ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12261

Wingless-type MMTV integration site family member 2 (*WNT2*) gene is associated with resistance to MAP in faecal culture and antibody response in Holstein cattle

A. Pauciullo*†, J. Küpper†, H. Brandt†, K. Donat‡, L. Iannuzzi* and G. Erhardt†

*Laboratory of Animal Cytogenetics and Gene Mapping, National Research Council (CNR) of Italy, ISPAAM, via Argine 1085, 80147 Naples, Italy. †Institute for Animal Breeding and Genetics, Justus-Liebig University, Ludwigstraße 21 B, 35390 Giessen, Germany. ‡Thüringer Tierseuchenkasse, Victor-Goerttler-Straße 4, 07745 Jena, Germany.

Summary

Mycobacterium avium subspecies *paratuberculosis* (MAP) is a pathogenic bacterium responsible for the lethal Johne's disease in cattle. So far, several genome-wide association studies (GWAS) have been carried out to identify chromosomal regions highly associated with Johne's disease. The aim of this study was to investigate the genetic variability within a pool of seven genes (*LAMB1, DLD, WNT2, PRDM1, SOCS5, PTGER4* and *IL10*) indicated by former GWAS/RNA-Seq studies as putatively associated with MAP infections and to achieve a confirmation study of association with paratuberculosis susceptibility in a population of 324 German Holstein cattle (162 cases MAP positive and 162 controls MAP negative) using ELISA and fecal cultural tests. SNP validation and genotyping information are provided, quick methods for allelic discrimination were set up and transcription factor binding analyses were performed. The *rs43390642:G>T* SNP in the *WNT2* promoter region is associated with paratuberculosis susceptibility ($P = 0.013$), suggesting a protective role of the *T* allele ($P = 0.043$; odds ratio 0.50 [0.25–0.97]). The linkage disequilibrium with the *DLD rs134692583:A>T* might suggest a combined mechanism of action of these neighboring genes in resistance to MAP infection, which is also supported by a significant effect shown by the haplotype $DLD^T/WNT2^T$ ($P = 0.047$). *In silico* analysis predicted *rs43390642:G>T* and *rs134692583:A>T* as essential parts of binding sites for the transcription factors GR, C/EBPβ and GATA-1, hence suggesting a potential influence on *WNT2* and *DLD* gene expression. This study confirmed the region on BTA 4 (UMD 3.1: 50639460–51397892) as involved in tolerance/resistance to Johne's disease. In addition, this study clarifies the involvement of the investigated genes in MAP infection and contributes to the understanding of genetic variability involved in Johne's disease susceptibility.

Keywords friesian breed, Johne's disease, MAP infection, *WNT2* gene

Introduction

Mycobacterium avium subspecies *paratuberculosis* (MAP) is a pathogenic intracellular bacterium known to be the causative agent for Johne's disease (paratuberculosis) in cattle and other ruminants. The disease was described for the first time by Johne & Frottingam (1895) as an atypical case of bovine tuberculosis. Nowadays, the pathogenesis of the disease is not completely clarified, but the chronic inflammation of the intestine with granulomatous lesions of the ileum is considered the main alterations of the intestinal mucosa caused by MAP. Animals with Johne's disease show progressive loss of weight, chronic diarrhea and reduction in productive performances (for a review, see Purdie *et al.* 2011).

Many reasons prevent the early diagnosis of the disease: the prolonged incubation time (up to 10 years) of MAP, the difficulties to clearly identify the infection status of the animals and the high incidence of subclinical infections and asymptomatic cases. Furthermore, the lack of quick and cost-efficient commercially available diagnostic tests with

Address for correspondence

A. Pauciullo, Laboratory of Animal Cytogenetics and Gene Mapping, National Research Council (CNR) of Italy, ISPAAM, via Argine 1085, 80147 Naples, Italy.
E-mail: alfredo.pauciullo@cnr.it

Accepted for publication 14 November 2014

high sensitivity makes the application of prophylaxis programs, as well as the eradication of MAP from an infected herd, very difficult (Beyerbach *et al.* 2001).

It is known that MAP infections are present worldwide, and this condition has a very negative impact on the economy of the dairy industry. Recently, Küpper *et al.* (2013) estimated a reduction in milk yield per day of life in MAP+ cows, and in the USA, the losses were estimated at approximately 200 million USD per year in terms of reduced milk production, limited reproduction efficiency and increased management costs (Ott *et al.* 1999).

Susceptibility to paratuberculosis has been showed to have a genetic component, and the heritability in cattle was estimated to range from 0.041 to 0.228 (Koets *et al.* 2000; Mortensen *et al.* 2004; Gonda *et al.* 2006; Hinger *et al.* 2008; Attalla *et al.* 2010; van Hulzen *et al.* 2012; Küpper *et al.* 2012). Despite many efforts and genetic approaches (including microsatellite genotyping, whole genome scanning for quantitative trait loci, SNP arrays) that have been attempted (for a review, see Purdie *et al.* 2011), the identification of the genetic component contributing to the phenotypic variance of MAP susceptibility is still not clear. However, independent genome-wide association studies (GWAS) conducted in the past few years proved that only a restricted number of chromosomal regions carry strongly significant SNPs involved in Johne's disease (Table 1). Although potential candidate genes have been indicated in each of these studies, so far no further investigation including characterization of the genes in Holstein cattle has been performed.

The understanding of the loci associated with susceptibility/resistance to disease is fundamental to incorporating them into breeding schemes and eradicating the disease. Therefore, the aim of this study was to investigate the genetic variability within a pool of genes putatively associated with MAP infection and accomplish a confirmation study of association with paratuberculosis susceptibility in a population of German Holstein cattle classified as MAP positive (MAP+) and MAP negative (MAP−) using both ELISA and fecal cultural results.

Table 1 Positional candidate genes recently indicated as putatively involved in Johne's disease after genome-wide association studies.

SNP ID	BTA	RefSeq genes (1 Mb)	References
ss86341066:A>C	3	*EDN2*	Settles *et al.* (2009)
rs43070062:C>G	9	–	Settles *et al.* (2009)
ss61491930:G>A	7	*IL4, IL5, IL13, IRF1*	Pant *et al.* (2010)
ss86328445:T>C	11	*SOCS5*	Pant *et al.* (2010)
ARS-BFGL-NGS-8531:A>G	9	*PRDM1*	Minozzi *et al.* (2010)
UA-IFASA-8974:A>C	20	*PTGER4*	Kirkpatrick *et al.* (2010)
rs41748405:A>C	15	GNA12	Zanella *et al.* 2011
ss66537488:C>T	4	*DLD, LAMB1, WNT2*	van Hulzen *et al.* (2012)

Materials and methods

Sample collection, nucleic acid isolation and diagnostic tests

A total of 324 German Holstein cows from 15 different farms located in Thuringia (Germany) were considered in this study. To eliminate potential stratification factors, a case–control study was designed. Therefore, 162 fecal culture positive (AVID 2007) animals (age >24 months) were chosen as cases, and 162 fecal culture negative animals from the same farm, from the same sire and at the same age were used as controls. Blood samples were collected to isolate the DNA according to the procedure described by Montgomery & Sise (1990). DNA concentration and $OD_{260/280}$ ratio of the samples were measured with the Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific Inc.).

Antibody levels were determined using a commercially available ELISA test (CATTLETYPE® MAP Ab; Labor Diagnostik) according to the manufacturer's information.

SNP selection and PCR conditions

Six genes falling within chromosomal regions indicated as highly associated with MAP resistance/susceptibility were investigated for genetic variability. The NCBI Reference SNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/) was used as a tool for the identification of SNPs to be genotyped in the MAP-tested German Holstein population. The complete list of the SNPs, the corresponding genes, their positions and the genotyping information are reported in Table 2, using as reference the UMD 3.1 sequence assembly. RNA-Seq data analysis from a current project indicated *interleukin-10* (*IL10*) as significantly upregulated during early MAP infection (data not shown); therefore, the *IL10* gene also was included in the analysis for SNP discovery and genotyping (Table 2). To reach this goal, the five exons and the intronic regions of the *IL10* gene underwent resequencing for test samples.

For genotyping, DNA fragments ranging from 103 to 402 bp were amplified by PCR and digested by specific restriction endonucleases. A typical PCR mix (25 μl) comprised: 50 ng of genomic DNA, 1X PCR Buffer (Promega), 2.5 mM of MgCl2, 5 pmol of each primer, dNTPs each at 200 μM and 1 U of Taq DNA polymerase (Promega). PCR was performed under the following conditions: 95 °C for 4 min, 35 cycles at 95 °C for 30 s, 56 °C for 30 s (with the exception of the protocol run for the *rs136770416* SNP, the annealing of which was set at 50 °C) and 72 °C for 30 s, and a final extension at 72 °C for 5 min. Product specificity was confirmed by ethidium bromide-stained 2% agarose gel electrophoresis.

Table 2 Reference SNPs chosen for a confirmation study of association with MAP susceptibility. Chromosome, gene location, alleles and amino acid AA changes are reported. Positions of the SNPs related to the gene (5′ flanking region/exon/intron) and related to BTA UMD 3.1 are also indicated. Primer sequences, amplicon size (bp) and restriction endonuclease information were provided for PCR amplification and RFLP genotyping.

Gene	BTA	RefSNP	Alleles	AA change	Gene Pos./UMD3.1	Primer	Size (bp)	Genotyping
LAMB1	4	*rs43388824*	A>G	–	Exon 8 49318278	5′-GGGAAGTAAACTTTACATAAG-3′ 5′-GCACGTACTTACCATTC-3′	345	BsmAI
LAMB1	4	*rs211391398*	C>T[1]	Ser>Asn	Exon 11 49312998	5′-TTGGTTAAAGATAAAATGAAGC-3′ 5′-TTTTGTGAAATTTGGAGGG-3′	402	DraI
DLD	4	*rs134692583*	A>T[1]	–	5′ flank. reg. 50639460	5′-TTACGCTCTTTACGACAGT-3′ 5′-TTCTGCCAAGGATTTCAC-3′	168	SphI
WNT2	4	*rs43390642*	G>T[1]	–	5′ flank. reg. 51397892	5′-GGGTGGATGAAATGATGGCAA-3′ 5′-TCTACCCCGAGCGCTTG-3′	167	HaeIII
PRDM1	9	*rs136669229*	A>C[1]	Phe>Val	Exon 2 44375813	5′-CAGAGTCATATCCGCGTC-3′ 5′-CGGGACAATGGGGATTAA-3′	103	NlaIV
PRDM1	9	*rs136770416*	A>C	Val>Gly	Exon 5 44359367	5′-TTGATGAGATTCACCGCCT-3′ 5′-CTGAAGGACAAGGCCTG-3′	140	DraIII
SOCS5	11	*rs134378401*	A>C	Tyr>Asp	Exon 2 29050621	5′-AGAGCGACTTCCTACAGT-3′ 5′-AGTTCACTGGATACGGATAAAA-3′	129	RsaI
PTGER4	20	*rs41944920*	C>G	Leu>Val	Exon 3 33764044	5′-TGATAAGTTCAGCGTTTCAC-3′ 5′-AGCCATAGAGAAGATCAAAT-3′	277	HinfI
IL10	16	Present work	A>G[1]	–	Intron 4	5′-CATGACCTTCCCAGCAG-3′ 5′-AATAAATATATGTGGGAGCTGAG-3′	353	Alw26I

[1]Polymorphic alleles confirmed in German Holstein.

Test samples and genotyping

A confirmation test of genetic variability was preliminarily run on 28 German Holstein (14 MAP+ and 14 MAP−) belonging to six different farms to validate the chosen SNP. For the negative SNPs, the confirmation test was extended to 48 additional German Holstein (24 MAP+ and 24 MAP−).

Restriction fragment length polymorphism protocols were set for each SNP. Digestion of 17 μl of each PCR amplification was accomplished with 10 U of endonuclease (Thermo Fisher Scientific Inc.) for 16 h at 37 °C. The digestion products were analyzed directly by electrophoresis in 2% agarose gel in 1X TBE buffer and stained with ethidium bromide. The samples with missing genotypes were amplified and digested at least twice before the exclusion to the statistical analysis.

In silico analysis of transcription factors binding sites

The *DLD rs134692583:A>T*; *WNT2 rs43390642:G>T*; and *IL10 AC_000173*:g.3625A>G SNPs were analyzed for potential transcription factor binding sites applying the online tool TFSEARCH, which is based on the TRANSFAC database (Heinemeyer *et al.* 1998). Transcription factors with predicted binding scores of ≥75 for each allele were included in the analysis (max. score = 100). For each SNP, major and minor alleles, including the flanking 15 nucleotides upstream (5′) and downstream (3′), were analyzed.

Statistical analysis

The allele frequency and Hardy–Weinberg equilibrium (χ^2 test) were calculated by means of POPGENE software ver 1.31 (University of Alberta, Canada). SAS system software (SAS 9.1; SAS Institute Inc.) was used to estimate differences between the allele frequencies of the different polymorphisms in the investigated genes between cases and controls. Analyses were carried out by Fisher's exact test using three different phenotypes: (i) fecal positive/negative, (ii) ELISA positive/negative and (iii) fecal or ELISA positive/negative. According to the same scheme, odds ratios (ORs) were calculated for the minor allele at each SNP by SAS system software considering 95% confidence intervals (CI). A test of H_0 for OR = 1 was calculated by Fisher's exact test. All tests were two-tailed considering P-values < 0.05 as significant.

The linkage disequilibrium parameters (D' and r^2) and the haplotype frequencies for the SNPs located on chromosome 4 (*rs211391398:C>T*; *rs134692583:A>T* and *rs43390642:G>T*) were estimated with HAPLOVIEW software according to the model proposed by Wang *et al.* (2002).

Results

Genetic variability occurred within a group of six genes (*LAMB1, DLD, WNT2, PRDM1, SOCS5* and *PTGER4*) indicated as putative candidates for the resistance/susceptibility to MAP infection. Eight SNPs falling within these genes were chosen from the NCBI Reference SNP database for genotyping of infected and uninfected animals. *In silico* analysis indicated that these SNPs were responsible for amino acid changes (Table 2) or had a potential effect on the transcriptional regulation (either enhancement or repression) because the affected nucleotide changes putative binding sites for gene transcription factors (Table 3).

Four of eight SNPs (*LAMB1 rs43388824:A>G; PRDM1 rs136770416:A>C; SOCS5 rs134378401:A>C; PTGER4 rs41944920:C>G*) were monomorphic in our test samples (38 MAP+ and 38 MAP−). Therefore, the total population of German Holstein (162 cases and 162 controls) was genotyped for the remaining four polymorphic sites (*LAMB1 rs211391398:C>T; DLD rs134692583:A>T; WNT2 rs43390642:G>T; PRDM1 rs136669229:A>C*) plus an additional SNP found in *IL10.* In fact, the comparison of the *IL10* sequences showed two new SNPs (g.1309C>T at intron 2 and g.3625A>G at intron 4; numbering is relative to the EMBL acc. no. AC_000173 used as reference) not previously reported in the NCBI database. The second polymorphic site is located only 14 bp upstream of exon 5, and it might potentially affect GATA binding sites; therefore, it was chosen for the genotyping of the complete population by the PCR-RFLP method.

The restriction patterns for each of the digestion protocols are reported in Fig. 1. Briefly, the transition *rs211391398: C>T* in *LAMB1* was restricted in two fragments for the *TT* samples (238 bp and 164 bp) and unrestricted (402 bp) for the *CC* genotype, whereas the heterozygote pattern was 402, 238 and 164 bp. A similar pattern for the transversion *rs134692583:A>T* was present for the *DLD* gene. The *AA* genotype was characterized by two fragments of 103 and 65 bp, the *TT* genotype showed an undigested band of 168 bp and the heterozygous sample resulted in three fragments (168, 103 and 65 bp). The transversion *rs43390642:G>T* at the *WNT2* locus showed the following restriction pattern: *TT,* 167 bp; *GT,* 167 bp, 95 bp and 72 bp; and *GG,* 95 bp and 72 bp.

The restriction pattern for the transversion *rs136669229: A>C* in the *PRDM1* gene was characterized by two fragments for the *AA* genotype (103 and 39 bp), three fragments for the *CC* genotype (75, 39 and 28 bp) and four fragments for the AC genotype (103, 75, 39 and 28 bp).

A similar pattern for the transversion *AC_000173: g.3625A>G* was detected in the *IL10* gene. The *AA* genotype showed two fragments of 282 and 71 bp. The 282-bp band was further digested into two fragments of 220 and 62 bp for the *GG* genotype. The heterozygous showed four fragments of 282, 220, 71 and 62 bp. The difference in size of the last two bands was only 9 bp, and they were not discriminated on the gel. Therefore, these bands appear as unique for the *AG* and *GG* genotypes (Fig. 1, lines 22 and 23).

The genotype distribution and the allelic frequencies are reported in Table 4. Missing genotypes varied between 0.3%

Table 3 Analysis of transcription factor binding sites in the DNA sequences surrounding SNPs: (a) *DLD rs134692583:A>T*; (b) *WNT2 rs43390642: G>T*; and (c) *IL10 AC_000173:g.3625A>G* (present work) by TFSEARCH software.

Transcription factor	Consensus sequence[1]	Position relative to SNP (5′>3′)	DNA strand	Binding allele score *T*	Binding allele score *A*
(a) *rs134692583:A>T* (8 bp upstream *DLD*) 5′-AGGCCGCGCTCGTGC***[A/T]***TGCGCAGGGCGGGGA-3′					
USF	NCACG**T**GN	−5 to +2	3′>5′	78.4	78.4
c-Myc	NANCACG**T**GNNW	−7 to +4	3′>5′	76.8	–
C/EBPb	NKN**T**TGCNYAAYNN	−3 to +10	5′>3′	76.0	–
Arnt	NDDNNCACG**T**GNNNNN	−8 to +5	3′>5′	75.9	–
N-Myc	NNCCACGTGN**N**N	−10 to +1	5′>3′	76.0	77.0
GATA-1	SNNGA**T**NNNN	−5 to +4	5′>3′	75.5	–

Transcription factor	Consensus sequence[1]	Position relative to SNP (5′>3′)	DNA strand	Binding allele score *G*	Binding allele score *T*
(b) *rs43390642:G>T* (164 bp upstream *WNT2*) 5′-AAACACCTCCGTGTG***[G/T]***CCTCGAGCACCCGCG-3′					
AML-1a	TGC**G**GT	−4 to +1	5′>3′	83.7	–
ZID	N**G**GCTCYATCAYC	−1 to +11	5′>3′	78.9	–
GR	NNNNNNCNNTNTG**T**NCTNN	−13 to +5	5′>3′	–	78.2
c-Ets-1	NC**M**GGAWGYN	−9 to +2	3′>5′	75.1	75.1

Transcription factor	Consensus sequence[1]	Position relative to SNP (5′>3′)	DNA strand	Binding allele score *A*	Binding allele score *G*
(c) *AC_000173:g.3625A>G* (14 bp upstream the exon 5 of *IL10*) 5′-CACTGAACACGTCTT***[A/G]***TCTCCCCACACAGCT-3′					
GATA-X	NGAT**A**AGNMNN	−4 to +6	3′>5′	93.1	–
GATA-1	NNC**W**GATARNNNN	−3 to +9	3′>5′	90.3	–
SREBP	K**A**TCACCCCAC	−1 to +9	5′>3′	90.1	85.9
GATA-3	NNGAT**A**RNG	−5 to +3	3′>5′	89.1	–

[1]Nucleotides in the genomic sequences according to the consensus sequences are underlined, and the polymorphic nucleotide is marked in bold. Exonic regions are double-underlined. Binding score threshold for each allele was set to ≥75.0. Nucleotide codes: K = *G* or *T*; M = *A* or *C*; S = *C* or *G*; W = *A* or *T*; Y = *C* or *T*; R = *A* or *G*; D = *A*, *G* or *T*; N = *A, G, C* or *T*.

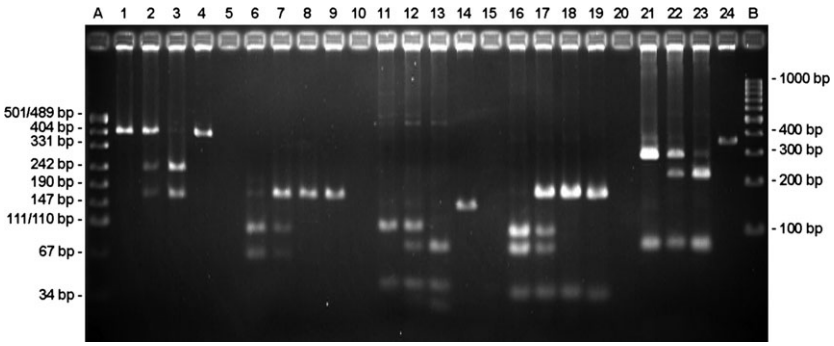

Figure 1 Genotyping of *LAMB1, DLD, WNT2, PRDM1* and *IL10* by PCR-RFLP in a German Holstein cattle population. Lines 1–4: locus *LAMB1 rs211391398:C>T*; genotypes *CC, CT* and *TT* reported in lines 1, 2 and 3 respectively. Lines 6–9: locus *DLD rs134692583:A>T;* genotypes *AA, AT* and *TT* reported in lines 6, 7 and 8 respectively. Lines 11–14: locus *PRDM1 rs136669229:A>C*; genotypes *AA, AC* and *CC* reported in lines 11, 12 and 13 respectively. Lines 16–19: locus *WNT2 rs43390642:G>T*; genotypes *GG, GT* and *TT* reported in lines 16, 17 and 18 respectively. Lines 21–24: locus *IL10 g.3625A>G*; genotypes *AA, AG* and *GG* reported in lines 21, 22 and 23 respectively. Lines 4, 9, 14, 19 and 24 show undigested PCR products each belonging to the relative locus. Lines 5, 10, 15 and 20 are empty. Line A shows the pUC 19 DNA/MspI (HpaII) marker, 23 (Fermentas); Line B shows GeneRuler™ 100 bp DNA Ladder (Thermo Scientific).

(*IL10*) and 7.7% (*DLD*). Chi-square values indicated no evidence of departure from Hardy–Weinberg equilibrium in the total population for all the analyzed SNPs. The distribution of genotypes according to the phenotypic test (fecal and ELISA) is provided in Table 5. Differences in the frequencies of *rs43390642:G>T*, located in the promoter region of *WNT2*, were observed in infected animals compared to healthy controls (Table 5). The *rs43390642:G>T* SNP was associated with paratuberculosis susceptibility using both fecal ($P = 0.035$) and ELISA ($P = 0.049$) diagnostic tests. Association ($P = 0.013$) also was found when a more restrictive phenotype was used (fecal or ELISA positive/negative). In the MAP+ group, the frequency of the rarer *T* allele of the *rs43390642:G>T* polymorphism was 0.04 (fecal/ELISA), whereas in the controls, it was 0.08. The OR value ($P = 0.043$, OR 0.50 [0.25–0.97]) suggested a protective effect of the minor allele in MAP infection. In contrast, no associations were observed for the other investigated SNPs with the exception of *rs134692583:A>T* in the *DLD* gene, whose OR showed a significant effect ($P = 0.046$, OR 0.58 [0.34–0.99]) for the minor allele (*T*) when fecal and ELISA tests were matched together.

A *D′* value of 0.968 and an r^2 value of 0.561 suggested a linkage disequilibrium between the two SNPs (*DLD rs134692583:A>T* and *WNT2 rs43390642:G>T*), which formed one haplotypic block covering a region of about 750 kb. The analysis allowed us to detect three haplotypes: *AG, TT* and *TG* whose frequencies were 0.894, 0.057 and 0.047 respectively. The χ^2 test run for each haplotype showed a significant association ($P = 0.047$) of the *TT* haplotype with resistance to MAP infection.

Discussion

The most recent approach to understanding which chromosomal region is involved in MAP susceptibility is represented by GWAS. So far, several studies have been carried out (Settles *et al.* 2009; Kirkpatrick *et al.* 2010; Minozzi *et al.* 2010, 2012; Pant *et al.* 2010; Zanella *et al.* 2011; van Hulzen *et al.* 2012), and a limited number of positional candidate genes have been indicated (Table 1). Despite this information, little investigation within the genes more or less biologically involved with MAP infection and falling in such regions has been carried out. Our study focused on the variability of seven genes (*LAMB1, DLD, WNT2, PRDM1, SOCS5, PTGER4* and *IL10*) indicated from independent GWAS and RNA-Seq studies as potentially associated with MAP infection.

Laminin, beta 1 (*LAMB1*), *dihydrolipoamide dehydrogenase* (*DLD*) and *wingless-type MMTV integration site family member 2* (*WNT2*) genes are located on chromosome 4 spread over a DNA region of about 2 Mbp and in the surrounding (~1 Mbp) of the *ss66537488:*C>T SNP, indicated by van Hulzen *et al.* (2012) as associated with paratuberculosis susceptibility by GWAS.

Laminin is a biologically active protein and an important structural component of the basement membrane, mediating the interactions between cells and matrix. It was described as one of the molecular determinants involved in the adherence of MAP to epithelial cells (Pethe *et al.* 2001). In human, three independent GWAS evidenced the controversial role of *LAMB1* in relation to Crohn's disease (Barrett *et al.* 2009; McGovern *et al.* 2010; van Sommeren *et al.* 2011). Although van Hulzen *et al.* (2012) indicated this gene as putatively involved in Johne's disease, our association study using fecal, ELISA and the combination of both diagnostic tests did not confirm such a role (Table 5). Moreover, the OR values (range 0.87–1.33) gave no evidence of a protective/susceptible allele effect for *LAMB1*. This result is in line with the findings of van Sommeren *et al.* (2011). Furthermore, our data showed that, although located within a DNA fragment of 2 Mbp, *LAMB1*

Table 4 Genotyping data, allele frequency and Hardy–Weinberg equilibrium ($P \leq 0.05$) of the SNPs chosen as molecular markers for a confirmation study of association with MAP susceptibility in German Holstein population. All chi-square tests have one degree of freedom.

			Observed genotypes				Allele frequency	
Locus	SNP		*CC*	*CT*	*TT*	Total	*C*	*T*
LAMB1	*rs211391398:C>T*	Obs. Exp. $\chi^2 = 0.013$	231 230.78	66 66.44	5 4.78	302	0.87	0.13
			AA	*AT*	*TT*	Total	*A*	*T*
DLD	*rs134692583:A>T*	Obs. Exp. $\chi^2 = 0.908$	240 238.42	54 57.15	5 3.42	299	0.89	0.11
			GG	*GT*	*TT*	Total	*G*	*T*
WNT2	*rs43390642:G>T*	Obs. Exp. $\chi^2 = 0.013$	287 287.11	36 35.77	1 1.11	324	0.94	0.06
			AA	*AC*	*CC*	Total	*A*	*C*
PRDM1	*rs136669229:A>C*	Obs. Exp. $\chi^2 = 0.009$	5 4.81	68 68.37	243 242.81	316	0.12	0.88
			AA	*AG*	*GG*	Total	*A*	*G*
IL10	*AC_000173:g.3625A>G*	Obs. Exp. $\chi^2 = 0.259$	2 1.43	39 40.14	282 281.43	323	0.07	0.93

rs211391398:C>T does not belong to the same haplotype block as *DLD rs134692583:A>T* (*LAMB1, DLD:* $r^2 = 0.002$) and *WNT2 rs43390642:*G>T (*LAMB1, WNT2:* $r^2 = 0.009$), thus suggesting the probable existence of a recombination hotspot among the first and the last two genes.

The *DLD* gene encodes for an enzyme called dihydrolipoamide dehydrogenase. Several recent GWAS have reported *DLD* as a putative candidate gene for resistance to mycobacterial infection both in human (Barrett *et al.* 2009; McGovern 2010) and in bovine (van Hulzen *et al.* 2012). This enzyme is involved in many biological pathways related to energy metabolism, including the degradation of essential amino acids. The different aptitude to use these nutrients as a result of genetic differences in *DLD* was proposed as a possible solution for the difference in development of the clinical stage of Johne's disease among infected animals (van Hulzen *et al.* 2012). Our investigation was focused on the *DLD rs134692583:A>T* SNP, located only 8 bp upstream of exon 1. The association study with MAP infection data showed controversial results. The Fisher's exact test did not confirm the association with all the available phenotypes (Table 5). However, the OR ($P = 0.046$; OR 0.58 [(0.34–0.99]) in restrictive phenotype conditions gave evidence of a protective effect of the *T* allele. The analysis of the putative transcription factor showed that this SNP is responsible for the alteration of several binding sites (Table 3). In particular, C/EBPβ is an important regulator of cytokine expression (Cloutier *et al.* 2009), whereas GATA-1 is a potent suppressor of T_h1-associated genes such as interferon-γ and chemokine receptor-3 (Sundrud *et al.* 2005). Both sites are putatively active only the in presence of the *T* allele with a binding allele score of 76.0 and 75.5% respectively, hence suggesting higher transcriptional activation of the *DLD* gene.

The *wingless-type MMTV integration site family member 2* (*WNT2*) gene, shown to be significantly associated with MAP infection ($P = 0.013$–0.049) using all available phenotypes, belongs to a family of structurally related genes (WNT) that encode glycoproteins and extracellular signaling molecules. Abnormal WNT signaling is linked to a range of diseases, especially cancer. The best-understood WNT-signaling pathway goes through the activation of the nuclear functions of β-catenin, which leads to changes in gene expression that influence cell proliferation and survival (Moon *et al.* 2004). Abnormal proliferation of fibroblasts in animals affected with Johne's disease is a key feature in granuloma formation, consisting of chronic inflammatory cells, which include macrophages, giant cells, lymphocytes,

Table 5 Frequencies observed for fecal and ELISA tests for each investigated SNP. Differences between the allele frequencies of the different polymorphisms were calculated by Fisher's exact test using three different phenotypes: (1) fecal positive/negative, (2) ELISA positive/negative and (3) fecal or ELISA positive/negative. Odds ratios (ORs) with 95% confidence intervals (CI) were calculated for each minor allele frequency (MAF) according to the same scheme.

	LAMB1 rs211391398:C>T					Alleles			
Test	*CC*	*CT*	*TT*	Total	*P* Fisher	*C*	*T*	OR [95% CI] MAF	*P* ORs
(1)									
+	116	29	3	148	0.642	261	35	0.87 [0.53–1.41]	0.581
–	115	37	2	154		267	41		
(2)									
+	71	25	2	98	0.510	167	29	1.33 [0.81–2.19]	0.256
–	160	41	3	204		361	47		
(3)									
+	124	37	3	164	0.926	285	43	1.11 [0.68–1.80]	0.670
–	107	29	2	138		243	33		
	DLD rs134692583:A>T								
	AA	*AT*	*TT*	Total		*A*	*T*		
(1)									
+	126	22	2	150	0.235	274	26	0.64 [0.38–1.09]	0.108
–	114	32	3	149		260	38		
(2)									
+	85	15	0	100	0.173	185	15	0.57 [0.31–1.05]	0.075
–	155	39	5	199		349	49		
(3)									
+	140	24	2	166	0.135	304	28	0.58 [0.34–0.99]	0.046*
–	100	30	3	133		230	36		
	WNT2 rs43390642:G>T								
	GG	*GT*	*TT*	Total		*G*	*T*		
(1)									
+	149	12	1	162	0.050*	310	14	0.56 [0.28–1.11]	0.098
–	138	24	0	162		300	24		
(2)									
+	103	7	1	111	0.050*	213	7	0.42 [0.18–0.97]	0.042*
–	184	29	0	213		397	31		
(3)									
+	166	13	1	180	0.019*	345	15	0.50 [0.25–0.97]	0.043*
–	121	23	0	144		265	23		
	PRDM1 rs136669229:A>C								
	AA	*AC*	*CC*	Total		*A*	*C*		
(1)									
+	3	30	124	157	0.583	36	278	0.85 [0.52–1.36]	0.505
–	2	38	119	159		42	276		
(2)									
+	2	26	77	105	0.543	30	180	1.29 [0.79–2.11]	0.295
–	3	42	166	211		48	374		
(3)									
+	3	36	134	173	0.926	42	304	0.95 [0.59–1.54]	0.864
–	2	32	109	143		36	250		

Table 5 (Continued)

	IL10 AC_000173:g.3625A>G								
	AA	*AG*	*GG*	Total		*A*	*G*		
(1)									
+	0	22	140	162	0.495	22	302	1.15 [0.61–2.17]	0.657
–	2	17	142	161		19	301		
(2)									
+	0	15	94	109	0.589	15	203	1.13 [0.58–2.19]	0.702
–	2	24	188	214		26	400		
(3)									
+	0	24	155	179	0.493	24	334	1.13 [0.59–2.16]	0.694
–	2	15	127	144		17	269		

*$P \leq 0.05$.

plasma cells and fibroblasts that deposit collagen and extracellular matrix proteins to create a dense fibrous region similar to a capsule (Ackermann 2013). Fibroblast proliferation and collagen synthesis are crucial in the repair of injured tissue associated with inflammatory lesions (for a review, see Flavel *et al.* 2008). WNT/β-catenin signaling is activated in this process (Cheon *et al.* 2004). In fact, the injection of soluble WNT inhibitor into adult mice inhibits intestinal cell proliferation and suggests the possibility of using WNT activators to regenerate gut epithelium as adjuvant therapy in inflammatory bowel disease (Kuhnert *et al.* 2004). Therefore, this suggests a possible role of the *WNT2* gene in Johne's disease. Our investigation was focused on the *rs43390642:G>T* SNP that is located in the promoter region of *WNT2* gene. The significant association was confirmed by Fisher's exact test and further validated through the OR values (Table 5), which suggested a protective role of the *T* allele in MAP infection. The presence of thymine is responsible for the creation of a putative consensus sequence for a glucocorticoid receptor (GR) element. Glucocorticoids have a vast array of functions within the body, including the potent suppression of immune response and inflammation. Their beneficial use as drugs in the treatment of human chronic inflammatory bowel diseases was first recognized nearly 60 years ago (Truelove & Witts 1954). The major anti-inflammatory effects of glucocorticoids appear to be due largely to the interaction between the activated GR and anti-inflammatory genes such as *annexin A1*, *interleukin-10* and the inhibitor of NF-kB (Hayashi *et al.* 2004). Such mechanism of interaction was well elucidated by Li *et al.* (2003) using a mouse mammary tumor virus (MMTV) promoter, which also characterizes the *WNT2* gene. Briefly, glucocorticoids inhibit expression of adhesion molecules and trafficking of inflammatory cells to target tissues (Hayashi *et al.* 2004), which might explain the protective role of the *T* allele for the *rs43390642*:G>T SNP in the *WNT2* promoter.

This SNP resulted in linkage disequilibrium with the polymorphic site found in the *DLD* promoter. The *TT* haplotype was associated ($P = 0.047$) with resistance to MAP infection, but it was less significant than *WNT2* alone, hence suggesting that the existence of cooperative action of these genes mediated by the activation of their specific transcription factors has to be deeply investigated. Functional studies are necessary to clarify the influence of C/EBPβ, GATA-1 and GR transcription factors on *DLD* and *WNT2* gene expression; however, the potential positive interaction between neighbor genes might partially explain the 'infinitesimal' effect expected from these loci in the control of such a complex trait.

PR domain containing 1 with ZNF domain (*PRDM1*) encodes a zinc finger-containing transcriptional repressor of beta-interferon. Our investigation on two A>C transversions responsible for an amino acid change within the *PRDM1* gene did not confirm any association with MAP susceptibility. However, recently a rare allele identified in *PRDM1* was associated with Crohn's disease in humans (Ellinghaus *et al.* 2013). Functional studies reported by the same authors showed that this rare risk allele led to increased peripheral blood lymphocytes expression of the L-selectin adhesion molecule, increased $CD4^+$ and $CD8^+$ T-cell proliferation, and increased IFN-γ secretion and upregulation of activation markers (Ellinghaus *et al.* 2013). Each of these factors may contribute to the pathogenic role of *PRDM1* in this disease.

The suppressor of cytokine signaling (SOCS) family of proteins is functionally closely related to interleukins. Cytokine signaling is in fact negatively regulated by SOCS proteins. We genotyped a SNP in the *SOCS5* gene, which is reported in the NCBI dbSNP as responsible for amino acid change *rs134378401*:$A^{Tyr}>C^{Asp}$, but this polymorphic site was monomorphic in the German Holstein population. However, *SOCS5* remains functionally interesting because it is preferentially expressed in Th1 cells, which promote cell-mediated effector responses to eliminate intracellular pathogens (Seki *et al.* 2002).

Prostaglandin E receptor 4 (subtype EP4) (*PTGER4*) has been identified in human as a Crohn's disease candidate gene by Libioulle *et al.* (2007) and Barrett *et al.* (2009). The same indication was reported after GWAS in Holstein cattle (Kirkpatrick *et al.* 2010). Our confirmation study was carried out on a SNP located at exon 3 and reported as responsible

for an amino acid change *rs41944920*:C^{Leu}>G^{Val}. However, this SNP was not polymorphic in the German cattle population. Also, the role of *PTGER4* related to MAP infection has to be taken into great consideration. Recently in human, Glas *et al.* (2012) demonstrated the strong Crohn's disease association of two SNPs (*rs4495224:A>C* and *rs7720838:G>T*) as part of binding sites for NF-kB and XBP1, suggesting that these transcription factors may modulate *PTGER4* gene expression. Therefore, future investigations are strongly required also in bovine to clarify the role of *PTGER4* as a candidate gene for paratuberculosis susceptibility.

IL10 is classified as a class-2 cytokine. Recent studies of gene expression indicated a greater upregulation of *IL10* in cow monocytes after 2 h of infection with MAP (Weiss *et al.* 2005). The intronic *AC_000173:g.3625A>G* SNP within *IL10* is putatively responsible for the alteration of GATA factor binding sites (Table 3). This family of transcription factors plays a key role for cytokine gene expression by Th2 cells (Zheng & Flavell 1997). In addition, intron GATA binding sites in the *IL4* gene were proved to be essential for acting both as a transcriptional enhancer and as a demethylation factor (Hural *et al.* 2000). Despite these biological functions, the result of the association of the g.3625A>G SNP with both fecal and ELISA tests did not confirm a role of *IL10* in susceptibility to MAP infection (Table 4).

Although the frequencies of minor alleles for all the investigated SNPs are low and genotyping larger sample sizes might better elucidate their role in susceptibility to paratuberculosis, many studies report that low-frequency and rare variants are involved in the etiology of complex traits (Bodmer & Bonilla 2008; Gibson 2011). Furthermore, even for diseases with a strong genetic component, the identified common variants usually explain only a small portion of the total genetic heritability. For instance, in a study of Crohn's disease, more than 30 loci were identified, but they explain <10% of the overall heritability (Barrett *et al.* 2008).

In conclusion, this study confirms the region (UMD 3.1:50639460–51397892) on chromosome 4 as a susceptibility locus for Johne's disease. *WNT2* is significantly associated with MAP infection on the basis of both diagnostic systems: fecal culture and ELISA tests. The *T* allele at the *rs43390642:G>T* locus showed a potential protective role against paratuberculosis as part of binding sites for a glucocorticoid receptor element, suggesting that this transcription factor may modulate *WNT2* expression.

The linkage disequilibrium with the *DLD rs134692583:A>T* SNP and the potential protective effect of the *T* allele also at this locus might suggest a combined mechanism of action of these neighboring genes in the resistance to MAP infection. However, functional assays are necessary to clarify whether these two SNPs modulate binding of transcription factors and thereby regulate their target gene expression and MAP infection susceptibility. The protective role of the *TT* haplotype could be a useful resource to support and verify the situation of the available 50K SNP chip data in Holsteins and eventually be implemented in genome-wide breeding programs.

No association with MAP infection was found for the other investigated genes, thus not confirming the results of previous reports (Table 1). Further investigation is also required to clarify the possible biological role of these genes in the pathogenesis of Johne's disease.

This study contributes to the understanding of genetic variability involved in Johne's disease susceptibility, and it clarifies the involvement of the investigated loci in MAP infection. The identification of loci associated with MAP susceptibility is the first step in setting up marker-assisted selection programs to make cattle populations more resistant, reduce the transmission of MAP to other animals in the herd, improve the health status by breeding and increase the productivity of the livestock industry.

Acknowledgements

This work was financially supported in part by the 7th EC Framework Programme Project 'MacroSys' FP7-KBBE-2007-1-1-2. We are grateful to the herd book breeders from Thuringia (Germany) for participating in the voluntary program and for providing the data. We are grateful to National Research Council of Italy (CNR) – Short Term Mobility Programme – for allowing the research activity of Dr. Alfredo Pauciullo in cooperation with the Justus-Liebig University of Gießen (Germany). The authors would also like to thank Margarete Falke and Anja Scheuermann for their excellent technical assistance.

References

Ackermann M.R. (2013) Inflammation and healing. In: *Pathologic Basis of Veterinary Disease* (Ed. by J.F. Zachary & M.D. McGavin), pp. 124–5. Elsevier, St. Louis, MO.

Attalla S.A., Seykora A.J., Cole J.B. & Heins B.J. (2010) Genetic parameters of milk ELISA scores. *Journal of Dairy Science* **93**, 1729–35.

AVID (2007) *Kultureller Nachweis von* Mycobacterium Avium *Subsp.* Paratuberculosis *in Kot- Oder Organproben*. Deutsche Veterinärmedizinische Gesellschaft (DVG) e. V., Gießen.

Barrett J.C., Hansoul S., Nicolae D.L. *et al.* (2008) Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature Genetics* **40**, 955–62.

Barrett J.C., Lee J.C., Lees C.W. *et al.* (2009) Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nature Genetics* **41**, 1330–4.

Beyerbach M., Gerlach G.F. & Kreienbrock L. (2001) Modelling of prevalence development in a paratuberculosis control program in a dairy herd. *Deutsche Tierärztliche Wochenschrift* **108**, 363–70.

Bodmer W. & Bonilla C. (2008) Common and rare variants in multifactorial susceptibility to common diseases. *Nature Genetics* **40**, 695–701.

Cheon S.S., Nadesan P., Poon R. & Alman B.A. (2004) Growth factors regulate beta-catenin-mediated TCF-dependent transcriptional activation in fibroblasts during the proliferative phase of wound healing. *Experimental Cell Research* **293**, 267–74.

Cloutier A., Guindi C., Larivée P., Dubois C.M., Amrani A. & McDonald P.P. (2009) Inflammatory cytokine production by human neutrophils involves C/EBP transcription factors. *The Journal of Immunology* **182**, 563–71.

Ellinghaus D., Zhang H., Zeissig S. *et al.* (2013) Association between variants of *PRDM1* and *NDP52* and Crohn's disease, based on exome sequencing and functional studies. *Gastroenterology* **145**, 339–47.

Flavel S.J., Hou T.Z., Lax S., Filer A.D., Salmon M. & Buckley C.D. (2008) Fibroblasts as novel therapeutic targets in chronic inflammation. *British Journal of Pharmacology* **153**, S241–6.

Gibson G. (2011) Rare and common variants: twenty arguments. *Nature Reviews Genetics* **13**, 135–45.

Glas J., Seiderer J., Czamara D. *et al.* (2012) *PTGER4* expression-modulating polymorphisms in the 5p13.1 region predispose to Crohn's disease and affect NF-κB and XBP1 binding sites. *PLoS ONE* **7**, e52873.

Gonda M.G., Chang Y.M., Shook G.E., Collins M.T. & Kirkpatrick B.W. (2006) Genetic variation of *Mycobacterium avium* ssp. *paratuberculosis* infection in US Holsteins. *Journal of Dairy Science* **89**, 1804–12.

Hayashi R., Wada H., Ito K. & Adcock I.M. (2004) Effects of glucocorticoids on gene transcription. *European Journal of Pharmacology* **500**, 51–62.

Heinemeyer T., Wingender E., Reuter I. *et al.* (1998) Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acids Research* **26**, 362–7.

Hinger M., Brandt H. & Erhardt G. (2008) Heritability estimates for antibody response to *Mycobacterium avium* subspecies *paratuberculosis* in German Holstein cattle. *Journal of Dairy Science* **91**, 3237–44.

van Hulzen K.J.E., Nielen M., Koets A.P., de Jong G., van Arendonk J.A.M. & Heuven H.C.M. (2012) Effect of herd prevalence on heritability estimates of antibody response to *Mycobacterium* avium *subspecies paratuberculosis*. *Journal of Dairy Science* **94**, 992–7.

Hural J.A., Kwan M., Henkel G., Hock M.B. & Brown M.A. (2000) An intron transcriptional enhancer element regulates *IL-4* gene locus accessibility in mast cells. *The Journal of Immunology* **165**, 3239–49.

Johne H.A. & Frottingam L. (1895) An atypical case of tuberculosis in cattle. *Deutsche Zeitschrift für Tiermedizin und vergleichende Pathologie* **21**, 438–54.

Kirkpatrick B.W., Shi X., Shook G.E. & Collins M.T. (2010) Whole-genome association analysis of susceptibility to paratuberculosis in Holstein cattle. *Animal Genetics* **42**, 149–60.

Koets A.P., Adugna G., Janss L.L., van Weering H.J., Kalis C.H., Wentink G.H., Rutten V.P. & Schukken Y.H. (2000) Genetic variation of susceptibility to *Mycobacterium avium* subsp. *paratuberculosis* infection in dairy cattle. *Journal of Dairy Science* **83**, 2702–8.

Kuhnert F., Davis C.R., Wang H.T., Chu P., Lee M., Yuan J., Nusse R., Kuo C.J. *et al.* (2004) Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1. *Proceedings of the National Academy of Science of the United States of America* **101**, 266–71.

Küpper J., Brandt H., Donat K. & Erhardt G. (2012) Heritability estimates for *Mycobacterium avium* subspecies *paratuberculosis* status of German Holstein cows tested by fecal culture. *Journal of Dairy Science* **95**, 2734–9.

Küpper J., Brandt H., Donat K. & Erhardt G. (2013) Associations between paratuberculosis status and milk production traits in Holstein cattle under consideration of interaction effects between test result and farm and lactation number. *Archives Animal Breeding* **56**, 94.

Li X., Wong J., Tsai S.Y., Tsai M. & O'Malley B.W. (2003) Progesterone and glucocorticoid receptors recruit distinct coactivator complexes and promote distinct patterns of local chromatin modification. *Molecular and Cellular Biology* **23**, 3763–73.

Libioulle C., Louis E., Hansoul S. *et al.* (2007) Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of *PTGER4*. *PLoS Genetics* **3**, e580538.

McGovern D.P.B., Gardet A., Törkvist L. *et al.* (2010) Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nature Genetics* **42**, 332–7.

Minozzi G., Buggiotti L., Stella A., Strozzi F., Luini M. & Williams J.L. (2010) Genetic loci involved in antibody response to *Mycobacterium avium* ssp. *paratuberculosis* in cattle. *PLoS ONE* **5**, e11117.

Minozzi G., Williams J.L., Stella A., Strozzi F., Luini M., Settles M.L., Taylor J.F., Whitlock R., Zanella R. & Neibergs H.L. (2012) Meta-analysis of two genome-wide association studies of bovine paratuberculosis. *PLoS ONE* **7**, e32578.

Montgomery G.W. & Sise J.A. (1990) Extraction of DNA from sheep white blood cells. *New Zealand Journal of Agricultural Research* **33**, 437–41.

Moon R.T., Kohn A.D., De Ferrari G.V. & Kaykas A. (2004) WNT and β-catenin signaling: diseases and therapies. *Nature Reviews Genetics* **5**, 691–701.

Mortensen H., Nielsen S.S. & Berg P. (2004) Genetic variation and heritability of the antibody response to *Mycobacterium avium* subspecies *paratuberculosis* in Danish Holsteins cows. *Journal of Dairy Science* **87**, 2108–13.

Ott S.L., Wells S.J. & Wagner B.A. (1999) Herd-level economic losses associated with Johne's disease on US dairy operations. *Preventive Veterinary Medicine* **40**, 179–92.

Pant S.D., Schenkel F.S., Verschoor C.P., You Q., Kelton D.F., Moore S.S. & Karrow N.A. (2010) A principal component regression based genome wide analysis approach reveals the presence of a novel QTL on BTA7 for MAP resistance in Holstein cattle. *Genomics* **95**, 176–82.

Pethe K., Puech V., Daffe M., Josenhans C., Drobecq H., Locht C. & Menozzi F.D. (2001) Mycobacterium smegmatis laminin-binding glycoprotein shares epitopes with Mycobacterium tuberculosis heparin-binding haemagglutinin. *Molecular Microbiology* **39**, 89–99.

Purdie A.C., Plain K.M., Begg D.J., de Silva K. & Whittington R.J. (2011) Candidate gene and genome-wide association studies of *Mycobacterium avium* subsp. *paratuberculosis* infection in cattle and sheep: a review. *Comparative Immunology, Microbiology and Infectious Disease* **34**, 197–208.

Seki Y., Hayashi K., Matsumoto A., Seki N., Tsukada J., Ransom J., Naka T., Kishimoto T., Yoshimura A. & Kubo M. (2002)

Expression of the suppressor of *cytokine signaling*-5 (*SOCS*5) negatively regulates *IL-4*-dependent STAT6 activation and Th2 differentiation. *Proceedings of the National Academy of Science of the United States of America* **99**, 13003–8.

Settles M., Zanella R., McKay S.D., Schnabel R.D., Taylor J.F., Whitlock R., Schukken Y., Van Kessel J.S., Smith J.M. & Neibergs H. (2009) A whole genome association analysis identifies loci associated with *Mycobacterium avium* subsp. *paratuberculosis* infection status in US Holstein cattle. *Animal Genetics* **40**, 655–62.

van Sommeren S., Visschedijk M.C., Festen E.A., de Jong D.J., Ponsioen C.Y., Wijmenga C. & Weersma R.K. (2011) HNF4α and CDH1 are associated with ulcerative colitis in a Dutch cohort. *Inflammatory Bowel Diseases* **17**, 1714–8.

Sundrud M.S., VanCompernolle S.E., Eger K.A., Bruno T.C., Subramaniam A., Mummidi S., Ahuja S.K. & Unutmaz D. (2005) Transcription factor GATA-1 potently represses the expression of the HIV-1 coreceptor *CCR*5 in human T cells and dendritic cells. *Blood* **106**, 3440–8.

Truelove S.C. & Witts L.J. (1954) Cortisone in ulcerative colitis; preliminary report on a therapeutic trial. *British Medical Journal* **2**, 375–8.

Wang N., Akey J.M., Zhang K., Chakraborty R. & Jin L. (2002) Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *The American Journal of Human Genetics* **71**, 1227–34.

Weiss D.J., Evanson O.A. & Souza C.D. (2005) Expression of *interleukin-10* and *suppressor of cytokine signaling-3* associated with susceptibility of cattle to infection with *Mycobacterium avium* subsp *paratuberculosis*. *American Journal of Veterinary Research* **66**, 1114–20.

Zanella R., Settles M.L., McKay S.D., Schnabel R., Taylor J., Whitlock R.H., Schukken Y., Van Kessel J.S., Smith J.M. & Neibergs H.L. (2011) Identification of loci associated with tolerance to Johne's disease in Holstein cattle. *Animal Genetics* **42**, 28–38.

Zheng W. & Flavell R.A. (1997) The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell* **89**, 587–96.