J. Dairy Sci. 93:1713–1721
doi:10.3168/jds.2009-2438

Identification of single nucleotide polymorphisms in the bovine solute carrier family 11 member 1 (*SLC11A1*) gene and their association with infection by *Mycobacterium avium* subspecies *paratuberculosis*

O. Ruiz-Larrañaga,* J. M. Garrido,† C. Manzano,* M. Iriondo,* E. Molina,† A. Gil,* A. P. Koets,‡ V. P. M. G. Rutten,§# R. A. Juste,† and A. Estonba*[1]

*Genetics, Physical Anthropology, and Animal Physiology Department, University of the Basque Country, Sarriena s/n, 48940 Leioa, Bizkaia, Spain
†Animal Health Department, NEIKER-Teknalia, Berreaga 1, 48160 Derio, Bizkaia, Spain
‡Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, the Netherlands
§Institute of Infectious Diseases and Immunology, Department of Immunology, Faculty of Veterinary Medicine, Utrecht University, the Netherlands
#Department of Veterinary Tropical Diseases, Faculty of Veterinary Science,University of Pretoria, Private Bag X04, Onderstepoort 0110, Republic of South Africa

ABSTRACT

Johne's disease is a chronic enteritis caused by *Mycobacterium avium* ssp. *paratuberculosis* (MAP) that causes substantial financial losses for the cattle industry. Susceptibility to MAP infection is reported to be determined in part by genetic factors, so marker-assisted selection could help to obtain bovine populations that are increasingly resistant to MAP infection. Solute carrier family 11 member 1 (*SLC11A1*) was adjudged to be a potential candidate gene because of its role in innate immunity, its involvement in susceptibility to numerous intracellular infections, and its previous association with bovine MAP infection. The objectives of this study were to carry out an exhaustive process of discovery and compilation of polymorphisms in *SLC11A1* gene, and to perform a population-based genetic association study to test its implication in susceptibility to MAP infection in cattle. In all, 57 single nucleotide polymorphisms (SNP) were detected, 25 of which are newly described in *Bos taurus*. Twenty-four SNP and two 3′-untranslated region polymorphisms, previously analyzed, were selected for a subsequent association study in 558 European Holstein-Friesian animals. The SNP c.1067C > G and c.1157–91A > T and a haplotype formed by these 2 SNP yielded significant association with susceptibility to MAP infection. The c.1067C > G is a nonsynonymous SNP that causes an amino acid change in codon 356 from proline to alanine (P356A) that could alter SLC11A1 protein function. This association study supports the involvement of *SLC11A1* gene in susceptibility to MAP infection in cattle. Our results suggest that SNP c.1067C > G may be a potential causal variant, although functional studies are needed to assure this point.

Key words: paratuberculosis, solute carrier family 11 member 1 (*SLC11A1*), single nucleotide polymorphism, haplotype

INTRODUCTION

Johne's disease, caused by *Mycobacterium avium* ssp. *paratuberculosis* (**MAP**), is a chronic enteritis characterized by progressive weight loss and profuse diarrhea. Johne's disease occurs worldwide and is highly prevalent in domestic ruminants, approximately 20% in cattle of several European countries (Nielsen and Toft, 2009). It is currently recognized as one of the main diseases of dairy cattle in industrialized countries and results in substantial financial losses every year for the dairy industry (Hasonova and Pavlik, 2006). Beyond this financial damage, bovine MAP infection may pose a public health risk, as MAP has been linked to Crohn's disease in humans (Juste et al., 2009).

Currently, there is no treatment for Johne's disease and the only control system that consistently yields good results is vaccination, which has proved notably successful in sheep and cattle (Köhler et al., 2009). Vaccination for MAP has met resistance in some countries because of its potential interference with the diagnosis of bovine tuberculosis, but it is highly beneficial financially because it reduces the number of clinically and subclinically infected animals (Juste et al., 2002). However, vaccination prevents only disease and not infection, so new strategies are required to help eradicate the infection and reduce its financial impact.

Genetic factors are involved in intracellular infections; MAP infection heritability is estimated to be in a moderate range, with 0.102 being the most reliable value (Koets et al., 2000; Mortensen et al., 2004; Gonda

Received June 1, 2009.
Accepted November 18, 2009.
[1]Corresponding author: andone.estonba@ehu.es

et al., 2006). The use of pre-existing genetic factors, selecting against susceptible animals by marker-assisted selection combined with classical breeding programs (Dekkers, 2004), may be a viable alternative in combating Johne's disease. To implement such a technique, genetic markers associated with susceptibility to MAP infection must be identified. Some of the studies conducted to date point to various chromosome regions or QTL associated with paratuberculosis (Gonda et al., 2007; Settles et al., 2009), and different genes have been highlighted as potentially involved in the mechanism of susceptibility to MAP infection (Estonba et al., 2005; Mucha et al., 2009; Pinedo et al., 2009a,b).

The solute carrier family 11 member 1 (***SLC11A1***, formerly *NRAMP1*) gene, is expressed mainly in the phagosomes of cells belonging to the immune system, such as macrophages and neutrophils. The role of the SLC11A1 protein is to prevent intracellular bacterial growth. The mechanism seems to be linked to the transporting of divalent metal ions, mainly Mn^{2+} and Fe^{2+}, with protein SLC11A1 playing a leading role in the cellular recycling of the latter (Soe-Lin et al., 2008). The involvement of the *SLC11A1* gene in infections by intracellular pathogens including mycobacteria has been shown in mice, humans, and various domestic animals (Vidal et al., 1993; Gazouli et al., 2008; Sanchez-Robert et al., 2008). Recently, a strong genetic influence of *Slc11a1* on the innate susceptibility of mice to infection with MAP has been established (Roupie et al., 2008). For cattle, a genetic association has been described between a microsatellite-type polymorphism $(GT)_n$ in the 3′-untranslated region (**UTR**) of the gene and *Brucella abortus* infection in Holstein-Friesian cattle (Adams and Templeton, 1998). In the same region of the gene, Estonba et al. (2005) analyzed a larger fragment described previously (Hořín et al., 1999), which includes the $(GT)_n$ microsatellite described by Adams and Templeton (1998) and a further adjacent $(GT)_n$ repeat. This study detected a genetic association with MAP infection in a naturally infected Holstein-Friesian herd.

The present study, however, is not only focused in one region of *SLC11A1* but goes further in an attempt to cover the whole functional variability of the *SLC11A1* gene in cattle, including all exons and their flanking intronic regions, UTRs, and part of the promoter region. With the aim of testing for its potential involvement in susceptibility to MAP infection, we conducted a SNP discovery approach on bovine *SLC11A1* gene and performed a candidate gene type genetic association test between these SNP and MAP infection in the Holstein-Friesian breed. In this study we also sought to establish the relationship between microsatellite $(GT)_n$ of Adams and Templeton (1998), the fragment analyzed by Estonba et al. (2005) in the same 3′-UTR, and the SNP covering the whole length of the *SLC11A1* gene.

MATERIALS AND METHODS

SNP Discovery and Selection

Comparative sequencing was used to identify novel SNP polymorphisms in *SLC11A1*. The gene was divided into 11 fragments (N1_1 to N1_11) covering all the exons, flanking intronic regions, promoter, and UTR (Supplemental Table 1; available online at http://www.journalofdairyscience.org/). Primers were designed with the Primer3 program, using Btau_3.1; ENSBTA00000015520 sequence from the Ensembl database (http://www.ensembl.org/index.html) as reference. Fragments N1_1 to N1_10 were sequenced in 15 individuals from 14 bovine breeds: Blonde, Limousin, Holstein-Friesian (2), Gelbvieh, Red Angus, Jersey, Guernsey, Salers, Pirenaica, Terreña, Betizu, Monchina, Beefmaster, and Brangus. The SNP found were then validated in a sample of 85 Holstein-Friesians. Because the N1_11 fragment includes several microsatellites, it was cloned and then sequenced in 17 Holstein-Friesians using a TOPO TA Cloning kit (Invitrogen, Carlsbad, CA) to avoid sequencing errors.

DNA was purified from blood samples with a QIAamp Mini Kit (Qiagen, Hilden, Germany). Fragments were amplified in a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA) and sequenced with a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) with both forward and reverse primers. Sequences were detected using 3100 Avant and 3130xl Genetic Analyzer equipment and analyzed with SeqScape v2.5 software (Applied Biosystems) to discover SNP.

As a complementary strategy, nucleotide sequences for the bovine *SLC11A1* gene stored in the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/) nucleotide database were compared with the reference sequence, and those SNP not described previously were noted. Finally, *SLC11A1* SNP from IBISS3 (Hawken et al., 2004), NCBI dbSNP, Ensembl, and Animal Genome databases were also compiled (http://www.livestockgenomics.csiro.au/IBISS3/; http://www.ncbi.nlm.nih.gov/snp/; http://www.ensembl.org/index.html; http://www.animalgenome.org/, respectively). For creating the SNPlex marker set for the association study, SNP were selected according to their frequencies in the 85 Holstein-Friesian samples and their methodological compatibilities.

Genetic Association Study

The population for the genetic association study was made up of 558 European Holstein-Friesian (**HF**) cows from 2 independent subpopulations. We identified cases and controls from a Spanish population of 243 HF animals belonging to 33 herds. All these samples were received at NEIKER (Derio, Spain) for confirmation of a clinical suspicion of paratuberculosis in the herd and had no further follow up. Animals were classified as infected (cases) if they had a positive fecal culture, PCR, or blood serum ELISA result. Animals with negative results for all 3 tests were considered healthy (controls). Independently, cases and controls were identified in a cohort of 315 HF animals of 8 Dutch dairy farms. These animals were sampled at a minimum of 4 time points and both fecal culture and blood serum ELISA tests were performed. A case was defined as having at least one positive fecal culture result. Animals tested at least 4 times and negative to all tests were considered controls.

Fecal cultures were prepared using the protocol described by Aduriz et al. (1995). The sensitivity and specificity of fecal culture are 60 and 99.9%, respectively (Collins et al., 2006). Isolation of MAP DNA from fecal samples was performed using a MAP DNA extraction and purification commercial kit (QIAamp DNA Blood minikit, Qiagen), and the purified DNA sequences were tested by real-time PCR (Adiagene, Saint Brieuc, France), which amplified a *Map* IS900. The sensitivity for the fecal PCR is highly dependent on the fecal shedding status (32% for light shedders to 91% for heavy shedders) but the overall sensitivity of the test is 60% and specificity is 100% (Taddei et al., 2004). The blood samples were analyzed by using a paratuberculosis indirect ELISA (Institut Pourquier, Montpellier, France) that used PPA3 as antigen. The sensitivity of ELISA is highly dependent on the stage of MAP infection but the overall sensitivity of the test is 30 to 45% and specificity is 99% (Sweeney et al., 1995; Collins et al., 2005). The final HF population consisted of 129 infected and 114 healthy animals in the Spanish subpopulation (**HFS**) and 138 infected and 177 healthy in the Dutch subpopulation (**HFN**).

All animals analyzed in this study were adults (2 yr of age and over). The average prevalence of infection in the herds of origin was 6% in Spain and 32.2% in the Netherlands. Paratuberculosis infection has been circulating unchecked in European ruminant populations at least since its first report in 1895, so this study was done under the assumption that MAP is widespread and that once MAP infection is present in a herd, all animals are potentially exposed to it.

To test for population stratification, 17 neutral microsatellites were genotyped in 100 unrelated cows, 50 from each subpopulation (25 infected and 25 healthy). Comparisons of allelic frequencies were made between infected and healthy subgroups, in both HFS and HFN subpopulations and in the European HF global population. Results did not indicate any potential stratification for bias association analyses (all P-values > 0.003 after Bonferroni correction).

DNA was extracted from blood samples using a QIAamp Mini Kit (Qiagen) and a Promega Wizard Genomic DNA kit (Promega, Madison, WI). The fragment described by Hořín et al. (1999) was amplified using the primers and PCR conditions described by the same author. The $(GT)_n$ microsatellite of Adams and Templeton (1998) was analyzed using the primers and PCR conditions described by Barthel et al. (2000). The SNP were genotyped using SNPlex technology by Applied Biosystems. The 2 UTR polymorphisms and the SNP were detected using 3100 Avant and 3130xl Genetic Analyzer equipment, and alleles were assigned using GeneMapper v3.7 software (Applied Biosystems).

Statistical Analysis

Departures from Hardy-Weinberg equilibrium (**HWE**) for the 3′-UTR polymorphisms were evaluated by exact test implemented in the GENEPOP v3.3 package (Rousset, 2008). Running the same program, we used Fisher's exact test to evaluate the genetic association between length polymorphisms and MAP infection and to estimate the genotypic disequilibrium between 3′-UTR polymorphisms and SNP, based on contingency tables. The haplotype phases between the 3′-UTR and SNP were reconstructed from population genotype data using the PHASE v2.1 program (Stephens et al., 2001) by a Bayesian statistical method.

Conformity of SNP genotype proportions to HWE, allelic frequencies, assessment of linkage disequilibrium for each SNP pair, and genetic association analyses were performed using Haploview version 4.1 statistical software (Barrett et al., 2005). Linkage disequilibrium was established by D' and r^2 parameters. The confidence intervals method was used to define haplotype blocks (Gabriel et al., 2002), and haplotypes were estimated using an expectation maximization algorithm based on the maximum likelihood. Analyses of genetic association between SNP or haplotypes and MAP infection were conducted only with SNP with minor allele frequency > 0.05 and in HWE. A Chi-square test was performed, and P-values were calculated for the allelic and haplotype frequencies in infected versus healthy individuals. To correct for the occurrence of false positives (type

I error), permutation procedures were performed as implemented in Haploview (1,000,000 permutations). This approach corrects for multiple testing but takes into account the correlation between markers. Permutation correction is thus less conservative than Bonferroni correction but it is appropriate for independent tests with multiple markers (Camargo et al., 2008). The odds ratios (**OR**) for associated alleles and haplotypes comparing infected versus healthy animals were calculated with a 95% confidence interval and the statistical tests of Cochran and Mantel-Haenszel using SPSS v16.0 (SPSS, Chicago, IL). The OR for haplotypes were obtained by testing the associated haplotypes against the rest of the haplotypes.

Finally, the Plink v 1.02 software package (Purcell et al., 2007; http://pngu.mgh.harvard.edu/purcell/plink) was used to analyze the genotypic association of SNP with MAP infection by a Chi-square test. This software offers the ability to perform 4 statistical tests simultaneously: Cochran-Armitage trend test for the additive allele effects, genotypic test (DD vs. Dd vs. dd), dominant (DD, Dd vs. dd), and recessive (DD vs.Dd, dd) models, D and d being the minor and major alleles, respectively. The *P*-values for these genotypic models were corrected by permutations as described above.

In Silico SLC11A1 Gene Analysis

For comparative analysis, the *SLC11A1* reference sequences from 6 ruminant species (*Bos taurus*, *Bos indicus*, *Ovis aries*, *Bubalus bubalis*, *Bison bison*, and *Cervus elaphus*) and 8 other mammals (*Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Pan troglodytes*, *Macaca mulatta*, *Sus scrofa*, *Equus caballus*, and *Canis familiaris*) were retrieved from the GeneBank database. Nucleotide sequences were aligned using ClustalW Multiple Alignment option from BioEdit Sequence Alignment Editor (http://www.mbio.ncsu.edu/BioEdit/BioEdit.html). Associated SNP positions were examined to assign the putative ancestral nucleotide codon.

RESULTS

3'-UTR Polymorphisms

The fragment described by Hořín et al. (1999) was analyzed in 536 animals, and 4 alleles of 211, 213, 215, and 217 bp were detected. Subpopulations HFS and HFN were in HWE. Allele 213 was predominant, whereas allele 211, associated with a lower susceptibility (Estonba et al., 2005) had a low frequency ($freq_{HFS}$: 0.079 and $freq_{HFN}$: 0.091). The differences in allele frequencies between infected and healthy animals were not significant (Supplemental Table 2; available online at http://www.journalofdairyscience.org/).

A total of 543 animals were analyzed for the microsatellite of Adams and Templeton (1998). Three alleles of 175, 177, and 179 bp were detected, which correspond to $(GT)_{13}$, $(GT)_{14}$, and $(GT)_{15}$, respectively (Paixao et al., 2006). Both subpopulations were in HWE. Allele $(GT)_{13}$, described as associated with a higher resistance to *Brucella abortus* infection (Adams and Templeton, 1998), predominated in both HFS and HFN ($freq_{HFS}$: 0.972 and $freq_{HFN}$: 0.983), whereas $(GT)_{14}$ and $(GT)_{15}$ had frequencies <0.02. Comparisons between infected and healthy animals showed no significant differences between allele distributions (Supplemental Table 2; available online at http://www.journalofdairyscience.org/).

SNP Discovery and Genetic Association Study

In total, 7,194 bp of the 12,164 bp that make up the bovine *SLC11A1* gene were sequenced. From the different strategies used to define SNP, 57 SNP were compiled over the entire gene, of which 25 were new SNP not previously described. All SNP were submitted to GenBank under accession numbers ss119336718 to ss119336751 and ss159815965 to ss159815987. Twenty-four SNP out of 57 were selected for genotyping of infected and healthy animals, according to their presence in the Holstein-Friesian panel and methodological compatibilities (Supplemental Table 3; available online at http://www.journalofdairyscience.org/). These SNP were genotyped in 486 animals (240 infected and 246 healthy). Sixteen proved to be polymorphic (Figure 1): 13 in HFS and 12 in HFN. All of them were in HWE, although only 2 (c.1067C > G and c.1157–91A > T) showed a minor allele frequency ≥0.05 (Supplemental Table 4; available online at http://www.journalofdairyscience.org/), and thus, were selected for the association analysis. Therefore, this study cannot prove that the SNP with minor allele frequency ≤0.05 do not affect MAP infection.

In HFS, intronic SNP c.1157–91A > T was genetically associated with susceptibility to MAP infection (corrected $P = 0.036$; Table 1), because its A allele had a significantly higher frequency in infected animals than in healthy ones. The OR for this allele comparing its presence in infected versus healthy animals also indicated a positive association with a higher risk to infection (OR = 1.721; CI 95% = 1.071–2.765; Cochran test $P = 0.024$; Mantel-Haenszel test $P = 0.033$). The C allele of exonic SNP c.1067C > G showed the same tendency toward association with susceptibility (nominal $P = 0.037$), although its corrected *P*-value fell short of significance (corrected $P = 0.052$). These 2 SNP were in linkage disequilibrium (D' = 1.0; $r^2 = 0.765$), forming a haplotype block for which 3 haplotypes were

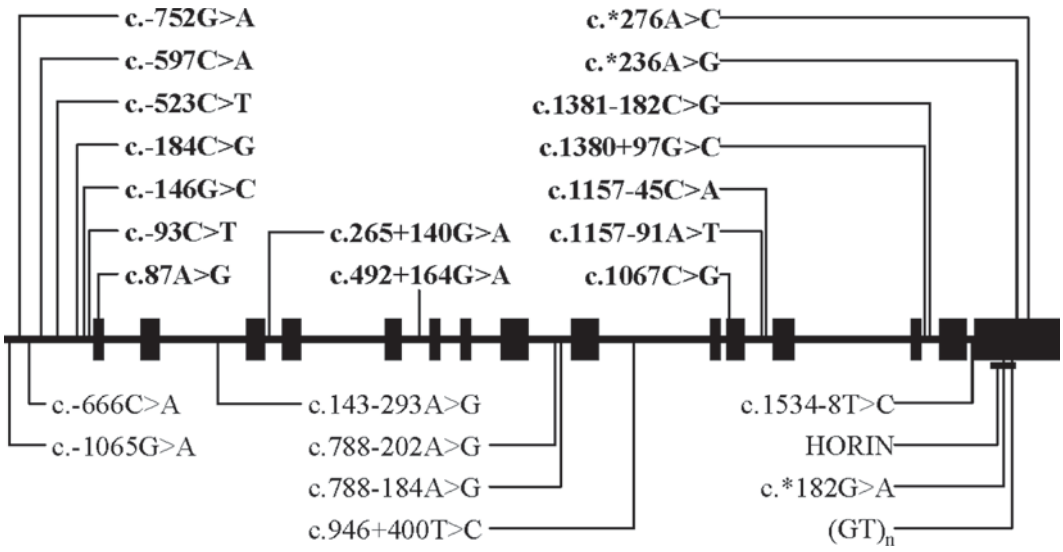

Figure 1. Genomic structure and analyzed polymorphisms of the bovine solute carrier family 11 member 1 (*SLC11A1*) gene. Black boxes indicate exons, and lines connecting them indicate introns; 3′ and 5′ untranslated regions (UTR) are included in exons 1 and 15, respectively. Polymorphic SNP in Spanish (HFS) and Dutch (HFN) Holstein-Friesian subpopulations are shown in bold.

detected: GT, GA, and CA (Table 1). The *P*-value for haplotype GT indicated a tendency toward association with a higher resistance to MAP infection (nominal $P = 0.024$; corrected $P = 0.062$), whereas the CA haplotype tended to be linked to susceptibility (nominal $P = 0.032$; corrected $P = 0.073$).

In HFN, no SNP significantly associated with the infection were detected, but it was observed that the 2 major alleles of SNP c.1067C > G and C.1157–91A > T (alleles C and A, respectively) were more frequent in infected animals than in healthy ones (Table 1), just as they were in HFS. The 2 SNP were in linkage disequilibrium (D' = 1.0; $r^2 = 0.905$), and the same haplotypes as in HFS could be observed (Table 1). No significant genetic association was detected between haplotypes and the infection (nominal $P > 0.05$).

Because the same tendency was observed in both HFS and HFN, a joint association analysis was conducted

Table 1. Chi-square and *P*-values from Haploview 4.1 (Barrett et al., 2005) for allelic and haplotypic frequencies of SNP c.1067C > G and c.1157–91A > T in infected and healthy animals of Spanish (HFS), Dutch (HFN), and European (HF) Holstein-Friesian populations

		SNP		Haplotype		
Item	n	c.1067C > G (allele C)	c.1157–91A > T (allele A)	CA	GT	GA
HFS	241					
Infected	127	0.815	0.857	0.814	0.143	0.042
Healthy	114	0.735	0.776	0.732	0.224	0.044
Chi-square		4.332	5.093	4.571	5.093	0.012
P-value		0.037*	0.024*	0.032*	0.024*	0.913
Corrected *P*-value[1]		0.052	0.036*	0.073	0.062	0.999
HFN	245					
Infected	113	0.911	0.925	0.907	0.075	0.018
Healthy	132	0.870	0.885	0.871	0.114	0.015
Chi-squared		2.004	2.151	1.576	2.072	0.047
P-value		0.156	0.142	0.209	0.150	0.828
Corrected *P*-value		0.220	0.206	0.379	0.317	0.981
HF	486					
Infected	240	0.861	0.889	0.859	0.111	0.030
Healthy	246	0.807	0.835	0.807	0.165	0.029
Chi-squared		5.009	6.007	4.736	5.882	0.020
P-value		0.025*	0.014*	0.029*	0.015*	0.886
Corrected *P*-value		0.037*	0.023*	0.081	0.039*	0.985

[1]*P*-value corrections by 1,000,000 permutations.

*$P < 0.05$.

to get a larger sample size and increase the statistical power of the study. All the SNP were in HWE in this European HF population (n = 486). For the 2 SNPs with minor allele frequency >0.05, both the C allele of SNP c.1067C > G and the A allele of SNP c.1157–91A > T were associated with susceptibility (Table 1). The GT haplotype formed by the 2 alternative alleles of these 2 SNP was found to be significantly more frequent in healthy animals than in infected animals (corrected $P = 0.039$). The OR comparing infected versus healthy for allele C (SNP c.1067C > G) and allele A (SNP c.1157–91A > T) separately, and for the GT haplotype, respectively, were also significant: $OR_{SNP\ c.1067C > G} = 1.484$ (CI 95% = 1.049–2.099; Cochran test $P = 0.025$; Mantel-Haenszel test $P = 0.032$); $OR_{SNP\ c.1157–91A > T} = 1.592$ (CI 95% = 1.095–2.314; Cochran test $P = 0.014$; Mantel-Haenszel test $P = 0.018$); and $OR_{GT\ haplotype} = 0.631$ (CI 95% = 0.434–0.917; Cochran test $P = 0.015$; Mantel-Haenszel test $P = 0.020$). These OR values indicated the relation between alleles C and A and a higher susceptibility to MAP infection, and the protective effect of the GT haplotype. Moreover, the CA haplotype was found to be more frequent in infected animals (nominal $P = 0.029$). Although the difference in frequency is not significant (corrected $P = 0.081$), the OR indicates a possible link with susceptibility to infection (OR = 1.465; CI 95% = 1.039–2.065; Cochran test $P = 0.029$; Mantel-Haenszel test $P = 0.036$). The genotypic association analysis conducted on this European HF population resulted in a best significance value for the dominant model (GG + CG vs. CC in SNP c.1067C > G and TT+AT vs. AA in SNP c.1157–91A > T; corrected P-values of 0.020 and 0.015, respectively).

Linkage Disequilibrium Between 3'-UTR Polymorphisms and SNP

The genotypic disequilibrium analysis indicates that the fragment described by Hořín et al. (1999) was linked to the haplotype block formed by SNP c.1067C > G and c.1157–91A > T ($P < 0.00001$; Supplemental Table 5; available online at http://www.journalofdairyscience.org/). Reconstruction of haplotype phases revealed that 95.8% of the 211-bp alleles (described previously as a resistance allele) were in phase with haplotype GT, whereas the rest were in phase with CA. With regard to the 213-bp alleles, 92.0% of them were in phase with haplotype CA, and 8% with GT or GA. The microsatellite of Adams and Templeton (1998) also showed genotypic disequilibrium with this haplotype block ($P < 0.00001$). Among the $(GT)_{13}$ alleles, 86.6% were in phase with CA, whereas 12.3% were in phase with GT. Finally, the 15 $(GT)_{14}$ alleles and 7 of the 8 $(GT)_{15}$ alleles were in phase with haplotype GA.

In Silico SLC11A1 Gene Analysis

The SNP c.1067C > G is located in exon 11 and is a nonsynonymous SNP, where the minor allele G results in an alanine at position 356 of SLC11A1 protein (GCA codon), whereas the C variant results in a proline (CCA). Sequence alignment of 6 ruminants and 8 other mammalian species suggests that the ancestral codon was GCG (alanine). *Sus scrofa* and the ruminants analyzed seem to have mutated to GCC and GCA, respectively, maintaining the alanine in that position. It is in genus *Bos* that GCA seems to have changed to CCA (proline).

DISCUSSION

The strategy used to detect and select SNP in bovine *SLC11A1* proved successful: 57 SNP were compiled, 25 of which (identified by comparative sequencing) are described in *Bos taurus* for the first time in this paper. Sixteen SNP out of the 24 selected for the association study proved to be polymorphic in Holstein-Friesians, although most of them showed a very low frequency for the minor allele, possibly because of a reduction in genetic diversity caused by a high level of artificial selection (Brotherstone and Goddard, 2005; Zenger et al., 2007). It is worth pointing out that the SNP c.-93C > T, c.87A > G, and c.1067C > G were described as monomorphic for the Holstein-Friesian breed in an earlier study (Martinez et al., 2008), probably because of the sample size (n = 10).

Previous references to the role of the *SLC11A1* gene in susceptibility to intracellular infections in cattle were limited to the association found by analyzing the variability of the 3′-UTR of the gene. The present study supports the role of the innate immunity gene *SLC11A1* in processes of this kind. A significant genetic association was detected between 2 SNP located at exon 11 and intron 11–12 and susceptibility to infection by MAP in Holstein-Friesian cattle. The C allele of c.1067C > G and the A allele of c.1157–91A > T were significantly in excess among infected European Holstein-Friesian animals. The genotypic association tests point to a dominant effect of the minor alleles (G of c.1067C > G and T of c.1157–91A > T). This suggests that the major alleles (C and A, respectively) confer susceptibility only in homozygosis. Nevertheless, these results should be taken with caution as the number of homozygote individuals for the minor alleles was low (Supplemental Table 6; available online at http://www.journalofdairyscience.org/).

As for the putative functional effect of these alleles, it is striking that c.1067C > G is located in exon 11, which encodes for transmembrane domain 8 (**TM8**), the most conserved region of the 12 putative transmem-

brane domains of the SLC11A1 protein (Feng et al., 1996). Indeed, SNP c.1067C > G is the only variable nucleotide that we found in the TM8 domain. The C allele, associated in the present study with susceptibility to MAP infection, results in a proline at AA position 356 of SLC11A1 (CCA codon), whereas the G allele results in an alanine (GCA codon). Sequence alignment analysis suggests that GCG (alanine) is the ancestral codon in mammals and it seems to have mutated to GCA (alanine) in ruminants. In genus *Bos* it would have changed to CCA, resulting in an AA substitution to proline in both *Bos taurus* and *Bos indicus*. In other words, because the G allele seems to be the ancestral variant in mammals and ruminants, one could speculate that the high prevalence of the C allele in Holstein-Friesians and other *Bos taurus* breeds (Martinez et al., 2008) could be caused by a prolonged artificial selection pressure, being the locus in close linkage with a nearby allele favorable to yield traits. In fact, the lowest frequency of allele C is found in *Bos indicus* (Martinez et al., 2008). It is known that artificial selection for productive traits may be detrimental to traits linked to the health of individual animals, such as resistance to infectious diseases (Detilleux, 2001), and this could agree with the idea of Holstein-Friesian cattle being more susceptible to different infections than Zebu cattle. In any event, it must be stressed that the effect of other factors on the frequency of the C allele, such as the age of the mutation or genetic drift, may be equally likely.

Various studies establish functional alterations related to proline or to alanine changes: the change in the functional binding domain of peroxisome proliferator-activated receptor-gamma 2 (PPARG2) protein, where alanine is related to resistance to development of diabetes in humans (Meshkani et al., 2007); the change in the catalytic subunit of 1,3-β-D-glucan synthase (Fks1p) gene of *Candida parapsilosis* that gives resistance to antifungal drugs (Garcia-Effron et al., 2008); and the proline to alanine change at the cyclooxygenase-2 (*COX-2*) gene associated with the risk of breast cancer (Li et al., 2009). Moreover, Pera et al. (2008) observed that the addition or omission of a proline or a charged AA in the solute carrier family 26 member 4 (SLC26A4) protein is detrimental to its function. Referring back to the *SLC11A1* gene, the nonconservative G to C change at c.1067C > G is predicted to generate a change in the secondary structure of the protein that would give rise to a helix instead of a strand structure in TM8 (Martinez et al., 2008). In this sense, bearing in mind that proline plays important structural as well as functional roles in membrane proteins (Joshi and Pajor, 2006), its presence could be considered to alter the stability or structure of SLC11A1 protein. This could influence the transport of Fe^{2+} and other divalent cations, resulting in a less effective control of the replication of MAP in macrophages by the host. Functional studies are needed to clarify the mechanism by which exonic SNP c.1067C > G might be involved in MAP susceptibility.

The other SNP associated with the infection (c.1157–91A > T) was found in intron 11–12. The presence of the allele associated with susceptibility to MAP infection could influence the control of *SLC11A1* gene expression, because intronic regions play an important role in posttranscriptional processes such as alternative splicing of exons, which can give rise to new isoforms (Bechtel et al., 2008). In this context, it must be pointed out that the existence of microRNA sequences derived from intronic regions is becoming increasingly important; however, we have not identified any microRNA or alternative splicing-related sequences in the vicinity of SNP c.1157–91A > T.

It is worth mentioning that found differences in *P*-value significance between subpopulations HFS and HFN, but in spite of this, the trends in the frequencies of alleles associated with susceptibility were maintained in both populations. These differences could be explained by one or more of several factors affecting the statistical power of the analysis: sample size, linkage disequilibrium patterns, linkage phase between the analyzed marker and the causal variant, genetic background, effect of each genetic variant on the trait, and allele frequency in each population. Among these possible explanations, allele frequency may be the main factor affecting *P*-values because the only significant association observed when the 2 populations are considered separately is that of c.1157–91A > T, which has a frequency twice as high in the HFS populations as in HFN. Our study also highlights the importance of sample size in association analysis: when the 2 subpopulations were merged into 1 (i.e., when we double the sample size), the trends observed in each population separately became significant.

A genetic association study based on haplotype blocks was performed to complete the analysis conducted with individual markers. The OR value found in this study for the CA haplotype, formed by c.1067C > G and c.1157–91A > T, indicates that this haplotype could be associated with susceptibility. Haplotype GT, in contrast, is in excess in healthy animals, so we consider that it contributes to reduce susceptibility to MAP infection in the European Holstein-Friesian population.

With regard to other polymorphisms in linkage disequilibrium with the haplotype formed by c.1067C > G and c.1157–91A > T, we must specially mention the fragment described by Hořín et al. (1999), and more specifically the allele 211 bp, for which Estonba et al. (2005) detected lower susceptibility to MAP infection.

Allele 211 bp corresponds to the combination of both the resistance allele to brucellosis $(GT)_{13}$ of Adams and Templeton (1998) and allele $(GT)_{10}$+G in the second microsatellite included in the fragment. Allele $(GT)_{13}$ shows genotypic disequilibrium with the 2 most frequent haplotypes (CA and GT) formed by c.1067C > G and c.1157–91A > T, but no association is detected for this microsatellite. With regard to the fragment of Hořín et al. (1999) we did not find genetic association, but we did find that allele 211 bp showed genotypic disequilibrium with haplotype GT, which was associated with a reduced susceptibility to MAP infection in the present study. Specifically, 95.8% of the alleles of 211 bp are in phase with haplotype GT. This linkage disequilibrium between allele 211 and haplotype GT could explain the genetic association detected by Estonba et al. (2005).

CONCLUSIONS

In summary, 2 SNP in exon 11 and intron 11–12 in the bovine *SLC11A1* gene showed a marginally significant association with susceptibility to MAP infection. Given that the procedure used in this study to identify SNP ensures that we have detected a significant part of the functional variability of this gene in the Holstein-Friesian breed, we have most probably identified a region in which the causative variant could be located. Furthermore, among the variants tested within this particular gene, the nonsynonymous SNP c.1067C > G (P356A) has the greatest probability of being the causative variant, which when interacting with other risk alleles and environmental factors might be contributing to the susceptibility to MAP infection. Further studies including resequencing of the wider genomic region around this gene, genotyping of other SNP around the location of c.1067C > G and c.1157–91A > T, and application of a more comprehensive statistical model for analysis will help us to clarify this issue. The results of this study could be considered as a first step toward implementing a marker-assisted selection method for MAP resistance in breeding programs. However, it must not be forgotten that Johne's disease is a complex intracellular infectious disease, and as such is controlled by many different genes that interact with one another and with environmental factors. The results of this and future studies seeking to identify further genes associated with MAP infection may help to reduce the impact of bovine paratuberculosis on the livestock industry.

ACKNOWLEDGMENTS

We acknowledge the helpful comments and thoughtful suggestions of the anonymous referees. The authors also thank Fernando Rendo and the Sequencing and Genotyping Facilities from the University of the Basque Country (Leioa, Spain) for technical assistance. This study was funded by the Spanish Ministry of Education and Science (MEC, AGL2006-14315-C02-02/GAN), the University of the Basque Country, *Bizkaiberri* program of Biscay Regional Government, S. L. Aberekin (Aberekin Insemination Centre), and the Agriculture and Fishing Department of the Basque Government.

REFERENCES

Adams, L. G., and J. W. Templeton. 1998. Genetic resistance to bacterial disease of animals. Rev. Sci. Tech. 17:200–219.

Aduriz, J. J., R. A. Juste, and N. Cortabarria. 1995. Lack of mycobactin dependence of mycobacteria isolated on Middlebrook 7H11 from clinical cases of ovine paratuberculosis. Vet. Microbiol. 45:211–217.

Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265.

Barthel, R., J. A. Piedrahita, D. N. McMurray, L. Payeur, D. Baca, F. S. Güemes, V. S. Perumaalla, T. A. Ficht, J. W. Templeton, and L. G. Adams. 2000. Pathologic findings and association of *Mycobacterium bovis* infection with the bovine *NRAMP1* gene in cattle from herds with naturally occurring tuberculosis. Am. J. Vet. Res. 61:1140–1144.

Bechtel, J. M., P. Rajesh, I. Ilikchyan, Y. Deng, P. K. Mishra, Q. Wang, X. Wu, K. A. Afonin, W. E. Grose, Y. Wang, S. Khuder, and A. Fedorov. 2008. The alternative splicing mutation database: A hub for investigations of alternative splicing using mutational evidence. BMC Res. Notes 1:3.

Brotherstone, S., and M. Goddard. 2005. Artificial selection and maintenance of genetic variance in the global dairy cow population. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360:1479–1488.

Camargo, A., F. Azuaje, H. Wang, and H. Zheng. 2008. Permutation-based statistical tests for multiple hypotheses. Source Code Biol. Med. 3:15.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229:1912–1919.

Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12:685–692.

Dekkers, J. C. M. 2004. Commercial application of marker-and gene selection in livestock: Strategies and lessons. J. Anim. Sci. 82(E. Suppl.):E313–E328.

Detilleux, J. 2001. Genetic improvement of resistance to infectious diseases in livestock. J. Dairy Sci. 84(E. Suppl.)E39–E46.

Estonba, A., M. Iriondo, O. Oyanguren, J. M. Garrido, M. V. Geijo, I. Sevilla, and R. A. Juste. 2005. Bovine 3′-UTR (GT_{11}) variant of the *NRAMP1* gene under-represented among paratuberculosis seropositive cattle. Pages 46–49 in Proc. 8th Intl. Colloq. on Paratuberculosis, The Royal Veterinary and Agricultural University, Copenhagen, Denmark. E. J. B. Manning and S. S. Nielsen, ed. Frederiksberg, Denmark.

Feng, J., Y. Li, M. Hashad, E. Schurr, P. Gros, L. G. Adams, and J. W. Templeton. 1996. Bovine Natural Associated Macrophage Protein 1 (*Nramp1*) Gene. Genome Res. 6:956–964.

Gabriel, S. B., S. F. Schaffner, H. Nguyen, J. M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, S. N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E. S. Lander, M. J. Daly, and D. Altshuler. 2002. The structure of haplotype blocks in the human genome. Science 296:2225–2229.

Garcia-Effron, G., S. K. Katiyar, S. Park, T. D. Edlind, and D. S. Perlin. 2008. A naturally occurring proline-to-alanine acid change in

Fks1p in *Candida parapsilosis, Candida orthopsilosis,* and *Candida metapsilosis* accounts for reduced echinocandin susceptibility. Antimicrob. Agents Chemother. 52:2305–2312.

Gazouli, M., V. Atsaves, G. Mantzaris, M. Economou, G. Nasioulas, K. Evangelou, A. J. Archimandritis, and N. P. Anagnou. 2008. Role of functional polymorphisms of *NRAMP1* gene for the development of Crohn's disease. Inflamm. Bowel Dis. 14:1323–1330.

Gonda, M. G., Y. M. Chang, G. E. Shook, M. T. Collins, and B. W. Kirkpatrick. 2006. Genetic variation of *Mycobacterium avium* ssp. *paratuberculosis* infection in US Holsteins. J. Dairy Sci. 89:1804–1812.

Gonda, M. G., G. E. Shook, B. W. Kirkpatrick, and M. T. Collins. 2007. Quantitative trait locus affecting susceptibility to *Mycobacterium paratuberculosis* infection identified on BTA 20 in US Holsteins. Anim. Genet. 38:389–396.

Hasonova, L., and I. Pavlik. 2006. Economic impact of paratuberculosis dairy cattle herds: A review. Vet. Med. (Praha) 51:193–211.

Hawken, R. J., W. C. Barris, S. M. McWilliam, and B. P. Dalrymple. 2004. An interactive bovine in silico SNP database (IBISS). Mamm. Genome 15:819–827.

Hořín, P., I. Rychlík, J. W. Templeton, and L. G. Adams. 1999. A complex pattern of microsatellite polymorphism within the bovine *NRAMP1* gene. Eur. J. Immunogenet. 26:311–313.

Joshi, A. D., and A. M. Pajor. 2006. Role of conserved prolines in the structure and function of the Na^+/dicarboxylate cotransporter, NaDC1. Biochemistry 45:4231–4239.

Juste, R. A., N. Elguezabal, A. Pavon, J. M. Garrido, M. V. Guijo, I. Sevilla, J. L. Cabriada, A. Tejada, F. Garcia-Campos, R. Casado, I. Ochotorena, and A. Izeta. 2009. Association between *Mycobacterium avium* ssp. *paratuberculosis* DNA in blood and cellular and humoral immune response in inflammatory bowel disease patients and controls. Int. J. Infect. Dis. 13:247–254. doi:10.1016/j.ijid.2008.06.034.

Juste, R. A., M. V. Geijo, I. Sevilla, G. Aduriz, and J. M. Garrido. 2002. Control of paratuberculosis by vaccination. Page 331 in Proc. 7th Intl. Colloq. Paratuberculosis, Animal Health Department, NEIKER-Teknalia, Derio, Spain. R. A. Juste, M. V. Geijo, and J. M. Garrido, ed.

Koets, A. P., G. Adugna, L. L. G. Janss, H. J. van Weering, C. H. J. Kalis, G. H. Wentink, V. P. M. G. Rutten, and Y. H. Schukken. 2000. Genetic variation of susceptibility to *Mycobacterium avium* ssp. *paratuberculosis* infection in dairy cattle. J. Dairy Sci. 83:2702–2708.

Köhler, H., F. Gierke, T. Seidler, and M. Ziller. 2009. Control of paratuberculosis by vaccination: A systematic review and meta-analysis. Page 134 in Abstract Book of 10th Intl. Colloq. Paratuberculosis, University of Minnesota, St. Paul.

Li, F., G. S. Ren, H. Y. Li, X. Y. Wang, L. Chen, and J. Li. 2009. A novel single nucleotide polymorphism of the cyclooxygenase-2 gene associated with breast cancer. Clin. Oncol. 21:302–305. doi:10.1016/j.clon.2008.12.005.

Martinez, R., S. Dunner, G. Barrera, and J. Cañon. 2008. Novel variants within the coding regions of the *Slc11A1* gene identified in *Bos taurus* and *Bos indicus* breeds. J. Anim. Breed. Genet. 125:57–62.

Meshkani, R., M. Taghikhani, B. Larijani, Y. Bahrami, S. Khatami, E. Khoshbin, A. Ghaemi, S. Sadeghi, F. Mirkhani, A. Molapour, and K. Adeli. 2007. Pro12Ala polymorphism of the peroxisome proliferator-activated receptor-gamma2 (PPARgamma-2) gene is associated with greater insulin sensitivity and decreased risk of type 2 diabetes in an Iranian population. Clin. Chem. Lab. Med. 45:477–482.

Mortensen, H., S. S. Nielsen, and P. Berg. 2004. Genetic variation and heritability of the antibody response to *Mycobacterium avium* subspecies *paratuberculosis* in Danish Hosltein cows. J. Dairy Sci. 87:2108–2113.

Mucha, R., M. R. Bhide, E. B. Chakurkar, M. Novak, and I. Mikula. 2009. Toll-like receptors TLR1, TLR2 and TLR4 gene mutations and natural resistance to *Mycobacterium avium* ssp. *paratuberculosis* infection in cattle. Vet. Immunol. Immunopathol. 128:381–388.

Nielsen, S. S., and N. Toft. 2009. A review of prevalences of paratuberculosis in farmed animals in Europe. Prev. Vet. Med. 88:1–14.

Paixao, T. A., C. Ferreira, A. M. Borges, D. A. A. Oliveira, A. P. Lage, and R. L. Santos. 2006. Frequency of bovine *Nramp1* (*Slc11a1*) alleles in Holstein and Zebu breeds. Vet. Immunol. Immunopathol. 109:37–42.

Pera, A., S. Dossena, S. Rodighiero, M. Gandia, G. Botta, G. Meyer, F. Moreno, C. Nofziger, C. Hernandez-Chico, and M. Paulmichl. 2008. Functional assessment of allelic variants in the *SLC26A4* gene involved in Pendred syndrome and nonsyndromic EVA. Proc. Natl. Acad. Sci. USA 105:18608–18613.

Pinedo, P. J., C. D. Buergelt, G. A. Donovan, P. Melendez, L. Morel, R. Wu, T. Y. Langaee, and D. O. Rae. 2009a. Association between CARD15/NOD2 gene polymorphisms and paratuberculosis infection in cattle. Vet. Microbiol. 134:346–352.

Pinedo, P. J., C. D. Buergelt, G. A. Donovan, P. Melendez, L. Morel, R. Wu, T. Y. Langaee, and D. O. Rae. 2009b. Candidate gene polymorphisms (*BoIFNG, TLR4, SLC11A1*) as risk factors for paratuberculosis infection in cattle. Prev. Vet. Med. 91:189–196. doi:10.1016/j.prevetmed.2009.05.020.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A tool set for whole genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559–575.

Roupie, V., V. Rosseels, V. Piersoel, D. K. Zinniel, R. G. Barletta, and K. Huygen. 2008. Genetic resistance of mice to *Mycobacterium paratuberculosis* is influenced by *Slc11a1* at the early but not at the late stage of infection. Infect. Immun. 76:2099–2105.

Rousset, F. 2008. Genepop'007: A complete re-implementation of the genepop software for Windows and Linux. Mol. Ecol. Resour. 8:103–106.

Sanchez-Robert, E., L. Altet, M. Utzet-Sadurni, U. Giger, A. Sanchez, and O. Francino. 2008. *Slc11a1* (formerly *Nramp1*) and susceptibility to canine visceral leishmaniasis. Vet. Res. 39:36 doi:10.1051/vetres:2008013.

Settles, M., R. Zanella, S. D. Mckay, R. D. Schnabel, J. F. Taylor, R. Whitlock, Y. Schukken, J. S. Van Kessel, J. M. Smith, and H. Neibergs. 2009. A whole genome association analysis identifies loci associated with *Mycobacterium avium* ssp. *paratuberculosis* infection status in US Holstein cattle. Anim. Genet. doi:10.1111/j.1365-2052.2009.01896.x.

Soe-Lin, S., A. D. Sheftel, B. Wasyluk, and P. Ponka. 2008. *Nramp1* equips macrophages for efficient iron recycling. Exp. Hematol. 36:929–937.

Stephens, M., N. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68:978–989.

Sweeney, R. W., R. H. Whitlock, C. L. Buckley, and P. A. Spencer. 1995. Evaluation of a commercial enzyme-linked immunosornbent assay for the diagnosis of paratuberculosis in cattle. J. Vet. Diagn. Invest. 7:488–493.

Taddei, S., C. Robbi, C. Cesena, I. Rossi, E. Schiano, N. Arrigoni, G. Vicenzoni, and S. Cavirani. 2004. Detection of *Mycobacterium avium* ssp. *paratuberculosis* in bovine fecal samples: Comparison of three polymerase chain reaction-based test with a conventional culture method. J. Vet. Diagn. Invest. 16:503–508.

Vidal, S. M., D. Malo, K. Vogan, E. Skamene, and P. Gros. 1993. Natural resistance to infection with intracellular parasites: Isolation of a candidate for *Bcg*. Cell 73:469–485.

Zenger, K. R., M. S. Khatkar, J. A. Cavanagh, R. J. Hawkwn, and H. W. Raadsma. 2007. Genome-wide genetic diversity of Holstein Friesian cattle reveals new insights into Australian and global population variability, including impact of selection. Anim. Genet. 38:7–14.