote for the cysteine-encoding allele was observed in a large population of more than 25,000 individuals despite a 6.7% allelic frequency and 122 expected homozygotes under neutrality assumption. Genotyping of 18 embryos collected from heterozygous parents as well as analysis on nonreturn rates suggested that most homozygous carriers died between 7 and 35 d postinsemination. The identification of this strong candidate mutation will enable the accurate testing of the reproducers and the efficient selection against this lethal recessive embryonic defect in the Montbéliarde breed.

INTRODUCTION

The recent development of genomic selection in cattle has generated an unprecedented large-scale genotyping activity, thus opening new possibilities in the search for causal variants. VanRaden et al. (2011) proposed using these genotyping data to identify haplotypes displaying a significant deficit in homozygous animals and thus to map putative loci responsible for recessive embryonic or perinatal lethality. Several similar studies were subsequently carried out in different breeds and subpopulations (Fritz et al., 2013; Sahana et al., 2013; Sonstegard et al., 2013; McClure et al., 2014; Venhoranta et al., 2014; Pausch et al., 2015; Adams et al., 2016; Schwarzenbacher et al., 2016), leading to the identification of tens of new recessive defects. Using this strategy, we previously reported the identification of 11 haplotypes displaying a significant deficit in homozygosity in the Montbéliarde breed (Fritz et al., 2013). Out of these, MH1 and MH2 were the 2 most frequent haplotypes (9 and 7%, respectively) and showed a significant negative effect on calving rate, confirming their association with embryonic lethal mutations. Then, mining whole-genome sequences (*WGS*) from carrier and noncarrier bulls, we identified 2 nonsense mutations in the sex hormone binding globulin (chr19 g.27956790C>T; SHBG p.Q52X) and solute carrier family 37 A2 protein (chr29 g.28879810C>T; SLC37A2 p.R12X) genes. We confirmed their association with the MH1 and MH2 haplotypes but did not validate their causality with large-scale genotyping or functional studies. Since then, both SNP have been added to the Illumina (San Diego, CA) EuroG10k SNP custom chip, widely used for genomic selection in France, and generated genotyping data for 128,743 Montbéliarde animals. Within

Key words: embryonic lethality, whole-genome sequencing, large-scale genotyping, phosphoribosylformylglycinamidine synthase

Received January 11, 2017.
Accepted June 21, 2017.
Corresponding author: aurelien.capitan@inra.fr
this large population, no homozygote was identified for the SLC37A2 nonsense mutation, whereas 1,043 were expected under neutrality assumption, thus confirming its strong candidate status for MH2-linked embryonic lethality. In contrast, we identified 242 homozygous carriers for the SHBG nonsense mutation (for 650 expected homozygotes), suggesting that it was not the causative mutation for MH1-linked embryonic lethality. Animals being homozygous for the nonsense mutation in SHBG were also detected in Simmental (H. Pausch, Institute of Agricultural Sciences, ETH Zurich, Switzerland, personal communication) and Vorderwälder (Reinartz and Distl, 2016) breeds. The purpose of this follow-up study was threefold: (1) to fine map the MH1 locus, (2) to identify the causative mutation, and (3) to further characterize its phenotypic effects using the large amount of phenotypes, genotypes, and WGS that have accumulated since the initial study.

MATERIALS AND METHODS

Genotyping Data and Haplotype Testing

Genotypes from 174,113 Montbéliarde animals analyzed from 2008 to 2017 with different Illumina chips [BovineSNP50 BeadChip (50K) v1 and v2, Matukumalli et al., 2009; LD chip, Boichard et al., 2012a; BovineHD chip, Illumina, 2015; EuroG10K] were available from the French genomic evaluation database. These were checked for quality and processed in the framework of the French genomic selection pipeline (Boichard et al., 2012b), including imputation and phasing with FI competing (Sargolzaei et al., 2014). Then, phased 50K genotype data from each animal were screened for the MH1 lethal haplotype as described by Fritz et al. (2013) to determine their carrier or noncarrier status. This haplotype consists of 27 consecutive markers from hapL (position 27,643,677 bp on the UMD3.1 assembly of the bovine genome; Zimin et al., 2009) to ARS-BFGL-NGS-111936 (position 29,383,514 bp on chromosome 19). Throughout this article, “MH1 haplotype” and “MH1 carriers” refer to this original haplotype, whereas “MH1 locus” refers to the embryonic lethal mutation. Following the fine mapping of the MH1 locus, a second haplotype test was performed to detect a subsection of 12 consecutive markers of the MH1 haplotype, denoted MH1s haplotype (from BTB-01692321 to ARS-BFGL-NGS-115178; position 28,484,744 to 29,048,973 bp). However, this haplotype test was found to be less reliable than the previous one for determining the carrier or noncarrier status of the animals for the MH1 mutation (see Results).

The SHBG p.Q52X (g.27956790C>T; rs383770500) polymorphism previously described by Fritz et al. (2013) and a new candidate polymorphism for MH1 [g.28511199C>T; rs455876205; phosphoribosylformylglycinamidine synthase (PFAS) p.R1205C] were included in 2 consecutive custom designs of the Illumina EuroG10K custom SNP BeadChip. At the time of the study, genotype information was available for 47,871 Montbéliarde animals for the first variant and for 26,868 Montbéliarde animals and 108,414 additional animals from 17 breeds for the second variant (see Results).

Pedigree Analysis

Pedigree analysis was performed using the French National database for the Montbéliarde breed and the Bayerische Zuchtwert-Informationen Rind database (http://www.lfl.bayern.de/itc/rind/bazi/) for the Simmental breed.

Fine Mapping of the MH1 Locus Using Recombinant Haplotypes

In line with previous studies (Sonstegard et al., 2013; Adams et al., 2016), we screened our population for animals that carried both the original lethal haplotype and a recombining haplotype. Because these animals are alive, they cannot be homozygous for the lethal mutation, assuming that the mutation is fully penetrant. This leads to the exclusion from the location interval of the identical-by-descent segment for which they are homozygous. Here we took advantage of the large contribution of the MH1 carrier bull Boislevin (MONTFRAM00186006232) to the Montbéliarde population. We defined 2 regions of 40 markers denoted hapL and hapR located 3 Mb upstream and downstream, respectively, of the MH1 haplotype. We subsequently identified the hapL and hapR source haplotypes of Boislevin linked to MH1. Among his genotyped descendants, we retrieved all the individuals homozygous for hapL. Then, the hapL haplotype was extended toward MH1 one SNP at a time, and individuals carrying a haplotype different from the initial one (therefore recombinant) were eliminated. The process was continued until no individual was retained. The same approach was repeated for hapR. On each side, the last remaining individuals defined the left or right boundary of the shortest interval carrying the lethal variant. The refined haplotype of 12 markers contained in this critical interval was denoted MH1s.

WGS Data

This study included WGS data from 274 individuals from 15 different French dairy and beef breeds (Boussa-
Evaluation of the Effect of the PFAS Mutation on Nonreturn Rate at 56 d

In our previous study (Fritz et al., 2013), analyzing AI data, we demonstrated a negative effect of the MH1 haplotype on calving rate (i.e., presence or absence of a calf after complete gestation), which was consistent with recessive embryonic lethality. Here, to narrow down the stage of embryonic loss, we evaluated the effect of the PFAS mutation (g.28511199C>T; rs455876205; PFAS p.R1205C) on nonreturn rate at 56 d (NRR56). The NRR56 was coded as 0 if a second insemination was observed in less than 56 d after the first insemination and as 1 otherwise. Because this locus has a recessive effect, the expected decrease in fertility in mating at risk is 0.25 × μ between 2 carriers (with μ being the average conception rate) and 0.5 × 0.5 × \( \frac{1}{2 - f} \) × μ between a carrier bull and the daughter of a carrier, where \( \frac{1}{2 - f} \) is the proportion of carriers among daughters of carriers bull and of dams of unknown genotype and f is the frequency of the deleterious allele. The mean conception rate μ reaches 44% in lactating cows and 55% in heifers. The model used to analyze NRR56 was the same as in Fritz et al. (2013). It accounted for several environmental effects, the fixed effect of the combination of the genotypes of the mated bull and of the sire of the cow for the PFAS mutation, and the random effects of the sire and maternal grandsire of the embryo.

Genotypes for the PFAS mutation were obtained from the Illumina EuroG10K SNP chip or deduced from a haplotype test considering the 27 markers from the 50K chip defining the MH1 haplotype. The haplotype test was based either directly on 50K genotypes or on 50K genotypes that have been previously imputed from the EuroG10K SNP chip. Because the MH1 haplotype is always associated with the PFAS g.28511199T allele (see Results), this procedure allows the reliable identification of at-risk matings. The strong but incomplete association between the PFAS g.28511199T allele and MH1 haplotype (because some recent recombinant animals carry the deleterious allele but not the complete MH1 haplotype) may lead to the misclassification of some true at-risk mating. This is expected to slightly reduce the power of the analysis but not bias the results. Only inseminations before May 2016 with known status for (1) males and females or (2) males and sires of females were considered. In total, 6,176,305 NRR56
data were analyzed, including 270,113 from at-risk matings. Tests were carried out separately for lactating cows and heifers. Finally a t-test was used to determine any significant difference in conception rate between at-risk matings and control groups.

Production and Genotyping of Embryos from Mating at Risk

Two bulls and 4 heifers heterozygous for mutation g.28511199C>T on chromosome 19 (PFAS p.R1205C) and for the MH1 haplotype were selected from among the animals genotyped for genomic selection. In the experimental station of Allice (Nouzilly, France), heifers were superovulated using standard protocols (Muñoz et al., 2014) and inseminated with the semen of the bulls. On d 7, embryos were collected. Their DNA was extracted, and a whole-genome amplification was carried out with the Repli-g Mini Kit (Qiagen, Hilden, Germany). Genotyping for the g.28511199C>T mutation was performed by PCR and Sanger sequencing. Forward (PFAS_F: AGTCCCCTTTCACTCCAGGT) and reverse (PFAS_R: TGGAGTTCCGGAGAAGAAAA) primers were designed from the UMD3.1 bovine genome assembly using Primer3 software (Rozen and Skaletsky, 2000). The PCR amplification was performed using the Go-Taq Flexi DNA Polymerase (Promega, Charbonnières-les-Bains, France) according to the manufacturer’s instructions on a Mastercycler Pro thermocycler (Eppendorf, Montesson, France). The PCR products were subsequently purified and bidirectionally sequenced by Eurofins MWG (Hilden, Germany) using conventional Sanger sequencing. Polymorphisms were detected with novoSNP software (Weckx et al., 2005).

RESULTS

Invalidation of the SHBG p.Q52X Mutation

Since 2013, 47,871 Montbéliarde animals have been genotyped for the SHBG p.Q52X polymorphism with the Illumina EuroG10K custom SNP BeadChip used for genomic selection. Comparison between the genotypes of these animals and their status for the MH1 haplotype revealed incomplete linkage disequilibrium (squared correlation coefficient; r² = 0.61). Notably, although no homozygous carrier of the MH1 haplotype was observed, we identified 242 animals homozygous for the SHBG nonsense mutation. This led us to consider that the SHBG p.Q52X mutation was not the causative mutation for MH1-associated embryonic lethality (Table 1).

Fine Mapping of the MH1 Causal Mutation

In line with previous studies (Sonstegard et al., 2013; Adams et al., 2016), we screened our population for animals carrying both the original lethal haplotype and a recombining haplotype among the inbred descendants of funder bulls who spread the mutation (e.g., Boislevin in our situation; see Materials and Methods; Figure 1a). Such animals cannot be homozygous for the lethal mutation and thus provide useful information to fine map the MH1 mutation. In doing so, we identified several recombinating haplotypes, among which 2 enabled us to reduce the critical interval from 1.74 Mb to a region of 702 kb (between positions 28,450,297 and 29,152,662 bp on chromosome 19 on the UMD3.1 assembly) from marker ARS-BFGL-NGS-95155 to ARS-BFGL-NGS-43359 (Figure 1a and b). Notably, this interval was included inside the MH1 source haplotype defined by Fritz et al. (2013) and excluded the SHBG mutation. These 2 recombinating haplotypes, mostly diffused by the AI sires Redon (MONFRAM0025294344, for HapR1) and Rapallo (MONFRAM007120640289, for HapR2), were observed in the heterozygous state with the MH1 haplotype (i.e., MH1/HapR1 and MH1/HapR2) in 259 and 7 individuals, respectively. Of note, MH1/HapR1 and, to a lesser extent, HapR1/HapR1 were the 2 most frequent genotypes observed among the homozygotes for the SHBG nonsense mutation mentioned previously. Finally, by analyzing pedigree and haplotype information, we were able to trace back each recombination event and to confirm that they were posterior to Boislevin (Figure 1c).

Analysis of the Association Between MH1 Haplotype, MH1s Haplotype, and the Embryonic Lethal Mutation

Following the fine mapping of the MH1 causal mutation, we analyzed, imputed, and phased 50K genotyping data of 26,868 Montbéliarde cattle to test the association between the MH1 haplotype (27 markers), its subsection MH1s (12 markers) contained in the critical interval, and the recessive embryonic lethal mutation

Table 1. Association between MH1 haplotype status and SHBG p.Q52X genotype in 47,871 Montbéliarde cattle

<table>
<thead>
<tr>
<th>Item</th>
<th>+/+</th>
<th>MH1/+</th>
<th>MH1/MH1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHBG Q52/Q52</td>
<td>38,233</td>
<td>3</td>
<td>0</td>
<td>38,236</td>
</tr>
<tr>
<td>SHBG Q52/X52</td>
<td>3,256</td>
<td>6,137</td>
<td>0</td>
<td>9,393</td>
</tr>
<tr>
<td>SHBG X52/X52</td>
<td>13</td>
<td>229</td>
<td>0</td>
<td>242</td>
</tr>
<tr>
<td>Total</td>
<td>41,502</td>
<td>6,369</td>
<td>0</td>
<td>47,871</td>
</tr>
</tbody>
</table>

+/- = MH1 noncarrier; MH1/+ = MH1 heterozygous carrier; MH1/ MH1 = MH1 homozygous carrier.
(Table 2). Whereas none of the animals studied were homozygous MH1/MH1, we observed a total of 131 MH1s/MH1s living individuals, indicating that MH1 haplotype was a better predictor than MH1s haplotype for determining the carrier or noncarrier status for the recessive embryonic lethal mutation (Table 2). Among the 131 MH1s/MH1s individuals, 23 did not carry the entire MH1 haplotype. Pedigree analysis revealed that none of them had Boislevin on both sides of their pedigree. Taken together, these results indicate that 2 versions of the MH1s haplotype segregate in the Montbéliarde population—one ancestral without the

![Figure 1](image)

**Figure 1.** Fine mapping of the MH1-associated embryonic lethal mutation. (a) Identification of recombinant haplotypes in the descendants of Boislevin using Illumina (San Diego, CA) BovineSNP50 genotyping data. The black and dotted lines correspond to the number of descendants of Boislevin that are homozygous for large identical by descent segments on the left and right sides, respectively, of the MH1 haplotype. Most of these animals carry one copy of the original segment containing MH1 and a recombining haplotype. (b) Diagram of recombining haplotypes HapR1 and HapR2. These were used to define the critical interval in which the MH1-associated embryonic lethal mutation should be located. (c) Pedigree of the 2 bulls that have spread HapR1 and HapR2 recombining haplotypes in the Montbéliarde population. Arrows indicate the first supposed carriers of these recombining haplotypes.
embryonic lethal mutation and a second, included in the MH1 haplotype, on which the mutation occurred.

**Screening of WGS Data Identifies a Strong Candidate Polymorphism in the PFAS Gene**

In this step, we considered only variants from 14 carriers of the MH1 haplotype that were never found in the homozygous state in 40 noncarrier Montbéliarde bulls and in additional controls. The SNP and insertion and deletion information was available for 1,296 genomes, whereas structural variations were identified in only 274 genomes (see Materials and Methods). In the critical interval, a total of 27 variants displayed a strong association with the MH1 haplotype (Fisher’s exact test \(P\)-values between \(6.46 \times 10^{-8}\) and \(1.86 \times 10^{-10}\) after Bonferroni correction; Figure 2a; Supplemental Table S1, https://doi.org/10.3168/jds.2017-12579). These comprised 25 small noncoding variants located in regions that are not conserved among 40 eutherian mammal genomes according to Ensembl (http://www.ensembl.org/; EPO_LOW_COVERAGE) and, for 21 of them, segregated in breeds that are not related to the Montbéliarde and Simmental breeds. The 27 variants also comprised 2 missense mutations. Mutation g.28979780C>T (rs4190651; PIK3R5 p.R363H), located in the phosphoinositide-3-kinase regulatory subunit 5 gene, is predicted to be tolerated according to SIFT (Kumar et al., 2009). Moreover, the alternative allele (at the DNA and protein level) is not conserved in mammals and corresponds to the ancestral allele in humans and other species, which led us to reject this candidate mutation. In contrast, mutation g.28511199T (rs455876205) is predicted to be damaging and to result in the substitution of an arginine that is entirely conserved among eukaryotes by a cysteine in a class I glutamine amidotransferase–like domain of the PFAS gene, is predicted to be always associated with the embryonic lethal mutation.

**Table 2. Association between MH1 and MH1s haplotype status in 26,868 Montbéliarde cattle**

<table>
<thead>
<tr>
<th>Item</th>
<th>+/+</th>
<th>MH1/+</th>
<th>MH1/MH1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+/+</td>
<td>21,888</td>
<td>0</td>
<td>0</td>
<td>21,888</td>
</tr>
<tr>
<td>MH1s/+</td>
<td>1,595</td>
<td>3,254</td>
<td>0</td>
<td>4,849</td>
</tr>
<tr>
<td>MH1s/MH1s</td>
<td>23</td>
<td>108</td>
<td>0</td>
<td>131</td>
</tr>
<tr>
<td>Total</td>
<td>23,506</td>
<td>3,362</td>
<td>0</td>
<td>26,868</td>
</tr>
</tbody>
</table>

+/- = noncarrier; MH1/+ = heterozygous carrier of the MH1 haplotype; MH1s/+ = heterozygous carrier of the MH1s haplotype; MH1/MH1 = homozygous carrier of the MH1 haplotype; MH1s/MH1s = homozygous carrier of the MH1s haplotype.

For further verification, we genotyped the PFAS g.28511199C>T polymorphism in a large population of 26,868 Montbéliarde cattle using the Illumina EuroG10K BeadChip. The MH1 haplotype was found to be always associated with the PFAS g.28511199T allele (PFAS p. 1205C allele), whereas the PFAS g.28511199T allele showed a strong but incomplete association with the MH1 haplotype (Table 3) because of recombinations between the MH1 haplotype and other haplotypes. Mining of phased genotype data enabled us to detect 3 main recombination events that account for a total of 201 (117, 51, and 33, respectively) of the 267 observations of heterozygous carriers of the PFAS mutation that are not carriers of the MH1 haplotype. The frequency of allele g.28511199T (PFAS p. 1205C) was 6.7%. No homozygous mutant was observed, even among the 131 animals that were homozygous for the short MH1s haplotype (Table 4), whereas 122 were expected in this data set (\(\chi^2\) \(P\)-value = 3.78 \(\times\) \(10^{-28}\)).

In addition, analysis of genotype data from 108,414 animals from 17 breeds (Table 5) confirmed the absence of the PFAS g.28511199T allele (PFAS p. 1205C) outside of the Montbéliarde as previously observed with WGS data. Surprisingly, the mutation was even absent from 364 genotyped Simmental animals when 3 of 216 whole-genome sequenced Simmental animals were found to be heterozygous. Pedigree analysis revealed that 2 of these 3 Simmental carriers (Joyeux SIMCHEM710488074272 and Samut SIMDEUM000930988845) descend from the Montbéliarde bull Rossli (MONFRAM000000160839, born in 1965) and that the third carrier (Mangfall SIMDEUM000935463615) had missing information before the 1980s on the maternal side of his pedigree. These results indicate that the PFAS g.28511199T (PFAS p. 1205C) allele is extremely rare in Simmental animals and that it has most probably been involuntarily introgressed in this breed through crossbreeding with the closely related Montbéliarde breed.
Figure 2. Identification of a missense variant in PFAS associated with the MH1 embryonic lethal haplotype. (a) Association of filtered polymorphisms with the MH1 haplotype in the 702-kb critical region. Variants with a significant P-value (<0.05) are represented in black. The arrow indicates the PFAS deleterious candidate variant (g.chr19:28511199C>T). Details on variants’ functional annotations are provided in Supplemental Table S1 (https://doi.org/10.3168/jds.2017-12579). (b) Integrative genome viewer screenshot showing sequenced reads supporting the g.28511199C>T substitution on chromosome 19 in one heterozygous carrier bull. (c) Electrophoregrams from a heterozygous carrier (up) of the g.28511199C>T substitution on chromosome 19 and from a wild-type animal (down). This substitution, located in exon 26 of PFAS, is predicted to result in a p.R1205C AA substitution at the protein level. (d) Diagram of the PFAS protein domains: a PurM-like N-terminal domain (PurM), 2 AIR synthase-related protein C-terminal domains (AIR), and the class I glutamine amidotransferase–like domain (ClassIGluT). The p.R1205C AA change is located in a ClassIGluT domain. The PFAS protein features and domain positions were retrieved from Ensembl (http://www.ensembl.org/index.html) and Uniprot (http://www.uniprot.org/). (e) Alignment of PFAS orthologs in different species showing a complete conservation of the p.R1205 AA among eukaryotes.
Analysis of Nonreturn Rate and Genotyping of Embryos Suggests that Mortality Occurs Between 7 and 35 d of Development

In 2013, we showed a negative effect of the MH1 haplotype on calving rate in mating at risk that was consistent with recessive embryonic lethality. This observation is also consistent with the complete absence of homozygous animals for the candidate mutation in the PFAS gene. To narrow down the stage at which the death of homozygous embryos occurs, we tested the effect of the g.28511199T allele on the NRR56 after insemination (Table 6). In comparison with the control group, we observed a significant loss in NRR56 for both heifers and adult cows in mating at risk between PFAS carriers as well as between PFAS carriers and daughters of PFAS carriers. Estimated values were close to the expected effect under the assumption of complete lethality in homozygous embryos (−12.5 and −6.5%, respectively, with an average conception rate of 50%). Considering that the reproductive cycle is on average 21 d in females, this result indicates that embryo development stopped in the first 35 d of gestation. For further verification, we produced embryos from mating at risk between heterozygous carriers of the PFAS g.28511199C>T (PFAS p.R1205C) polymorphism. In total, 18 embryos were recovered at 7 d of gestation and genotyped (Table 7). At this stage, we observed 4 homozygous for the derived allele, a proportion in agreement with Mendel’s rules. None of them showed obvious signs of abnormal development upon visual examination, suggesting that mortality occurs between 7 and 35 d of development.

DISCUSSION

In a previous study, we proposed a candidate variant in the SHBG gene for embryonic lethality segregating with the MH1 haplotype in Montbéliarde cattle (Fritz et al., 2013). Here, with the support of genotyping data, we demonstrated that this variant is not recessive embryonic lethal. Subsequently, we fine mapped the embryonic lethal mutation in a 702-kb region on chromosome 19 and identified 27 variants in association with MH1. Among them, the most compelling candidate variant was the PFAS g.28511199C>T (PFAS p.R1205C) substitution that (1) affects a region con-
served among mammals, (2) is predicted to cause a deleterious AA substitution, and (3) is absent from breeds other than Montbéliarde and the closely related Simmental breed. Moreover, genotyping of more than 20,000 animals revealed a complete linkage disequilibrium between allele g.28511199T and the embryonic lethal mutation. In our initial study (Fritz et al., 2013), we accidentally missed this strong candidate mutation because we used strict filtering criteria. The candidate mutation had to be absent from noncarrier bulls and present in the heterozygous state in the 2 carrier bulls sequenced at that time. Unfortunately, one of them (Boislevin) had a no read coverage for position 28,511,199 bp on chromosome 19. This compromised its genotype call for the corresponding SNP and misled our screen for potential causal mutations. These results demonstrate the importance of having substantial numbers of whole-genome sequenced individuals and of using relaxed filtering criteria (e.g., as implemented here or in Pausch et al., 2014) to uncover causal variations when sequencing is performed at low to medium coverage. They also underline the need for large-scale genotyping to statistically validate or invalidate the causality of candidate embryonic lethal variants.

The PFAS g.28511199C>T variant is predicted to result in the substitution of an arginine that is entirely conserved among eukaryotes by a cysteine in a class I glutamine amidotransferase-like domain of the protein (PFAS p.R1205C). This perfect conservation reveals a strong evolutionary constraint and provides evidence that this residue plays an important role in the stability or activity of the protein. Phosphoribosylformylglycinamidine synthase is involved in the fourth step of the de novo purine synthesis, a universal pathway among eukaryote organisms (with the exception of some parasites) due to the central function of purine molecules in cell metabolism such as DNA and RNA synthesis and energy molecule supply (Henikoff, 1987; Bønsdorff et al., 2004). Although purine synthesis is a regulated balance between de novo and salvage pathways, it has been shown that de novo synthesis is more active in proliferating cells and therefore essential for meeting the high demand of purines for nucleic acid synthesis in early embryonic development (Alexiou and Leese, 1992). Interestingly, mutations of several genes of this pathway have been reported to cause recessive embryonic lethality in different eukaryotic species (e.g., PRAT, GART, PAICS; Clark, 1994; Ng et al., 2009; Fritz et al., 2013). Among them, we should highlight GART, in which we previously identified a missense variant (chr2:g.1277227A>C; p.N290T) associated with the embryonic lethal haplotype HH4 segregating in Holstein cattle (Fritz et al., 2013). Concerning PFAS, several mutations have been described in mice, causing a dominant short-face phenotype and most probably recessive embryonic lethality because no homozygous mutant has ever been obtained from mating heterozygous animals (Palmer et al., 2016). In addition, homozygous deficiency of the PFAS ortholog gene ade2 is lethal during pupal development in Drosophila melanogaster (Henikoff et al., 1986; Holland et al., 2011). Taken together, these arguments strongly support a causative role for the g.28511199C>T variant in MH1-associated embryonic lethality.

By analyzing the effect of this mutation on NRR56 and by genotyping embryos from mating at risk, we demonstrated that mortality in homozygotes occurs between 7 and 35 d of gestation. At 7 d, homozygous mutant blastocysts were normal in appearance, indicating that they had undergone normal early embryonic

---

**Table 6.** Estimation of loss in nonreturn rate at 56 d (NRR56) in mating at risk between carriers of the PFAS p.R1205C substitution

<table>
<thead>
<tr>
<th>Mating at risk</th>
<th>Category</th>
<th>No. of matings at risk</th>
<th>Proportion of mating at risk compared with all matings (%)</th>
<th>Loss in NRR56 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier × daughter of a carrier</td>
<td>Heifer</td>
<td>76,723</td>
<td>4.06</td>
<td>−4.77</td>
</tr>
<tr>
<td></td>
<td>Cow</td>
<td>193,390</td>
<td>4.51</td>
<td>−4.94</td>
</tr>
<tr>
<td>Carrier × carrier</td>
<td>Heifer</td>
<td>774</td>
<td>1.49</td>
<td>−11.15</td>
</tr>
<tr>
<td></td>
<td>Cow</td>
<td>797</td>
<td>1.38</td>
<td>−12.58</td>
</tr>
</tbody>
</table>

1 All losses in NRR56 are significant ($P < 0.05$).

**Table 7.** Genotypes of 7-d embryos produced from heterozygous carriers of the PFAS p.R1205C substitution

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Embryos observed (no.)</th>
<th>Embryos expected supposing recessive embryonic lethality before 7 d (no.)</th>
<th>Embryos expected supposing no recessive embryonic lethality before 7 d (no.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1205/R1205</td>
<td>2</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>R1205/C1205</td>
<td>11</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>C1205/C1205</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
</tbody>
</table>
embryonic development in humans. Mutations in pregnancy and a potential effect of cattle provides an interesting model for investigating observed; no homozygote). Thus, the discovery of this allele, which segregates at a very low frequency in database the same p.R1205C substitution in the PFAS we found in the human Exome Aggregation Consortium detectable forms (Baresova et al., 2016). Interestingly, prenatal or early death in their most severe and hardly investigated as potential candidates for severe metabolic disorders in humans, which are supposed to cause prenatal or early death in their most severe and hardly detectable forms (Baresova et al., 2016). Interestingly, we found in the human Exome Aggregation Consortium database the same p.R1205C substitution in the PFAS protein (GRC/hg19; Chr17:8172801C>T; Lek et al., 2016), which segregates at a very low frequency in European and Asian populations (4 alleles over 117,110 observed; no homozygote). Thus, the discovery of this recessive lethal candidate mutation in Montbéliarde cattle provides an interesting model for investigating a potential effect of PFAS mutations in pregnancy and embryonic development in humans.

**CONCLUSIONS**

Combining a genome scan for homozygous haplotype deficiency with WGS data analysis is a powerful approach for quickly identifying embryonic lethal mutations. However, it is of primary importance to statistically confirm the perfect linkage disequilibrium between the candidate mutations identified and the embryonic lethal phenotype by genotyping large populations. Here, after eliminating a previous candidate mutation in the SHBG gene and analyzing new WGS data, we propose a strong candidate mutation (g.28511199C>T; rs455876205; PFAS p.R1205C) for recessive embryonic lethality associated with the MH1 haplotype in Montbéliarde dairy cattle. The absence of homozygous mutants among tens of thousands of genotyped animals, the perfect conservation of the affected residue among eukaryotes, and the critical function of PFAS and of the de novo purine synthesis pathway are convincing arguments for its embryonic lethal effect. This discovery provides the optimal tool for carrying out an efficient selection against this embryonic lethal defect present at a high frequency in Montbéliarde and, to some extent, Simmental dairy cattle after recent admixture with Montbéliarde.

**ACKNOWLEDGMENTS**

This study is part of the BOVANO project (ANR-14-CE19-0011) funded by the French Agence Nationale de la Recherche (Paris, France) and APIS-GENE (Paris, France). The authors also acknowledge the financial contribution of the European Union’s Research and Innovation funding program EU-FP7 (FECUND grant no. 312097-2013-2017). P. Michot is a recipient of a CIFRE PhD grant from Allice (Paris, France), with the financial support of the French Association Nationale de la Recherche et de la Technologie (Paris, France) and APIS-GENE. The authors are grateful to the partners of the 1000 bull genomes consortium for their excellent collaboration.

**REFERENCES**


Alexiou, M., and H. J. Leese. 1992. Purine utilisation, de novo AMP activation. In such a situation, partial PFAS activity could be sufficient to meet purine requirements during early embryonic development but not during blastocyst elongation and trophoblast differentiation. Implantation of homozygous mutants and control embryos, close monitoring of gestations, and gene expression analyses would be required to test these hypotheses and further understand molecular mechanisms causing the death of homozygous mutants.

Finally, mutations affecting genes belonging to the purine de novo and salvage synthesis pathways are investigated as potential candidates for severe metabolic disorders in humans, which are supposed to cause prenatal or early death in their most severe and hardly detectable forms (Baresova et al., 2016). Interestingly, we found in the human Exome Aggregation Consortium database the same p.R1205C substitution in the PFAS protein (GRC/hg19; Chr17:8172801C>T; Lek et al., 2016), which segregates at a very low frequency in European and Asian populations (4 alleles over 117,110 observed; no homozygote). Thus, the discovery of this recessive lethal candidate mutation in Montbéliarde cattle provides an interesting model for investigating a potential effect of PFAS mutations in pregnancy and embryonic development in humans.

CONCLUSIONS

Combining a genome scan for homozygous haplotype deficiency with WGS data analysis is a powerful approach for quickly identifying embryonic lethal mutations. However, it is of primary importance to statistically confirm the perfect linkage disequilibrium between the candidate mutations identified and the embryonic lethal phenotype by genotyping large populations. Here, after eliminating a previous candidate mutation in the SHBG gene and analyzing new


