A genome scan for QTL affecting resistance to Haemonchus contortus in sheep.
— **Source link** ⧉

Guillaume Sallé, Philippe Jacquiet, L. Gruner, Jacques Cortet ...+16 more authors

Institutions: Institut national de la recherche agronomique

Published on: 01 Dec 2012 - Journal of Animal Science (American Society of Animal Science)

Topics: Haemonchus contortus, Population and Linkage disequilibrium

Related papers:

- The distribution of SNP marker effects for faecal worm egg count in sheep, and the feasibility of using these markers to predict genetic merit for resistance to worm infections

- Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs

- Quantitative trait loci associated with parasitic infection in Scottish blackface sheep

- A joint analysis to identify loci underlying variation in nematode resistance in three European sheep populations.

- Discovery of quantitative trait loci for resistance to parasitic nematode infection in sheep: I. Analysis of outcross pedigrees

Share this paper: ⓕ 🐦 in ✉

View more about this paper here: https://typeset.io/papers/a-genome-scan-for-qtl-affecting-resistance-to-haemonchus-4nkk2kt0mq

A genome scan for QTL affecting resistance to Haemonchus contortus in sheep

G. Sallé, P. Jacquiet, L. Gruner, J. Cortet, C. Sauvé, F. Prévot, C. Grisez, J. P. Bergeaud, L. Schibler, A. Tircazes, D. François, C. Pery, F. Bouvier, J. C. Thouly, J. C. Brunel, A. Legarra, J. M. Elsen, J. Bouix, R. Rupp and C. R. Moreno

The online version of this article, along with updated information and services, is located on the World Wide Web at:

http://www.journalofanimalscience.org/content/90/13/4690

Society of Animal Science

www.asas.org

A genome scan for QTL affecting resistance to *Haemonchus contortus* in sheep[1]

G. Sallé,*†[2] P. Jacquiet,† L. Gruner,‡ J. Cortet,‡ C. Sauvé,‡, F. Prévot,†
C. Grisez,† J. P. Bergeaud,† L. Schibler,§ A. Tircazes,* D. François,* C. Pery,# F. Bouvier,#
J. C. Thouly,# J. C. Brunel,# A. Legarra,* J. M. Elsen,* J. Bouix,* R. Rupp,* and C. R. Moreno*

*INRA, UR631, Station d'Amélioration Génétique des Animaux, BP 27, F-31326, Castanet-Tolosan, France; †INRA, UMR1225, Interactions Hôtes—Agents Pathogènes, BP 87614, F-31076 Toulouse, France; ‡INRA, UR1282, Infectiologie Animale et Santé Publique, F-37880 Nouzilly, France; §INRA, UMR1313, Laboratoire de Génétique Biochimique et de Cytogénétique, F-78252 Jouy-en-Josas, France; and #INRA, UE332, Domaine de la Sapinière, F-18390, Osmoy, France

ABSTRACT: Gastrointestinal nematodes are one of the main health issues in sheep breeding. To identify loci affecting the resistance to *Haemonchus contortus*, a genome scan was carried out using 1,275 Romane × Martinik Black Belly backcross lambs. The entire population was challenged with *Haemonchus contortus* in 2 consecutive experimental infections, and fecal egg counts (FEC) and packed cell volumes were measured. A subgroup of 332 lambs with extreme FEC was necropsied to determine the total worm burden, length of female worms, sex ratio in the worm population, abomasal pH, and serum and mucosal G immunoglobulins (IgG) responses. Pepsinogen concentration was measured in another subset of 229 lambs. For QTL detection, 160 microsatellite markers were used as well as the Illumina OvineSNP50 BeadChip that provided 42,469 SNP markers after quality control. Linkage, association, and joint linkage and association analyses were performed with the QTLMAP software. Linkage disequilibrium (LD) was estimated within each pure breed, and association analyses were carried out either considering or not the breed origin of the haplotypes. Four QTL regions on sheep chromosomes (OAR)5, 12, 13, and 21 were identified as key players among many other QTL with small to moderate effects. A QTL on OAR21 affecting pepsinogen concentration exactly matched the pepsinogen (*PGA5*) locus. A 10-Mbp region affecting FEC after the 1st and 2nd infections was found on OAR12. The SNP markers outperformed microsatellites in the linkage analysis. Taking advantage of the LD helped to refine the locations of the QTL mapped on OAR5 and 13.

Key words: *Haemonchus contortus*, linkage disequilibrium, quantitative trait loci, resistance, sheep

INTRODUCTION

Gastrointestinal nematodes (**GIN**) are one of the main health issues in grazing ruminants (Coop et al., 1985; Man-

donnet et al., 2003; Mandonnet et al., 2005; Davies et al., 2006). Among GIN, *Haemonchus contortus* settles in the sheep abomasum where it sucks blood from its host, thus leading to anemia and even to death in the absence of treatment. This issue has worsened over recent years because GIN have developed resistance against anthelmintics, the usual means of control. Resistance is constantly increasing in terms of prevalence, geographical repartition, and severity (Sargison et al., 2007; Traversa et al., 2007; Howell et al., 2008; Hoglund et al., 2009; Cezar et al., 2010).

Selection for animals that could resist nematode infection may provide a feasible long-term control strategy (Bishop and Morris, 2007). Furthermore, subtropical and tropical sheep breeds, such as Gulf Coast Native, St Croix, Santa Ines, or Barbados Blackbelly are considered to be relatively resistant to *H. contortus* (Aumont et al., 2003; Amarante et al., 2004), hence contrasting with other breeds from temperate areas, like

4690

the French Romane breed, formerly known as INRA401 (Terefe et al., 2007).

Taking advantage of the resistance contrast between the subtropical Martinik Black Belly breed (**MBB**) and the RMN in a backcross population consisting of 1,275 experimentally infected lambs, the aim of this study was to perform in-depth QTL detection for the resistance to *H. contortus*. Extensive phenotyping was achieved through measurement of fecal egg count (**FEC**), packed cell volumes (**PCV**), total worm burden, worm female length, abomasal pH, blood serum and mucosal specific IgG, and serum pepsinogen. Genotyping strategy firstly was a microsatellite selective genotyping of a subset of 332 lambs with 160 microsatellite markers before switching to SNP DNA markers from the Illumina Ovine SNP50 BeadChip genotyped for 1,000 lambs.

MATERIALS AND METHODS

All animals were kept indoors, handled with care, and managed as a commercial flock following the INRA ethics policy. At the end of the experimental infection, animals were slaughtered at the INRA-Nouzilly abattoir following the EU rules.

Study Populations

Backcross population. A backcross design between resistant MBB and susceptible RMN was set up as follows (Table 1): 5 F_1 (MBB × RMN) rams were backcrossed by intrauterine AI with 600 purebred RMN ewes at the INRA experimental farm, La Sapinière (UE0332, Osmoy, France). These inseminations resulted in the birth of 1,046 male and female lambs forming a first population born in 2003 and hereafter denoted **BC1**. According to the preliminary QTL mapping study performed in the BC1 population, 4 out of the 5 original F_1 rams were identified as segregating for most of the QTL. In 2006, these 4 sires were mated to 134 RMN ewes that gave birth to 229 backcross lambs (denoted **BC2**).

Additional Purebred Populations. To estimate the linkage disequilibrium (**LD**) extent in the 2 pure breeds that composed the backcross population, additional animals were used. A 90-individual MBB population structured in nuclear families (i.e., sire, dam, and progeny) was selected to be representative of the entire population. The RMN population was composed of an 8-family granddaughter design of 1,050 individuals.

Measured Traits

Infection Procedure. Lambs were challenged artificially twice at 3 and 5 mo of age with 10,000 infective *H. contortus* larvae. At the end of the first infection (i.e.,

Table 1. Summary of experimental design for each family available[1,2]

Item	Total no. lambs Infected	No. lambs with extreme FEC_a (used for microsatellite selective genotyping)	Available SNP data after quality control
Sire No.			
14952	268 (48)	72	249
14971	191 (0)	64	0
14976	301 (63)	66	284
14988	269 (75)	64	247
16754	245 (43)	66	220
Total	1,275 (229)	332	1,000
Measured traits[2]	FEC, PCV	WB, L, IgGm, IgGs	
	Peps (for BC2 only)	SexR, pH	

[1]For each founder sire, the respective number of progeny, among which the number of extreme resistant/susceptible BC1 (first backcross between resistant Martinik Black Belly and susceptible Romane breeds) progeny used for fine phenotyping, are reported. The respective numbers of BC2 (second Martinik Black Belly × Romane backcross, with rams selected for specific QTL) progeny are indicated in parentheses. Every trait that was measured for each subset of lambs is indicated and number of individuals that were genotyped for the SNP chip (4 families).

[2]FEC, fecal egg count; FEC_a, animal solution of a mixed model equation with the infection rank added to other fixed effects and animal fitted as a random variable; PCV, packed-cell volume; Peps, pepsinogen concentration; WB, worm burden; L, female worm length; IgGm, IgG concentration in abomasal mucus; IgGs, IgG concentration in serum; SexR, sex ratio in adult worm population; pH, abomasal pH.

41 d postchallenge; **dpc**), lambs were drenched (7.5 mg of levamisole 5%/kg BW; Virbac S. A., Carros, France) and remained uninfected for an 8-d washout period before being reinfected with the same number of larvae.

FEC and PCV Traits Measured on BC1 and BC2 Populations. Fecal samples were taken twice at 25 and 35 dpc for the 1,275 backcross lambs of the BC1 and BC2 populations. Fecal egg counts were determined following the McMaster technique modified by Raynaud (1970), and FEC values were averaged for each infection (hereafter denoted **FEC12** for the first infection and **FEC34** for the second infection). Blood samples were taken in EDTA-coated tubes to determine PCV at d 0 (denoted **PCV0**). A second PCV measure was done at 41 dpc in both infections in BC1 population and at 35 dpc in BC2 population (denoted **PCV1** and **PCV2** for 1st and 2nd infections, respectively).

Fine Phenotyping of 332 BC1 Lambs with Extreme FEC. After the second infection, the 15% most resistant and 15% most susceptible BC1 lambs of each ram family were selected according to their overall FEC measurements. These selected subsets of 332 lambs were slaughtered on d 42 ± 2 after the second infection to gather additional phenotypes. Measurement techniques have been fully described elsewhere (Lacroux et al., 2006). Total worm burden (**WB**) was determined and the sex ratio (**SexR**) of adult worms (females:males)

was computed. Total lengths (**Len**) of 20 randomly chosen female worms per lamb were measured using a video camera. The abomasal pH (**pH**) was evaluated by using pH paper in direct contact with the mucous layer. Serum and mucus IgG (**IgGs** and **IgGm**, respectively) concentrations were determined by indirect ELISA (Lacroux et al., 2006) performed on serum and abomasal fundic mucosa collected from each animal just after death.

Pepsinogen Concentration Measured in the 229 BC2 Lambs. Serum pepsinogen concentrations were determined using a micromethod for routine determination adapted from Dorny and Vercruysse (1998). This trait is a direct indicator of the mucosal damage caused by *H. contortus* infection. For each infection, the pepsinogen concentration was measured twice at d 0 and at 15 dpc. These measurements are hereafter denoted **Peps1** and **Peps2** for the 1st infection and **Peps3** and **Peps4** for the 2nd infection.

Molecular Quality Checks and Map Construction

DNA extraction and genotyping were performed at LABOGENA (www.labogena.fr; verified November 9, 2012).

Microsatellite Genotypes. A panel of 160 microsatellite markers distributed across the whole genome was used. The 2 groups of BC1 lambs with extreme FEC were genotyped, resulting in a so-called selective genotyping procedure (Lander and Botstein, 1989). In addition to these 332 lambs, their 5 sires were also genotyped (Table 1). Raw microsatellite data were analyzed with Genetic Profiler v. 1.5 software (Amersham Biosciences, Uppsala, Sweden). The relative positions of markers were established with the Cri-map software (Green et al., 1990) using the international sheep map (http://www.ncbi.nlm.nih.gov/mapview; verified November 9, 2012) as well as both the human and bovine genome sequences for a subset of unpublished markers developed by the INRA institute.

SNP Genotypes. The 4 backcross families segregating for most of the QTL found with microsatellites were genotyped with the IlluminaOvineSNP50 Beadchip (Illumina, Inc., San Diego, CA). In total, 1,044 backcross lambs were genotyped, as well as their 4 respective F_1 sires and the 4 MBB founders (Table 1). In addition, the 2 additional purebred MBB and RMN populations were genotyped, hence providing full 50K SNP chip data for 90 MBB and 1,050 RMN individuals.

Individuals with a call rate below 98% were discarded, and a 99.9% technical reliability was established by duplicated genotyping in 50 animals. In addition, Mendelian inconsistencies (i.e., no allele shared in common between a progeny and its sire for a given SNP) were checked. Intrinsic SNP quality criteria were also considered. Useless SNP that had been eliminated in the frame of the sheep HAPMAP project (i.e., SNP that Illumina annotated as abnormal; SNP with minor allele frequencies equal to 0; SNP that displayed discordant genotypes between experiments; or SNP showing Mendelian inconsistencies within the International Mapping Flock) were discarded (J. W. Kijas, personal communication). Within our population, SNP with a call rate < 97% were removed. A minor allele frequency < 1% was applied. A test for Hardy-Weinberg disequilibrium was also considered in purebred populations to eliminate SNP with abnormal behavior ($P < 10^{-6}$). The SNP discarded in purebred populations were also eliminated from the backcross genotype data. Furthermore, SNP for which more heterozygotes than expected (i.e., 50% for a heterozygous sire) were counted, or SNP with too great a recombination rate were not included in the QTL analysis. Sex chromosomes were not considered for analysis. Finally, 1,000 backcross individuals, 939 RMN individuals, 90 MBB sheep, and 42,469 autosomous markers were retained for subsequent analysis.

The SNP positions were obtained from the sheep Genome Browser v. 2.0 on http://www.livestockgenomics.csiro.au/cgi-bin/gbrowse/oarv2.0/ (verified November 9, 2012; Archibald et al., 2010). For mapping purposes, 1 Mbp was considered as equivalent to 1 cM.

Statistical Handling of Phenotypes

Transformations Applied to Phenotypes. Basic statistics and correlations of phenotypes were computed and data were tested for normality using the Shapiro-Wilk test (SAS Inst. Inc., Cary, NC). A quasinormal distribution was observed for PCV, Len, and IgGs traits. A fourth root transformation was applied to FEC traits (**FEC12t** and **FEC34t** for 1st and 2nd infections, respectively) as well as to pH and IgGm (denoted **pHt** and **IgGmt**). Square root transformation corrected departures from normality for worm burden and sex ratio (**WBt** and **SexRt**, respectively). Finally, the difference between pepsinogen concentrations under a naive state and during infection was considered, and a fourth root was applied (denoted **Peps12t** and **Peps34t** for the 1st and 2nd infections, respectively).

Correction for Fixed Effects. Transformed phenotypes were subsequently corrected for the usually encountered environmental effects (i.e., sex, management group, litter size, and age at infection). The FEC and PCV were also considered as longitudinal traits over both infections by estimating the animal solution of a mixed model equation (SAS PROC MIXED), with the infection rank added to other fixed effects and animal fitted as a random variable. These traits were denoted **FEC_a** and **PCV_a**. To account for within-animal physiological variation in PCV, PCV values obtained after infection were corrected with PCV0 fitted as a covariable and denoted **PCV1c** and

PCV2c. Determination of the significant fixed effects was performed using SAS PROC GLM.

Phenotypic Correlations. Phenotypic correlations between transformed traits corrected for environmental effects were calculated using the Pearson correlation coefficient (SAS PROC CORR). In the BC1 flock, immunopathological traits were measured in 2 pools of animals selected on their FEC; this resulted in a bias in the observed correlations between other traits. Therefore, correlation coefficients were estimated by taking into account the increase in variance of the selected population (see formulae in APPENDIX 1).

Methods for Analysis of Marker Data

Comparison of LD and LD Phase. The r^2 LD measure (Hill and Robertson, 1968) was computed within pure breeds between SNP pairs less than 1 Mbp apart using SNP data and 2 specific pure breed populations. The correlation of r values across breeds expressed as a function of genomic distance was considered to investigate the persistence of LD phase between the 2 breeds, as reported by de Roos et al. (2008).

Linkage Analysis. Linkage analysis (**LA**) was performed with microsatellite markers in the selective genotyping design (denoted **LA-micro**) and with the 4-family SNP genotyped dataset (denoted **LA-SNP**). Within each family, the presence of a QTL was tested against the null hypothesis, which was absence of a QTL at every 0.1-cM interval (approximately 1 Mbp/1 cM), by likelihood computation using the QTLMAP software (Elsen et al., 1999). Chromosome-wise significance was determined for each chromosome-trait pair by testing with 10,000 permutations (Churchill and Doerge, 1994). To prevent any overestimations of thresholds due to the selective phenotyping occurring for some traits, permutations were restricted to animals that were both genotyped and phenotyped for the considered traits. These permutations were used to calculate the genome-wise suggestive threshold (1 false-positive result was expected for a whole genome scan) and genome-wise significant threshold (5% genome-wise significance threshold) as indicated in Lander and Kruglyak (1995). These genome-wise thresholds were obtained by applying the Bonferroni correction $P_{genome-wise} = 1 - (1 - P_{chromosome-wise})^n$, where n is the number of chromosomes (i.e., 26 in sheep; Knott et al., 1998). Confidence intervals were determined using the 2-LOD (i.e., logarithm of the odds) drop-off criterion (Ooijen, 1992) and assuming 1 LOD = 4.61 LRT (i.e., long terminal repeats; Lynch and Walsh, 1998). Estimated QTL effects were corrected for the bias due to selective genotyping (see formulae in APPENDIX 2) following the recommendations of Bovenhuis and Spelman (2000). The QTL effects were hence expressed in phenotypic SD (σ_p) corrected for selective phenotyping as described in APPENDIX 1. A *t*-test was performed to identify the 5% significant QTL effects within each sire family.

Association Analysis and Joint Linkage and Association Analysis. The LD-decay model proposed by Legarra and Fernando (2009) was implemented in the QTLMAP software and applied to our data. In this application of the LD-decay model, observed phenotypes were regressed on the conditional probability of having inherited a 4-SNP haplotype from each of its 2 parents, the substitution effect of the sire haplotypes being weighted by the transmission probabilities. In addition, a joint association and linkage analysis (**LDLA**) was performed to take advantage of both LD and pedigree information in the experimental population (Legarra and Fernando, 2009). In this LDLA model, within sire QTL effects are added to the sire haplotypes effects of the LD-decay model to account for a possible between sire variability of the QTL effect beyond that reflected by the haplotype.

For both genome-wide association studies (**GWAS**) and LDLA analyses, the minor haplotype frequency was set at 1% for analysis of the FEC and PCV traits and at 5% for other traits because fewer animals were measured, hence ensuring at least 10 observations for estimation of the maternal haplotype effect.

In both models described by Legarra and Fernando (2009), it is assumed that every founder originated from the same common base population. Because our working population was a mixture of 2 breeds, additional analysis was performed by clustering founders haplotypes according to their breed origin (i.e., 2 identical haplotypes were considered to be different if originating from different breeds). This is similar to Pérez-Enciso and Varona (2000). These analyses were annotated with a *b* index (**GWASb** and **LD-LAb** for association and joint analysis, respectively).

The chromosome-wise P values were estimated for each trait assuming the LRT statistics asymptotically followed a χ^2-distribution with *k* degrees of freedom, *k* being the number of QTL effects (Piepho, 2001). Hence, *k* was equal to the number of haplotypes minus 1 for GWAS and the number of haplotypes plus the number of families minus 1 for LDLA. Genome-wise *P* values were derived by applying a Bonferroni correction as described above. Any association reaching the 1% genome-wise significant *P* value was declared significant. Associations reaching the 5% genome-wise significant threshold were only reported if a suggestive QTL had already been found for the same trait by LA-SNP analysis.

RESULTS

Phenotypes and Correlations

A summary of PCV and FEC statistics for challenged lambs is reported in Table 2. During the primary infection, large FEC values were measured and the mean PCV of infected lambs decreased by 11%. Upon reinfection, FEC were significantly reduced in comparison with the 1st infection (–76% on average) and mean PCV were less reduced (Table 2). After worm development in the abomasum, mean pepsinogen concentrations followed a significant 2.6- and 3-fold increase between d 0 and 15 dpc in 1st and 2nd infections, respectively.

At the end of the second infection period of the BC1 population, a pool of 332 BC1 lambs with extremely large (susceptible; **S group**) or small (resistant; **R group**) FEC values ($P < 10^{-4}$) were assessed for additional phenotypes (i.e., characteristics of worm populations, abomasal pH, and IgGs and IgGm concentrations; Table 3). The mean

total number of worms in the S group was 3.6-fold greater than in the R group. The R group also exhibited a greater number of immature stages. Moreover, female worms measured in the S lambs were significantly longer ($P < 10^{-4}$). SexR in adult worms were slightly different between the 2 groups with more females counted in the R group ($P = 0.02$). Resistant lambs maintained greater PCV values during both infections ($P < 10^{-4}$) but exhibited lesser serum IgG concentrations ($P < 2.10^{-4}$). No significant differences were observed between R and S groups as regards to abomasal pH and mucosal IgG values.

Phenotypic correlations between the different variables in the BC1 population are listed in Table 4. Mean FEC and PCV values were negatively correlated in both infections whatever the population, whereas PCV2 was positively correlated to PCV1. Total worm burden, length of female worms, and serum IgG concentrations were all positively correlated to FEC and negatively correlated to PCV in both infections. In addition, it was observed that the more worms counted in the abomasum, the greater the proportion of males, the longer the females, and the greater the serum IgG concentration. Most of the other correlations were not significant. Regarding pepsinogen concentrations measured in BC2 lambs, 2 significant correlations of –0.21 and –0.24 were found between FEC34t and Peps4 and PCV1 and Peps3, respectively.

Table 2. Basic statistics of packed-cell volume (PCV) and fecal egg counts (FEC) for the whole backcross population (n = 1,275 lambs) and summary statistics of pepsinogen concentration for the BC2 flock (n = 229 lambs), before and after 2 successive infections with *Haemonchus contortus*[1]

Trait	n	Mean	±SD	Minimum	Maximum
FEC and PCV in the whole population					
Before challenge					
Packed cell volume, %	1,245	39	5	26	55
First challenge					
FEC mean of 25 and 35 dpc,[2] eggs/g of feces	1,224	11,034	10,964	0	87,950
Packed cell volume before drenching, %	1,229	33	8	10	55
Second challenge					
FEC mean of 25 and 35 dpc,[2] eggs/g of feces	1,190	2,620	4,322	0	42,667
Packed cell volume before drenching, %	1,191	35	7	6	55
Pepsinogen concentration in the BC2 population					
Before any challenge, U/L	200	0.32	0.19	0	1.51
35 d after the 1st challenge, U/L	200	0.83	0.37	0	2.43
Before the 2nd challenge, U/L	201	0.35	0.16	0	1.03
35 d after the 2nd challenge, U/L	201	1.02	0.56	0.19	3.43

[1]BC2 represents the second of Martinik Black Belly × Romane backcross, with rams selected for specific QTL.

[2]dpc, days postchallenge.

Table 3. Egg excretion and parasitological examination of necropsied animals (n = 332) from the extreme resistant and susceptible groups (R and S group, respectively)

Trait	R group (n = 169)		S group (n = 163)		
	Mean	±SD	Mean	±SD	P value
Total worm burden	1,141	1,277	4,117	2,088	10^{-4}
L4 larvae, %	8%	18%	4%	10%	10^{-2}
Juveniles, %	6%	13%	2%	4%	10^{-3}
Adult males, %	38%	16%	45%	9%	10^{-4}
Adult females, %	49%	18%	49%	7%	NS
Sex ratio in adult worms	1.39	0.83	1.19	0.81	2.10^{-2}
Adult female length, μm	17,080	2,158	18,548	1,337	10^{-4}
pH in abomasum[1]	3.5	0.8	3.3	0.9	NS
IgG in serum[2]	44	26	55	28	2.10^{-4}
IgG in mucus[2]	53	71	56	76	NS
PCV0, %	40	5	40	6	NS
PCV1, %	37	7	32	9	10^{-4}
PCV2, %	39	7	32	8	10^{-4}
FEC12,[3] eggs/g of feces	4,427	6,616	21,463	15,572	10^{-4}
FEC34,[3] eggs/g of feces	292	709	7,014	6,705	10^{-4}

[1]For pH, only 119 resistant and 115 susceptible lambs were measured.

[2]IgG concentrations are expressed as a percentage of the basal value.

[3]FEC12, fecal egg count mean between 25 and 35 d after the 1st challenge; FEC34, fecal egg count mean between 25 and 35 d after the 2nd challenge.

Table 4. Phenotypic correlation (Pearson coefficient) between FEC, PCV, and parasitological examination traits[1]

	FEC12	FEC34	PCV1	PCV2	FEC_a	PCV_a	WB	SexR	Len	pH	IgGm	IgGs
PCV0	–0.07*	0.04	0.00	0.01	–0.02	0.06	–0.04	0.04	0.11*	0.00	0.02	–0.03
FEC12		0.28**	–0.33**	–0.14**	0.80**	–0.31**	0.40**	–0.11*	0.26**	–0.11*	0.11*	0.23**
FEC34			–0.16**	–0.42**	0.81**	–0.41**	0.69**	–0.19**	0.33**	–0.12*	–0.05	0.13*
PCV1				0.13*	–0.30**	0.72**	–0.21**	0.02	–0.16*	0.11	–0.10*	–0.18*
PCV2					–0.36**	0.77**	–0.51**	0.15*	–0.14*	0.06	–0.02	–0.19*
FEC_a						–0.45**	0.60**	–0.16**	0.31**	–0.13*	0.02	0.18**
PCV_a							–0.45**	0.09	–0.18*	0.10	–0.07	–0.22**
WB								–0.30**	0.23**	–0.17*	0.03	0.18*
SexR									0.00	0.09	–0.09	–0.02
L										–0.04	–0.07	0.03
pH											0.19*	0.01
IgGm												0.29**

*Significantly different from zero, $P < 0.01$.

**Significantly different from zero, $P < 0.0001$.

[1]FEC, fecal egg count, 12 and 34 indicate FEC after 1st and 2nd challenge, respectively; PCV, packed-cell volume, 1 and 2 indicate PCV after 1st and 2nd challenge, respectively; FEC_a and PCV_a, within-animal physiological variation accounted for in the analysis; WB, worm burden; SexR, sex ratio in adult worm population; Len, female worm length; pH, abomasal pH; IgGm, IgG in abomasal mucus; IgGs, IgG in serum.

Linkage Analysis with Microsatellites or SNP Markers

Full results are provided for the linkage analysis with microsatellites (Supplementary Data 1; see on-line version of article to access supplemental material, at http://journalofanimalscience.org) or SNP markers (Supplementary Data 2), as well as a comparison of the 2 types of molecular information (Supplementary Data 5).

Five-Family Selective Genotyping Design with Microsatellites. Linkage analysis was performed on the animals genotyped with the microsatellite panel (i.e., extreme R and S groups from the BC1 flock). A total of 8 suggestive and 3 significant QTL regions for resistance to *H. contortus* were detected (Table 5, Figure 1A, 1B). The QTL mostly affected FEC traits, PCV, and total worm burden, but also specific variables such as IgGmt, SexRt, Len, and pH. Chromosomes 3, 12 (see Figure 1A, 1B), and 23 were tagged as key players as they each carried significant QTL for several correlated traits (FEC12t, FEC34t, FECt_a, or WBt). In addition, these QTL were found in the same families and showed overlapping confidence intervals. Other QTL affecting pH, SexRt, Len, and IgGmt were mapped in clearly distinct regions on chromosomes 2, 5, 12, 13, 18, and 19.

Linkage Analysis with SNP Data. Four out of the 5 original families were genotyped with SNP (Table 1). In total, 9 significant QTL were mapped on sheep chromosome (**OAR**)5, 7, 10, 12, 13, and 21. As in the selective genotyping design with microsatellites, most of the QTL were related to FEC traits (16 out of the 38 at least suggestive QTL). Mean QTL effects for FEC and PCV ranged between 0.13 and $0.3\sigma_p$, whereas QTL effects estimated on other parasitological phenotypes were greater (Table 5).

Almost 50% of the QTL were clustered on 3 chromosomes (i.e., OAR5, 7, and 12). Of particular interest was the 10-Mbp region on OAR12 (Figure 2A, 2B) that was significantly associated with each of the FEC traits and exhibited some of the greatest effects for this trait (0.22 σ_p for FECt_a and 0.19 σ_p for the 2 other FEC traits). On OAR5, a suggestive QTL region located at the end of the chromosome was repeatedly associated with 7 correlated traits, but the confidence intervals usually spanned more than half of the chromosome (Figure 3A, 3B). It is interesting to note that 4 QTL were mapped on OAR7. A significant QTL with a narrow confidence interval was obtained for FEC12t but other suggestive QTL on this chromosome were distributed in a rather erratic way, either at the right (FEC34t) or at the left end of the chromosome. Last but not least, a QTL region on OAR21 was significantly associated with Peps34t at 37.8 Mbp (i.e., 3 Mbp from another suggestive QTL related to FECt_a; Figure 4). Other chromosomes showing multiple suggestive QTL (i.e., OAR16, 17, 23, and 25) also exhibited at least 10 Mbp between the peaks and/or had large confidence intervals (Table 5). It is worth noticing that no QTL were detected on OAR3 (but detected with microsatellites), but 1 suggestive region was reported for FEC34t on OAR20.

Taking Advantage of LD in the Experimental Population

LD Estimation in Pure Breeds. The LD between SNP markers was rather small in both breeds. The decrease in LD was slightly less in MBB compared with their RMN counterparts (Supplementary Figure 1A). At the average marker spacing (i.e., 57 kb), r^2 value was 0.14 and 0.13 for MBB and RMN, respectively. As expected, phase persistence between the 2 breeds decreased rapidly with genomic distance: r coefficients decreased

Table 5. Summary of linkage analyses (LA) studies with microsatellites or SNP data

OAR	Trait[1]	Selective genotyping with microsatellites		SNP data		Avg QTL Effect[3]
		Significance	Position, cM [CI][2]	Significance	Position, Mb [CI]	
1	PCV_a	–	–	*	146.5 [21.1–236.6]	0.15
2	PCV1c	–	–	*	244.3 [128.4–248.5]	0.19
2	SexRt	*	52 [9–80]	–	–	–
3	FEC34t	*	158 [141–167]	–	–	0.68
3	WBt	**	155 [144–162]	–	–	0.86
4	Len	–	–	*	45.6 [32.6–72.5]	0.36
4	pHt	*	73 [69–86]	–	–	0.89
5	FEC12t	–	–	*	86.7 [35.1–94.2]	0.16
5	FEC_a	–	–	*	86.7 [35.2–94.1]	0.17
5	IgGst	–	–	*	87.3 [60.5–94.2]	0.54
5	WBt	*	82 [72–89]	*	80.6 [37.4–94.4]	0.42
5	PCV1c	–	–	*	68.3 [32.2–91.3]	0.15
5	PCV2c	–	–	*	87.4 [7.6–96.7]	0.14
5	PCV_a	–	–	**	71.8 [53.3–92.7]	0.15
7	FEC12t	–	–	**	46.4 [28.8–52.4]	0.17
7	FEC34t	–	–	*	97.3 [21–100]	0.15
7	FEC_a	–	–	*	46.3 [20.4–50.5]	0.17
7	Len	–	–	*	13.8 [4.4–21.7]	0.55
9	FEC_a	–	–	*	52.3 [24.8–94.9]	0.13
10	Peps34t	–	–	**	83.5 [80.3–83.5]	0.66
12	FEC12t	**	55 [17–72]	**	56 [45.8–60.2]	0.19
12	FEC34t	*	59 [53–65]	**	47 [35.6–58.3]	0.19
12	FEC_a	**	58 [52–68]	**	46.3 [36.5–57.4]	0.22
12	PCV2c	–	–	*	45.4 [29.9–57.7]	0.20
12	PCV_a	–	–	*	35.6 [19.7–56.8]	0.30
12	pHt	*	77 [72–82]	–	–	0.84
13	FEC34t	–	–	**	72.3 [70.1–77.8]	0.22
13	IgGmt	*	109 [98–117]	*	0.6 [0.2–81.6]	0.30
14	Peps12t	–	–	*	61.2 [54.3–62.2]	0.36
15	FEC12t	no marker	no marker	*	43.5 [1.6–52.3]	0.18
16	FEC34t	–	–	*	27.2 [10.8–41.1]	0.18
16	WBt	–	–	*	19.2 [11.2–29.1]	0.44
17	pHt	–	–	*	63.7 [47.5–64.7]	0.31
17	PCV1c	–	–	*	18.1 [12.6–66.7]	0.13
17	PCV_a	–	–	*	65.1 [2.8–72.4]	0.17
18	Len	*	43 [49–81]	–	–	0.53
19	Len	*	51 [61–66]	–	–	0.48
20	FEC34t	–	–	*	31.3 [19.8–41.5]	0.17
21	FEC_a	–	–	*	41 [1.6–46.8]	0.14
21	Peps34t	–	–	**	37.8 [31.8–46.1]	0.79
23	FEC34t	*	51 [40–80]	*	59.6 [0.8–62.7]	0.15
23	FEC_a	*	54 [43–74]	*	32.2 [0.3–62.7]	0.27
23	WBt	*	63 [52–70]	*	44.1 [15–59.8]	0.58
25	SexRt	–	–	*	39.5 [0.4–40.7]	0.56
25	PCV2c	–	–	**	41.4 [16.6–44]	0.17

*Suggestive threshold.

**5% genome-wise significant.

[1]For fecal egg count (FEC) traits, results for SNP data are based on the four families backcross population., FEC12 and FEC34 indicate FEC after 1st and 2nd challenge, respectively; PCV, packed-cell volume, 1 and 2 indicate PCV after 1st and 2nd challenge, respectively, and c indicates values corrected with PCV0 fitted as a covariable; WB, worm burden; SexR, sex ratio in adult worm population; Len, female worm length; pH, abomasal pH; IgGm, IgG in abomasal mucus; IgGs, IgG in serum; Peps, pepsinogen; t, fourth root transformation of the variable; _a, within-animal physiological variation accounted for.

[2]CI, confidence interval.

[3]From the SNP analysis if the QTL were found in both analyses; given in phenotypic SD.

from 90% at 10 Kb to 50% at the average SNP spacing, and a 50% correlation was observed at the average SNP spacing (Supplementary Figure 1B). The LD phase persistence between MBB and RMN breeds decreased rapidly with genomic distances, and even at relatively short distances (<10 Kb) some correlations between *r* exhibited opposite signs (Supplementary Figure 1C).

Association Analysis and Joint Linkage and Association Analysis. Both association and joint linkage and association analyses were performed twice, either considering or not the breed origin of the haplotypes. Significant QTL are reported in Table 6, and Manhattan plots of the LDLA analysis for FEC12t and FEC34t are presented in Figures 5 and 6. Extensive results of the LD-based analyses performed can be found in Supplementary Data 3 (GWAS and LDLA analyses) and 4 (GWASb and LDLAb analyses).

In total, 41 haplotype-trait associations were declared significant, 24 of which were detected by LDLA analysis only (Table 6). By contrast, GWAS detected only 1 significant QTL, whereas both GWASb and LDLAb performed similarly (n = 11 and 15 significant QTL, respectively).

The chromosomes already highlighted by LA analysis (i.e., OAR5, 7, 12, 13, and 21) showed consistent significant associations using 3 different analysis methods (Figures 2 to 6 and Table 6). In addition, at least 4 trait-haplotype associations were found on OAR5, 12, and 21 (Figures 2, 3, and 4). An association hot-spot was associated with FEC, PCV, and pepsinogen concentration at the end of OAR5, and OAR12 was the only chromosome to be associated with both FEC traits (Table 6). The QTL region found by LA for Peps34t on OAR21 was associated with both Peps12t and Peps34t and pinpointed a 1-Mbp interval between 36.7 and 37.7 Mbp (Tables 5 and 6). Three additional associations were found on OAR20 and OAR23, but the SNP haplotypes were at least 10 Mbp apart from each other. Other chromosomes (OAR1, 2, 4, 8, 14, 17, 18, 22, 25, and 26) also exhibited 1% genome-wise significant associations but with the LDLA method only (Table 6). Among these, no previous QTL had been mapped using LA on chromosomes 8, 22, and 26 hence suggesting they segregate in the RMN breed only or that they are spurious signals (Table 5).

DISCUSSION

Overview

The aim of our study was to provide an exhaustive characterization of the genetic basis of breed differences in the response to *H. contortus*. This was not only achieved by using dense molecular information but also by studying a broad range of phenotypes (i.e., FEC, PCV, worm burden, length of females, IgG, and pepsinogen concentration) in more than 200 animals. After a first QTL mapping exercise with a selective genotyping strategy using microsatellites, 50k SNP markers were genotyped for 4 backcross families (1,000 genotyped individuals after quality control) in an attempt to more precisely map the QTL locations. This work complements another study by Kemper et al. (2011) that also used ovine SNP data to investigate the genetic architecture of host resistance to *Trichostrongylus colubriformis* and *H. contortus* in a large mixed breed population (more than 2,000 animals), again using FEC data.

Infection Dynamics

Considering basic traits, a sharp decrease in egg counts was observed upon reinfection, whereas large FEC values were associated with low PCV. In addition, the dynamics and intensity of FEC were similar between the 2 BC1 (mean FEC12 = 11,697 and mean FEC34 = 2,706) and BC2 flocks (mean FEC12 = 8,665 and mean FEC34 = 2,347). However, no increase in FEC was reported between 28 and 35 dpc during the first exposure period to *H. contortus* in the BC1 flock. It may be possible that the onset of the immune response in BC2 lambs at the 1st infection was not prompt enough to regulate worm growth. The R subgroup exhibited a smaller total worm burden and more immature worms in comparison with the S group. In addition, the length of female worms, which is correlated to female worm fertility (Stear et al., 1996), was shorter in the R group. Hence, the resistance of backcross lambs relied both on a reduction in worm burden and female fertility. Interestingly, this is consistent with findings in Scottish Blackface, in which the within breed variation of the control of another abomasal trichostrongylid (*Teladorsagia circumcincta*) is due to both a reduction in female fertility and worm burden (Stear et al., 1996). By contrast, Good et al. (2006) reported that the differences observed between Texel and Suffolk were only due to a lessening of the worm population. Interestingly, the control of *H. contortus* in the backcross population, as well as the differences observed between RMN and MBB breeds (Terefe et al., 2007), appear to follow the same pattern as reported within Scottish Blackface. The IgG titers were greater in the S group and positively correlated with FEC, hence strongly contrasting with results obtained in RMN lambs experimentally infected with *H. contortus* (Lacroux et al., 2006). This discrepancy may lie in the smaller number of animals (n = 26) considered for correlation computation in the study performed by Lacroux et al. (2006).
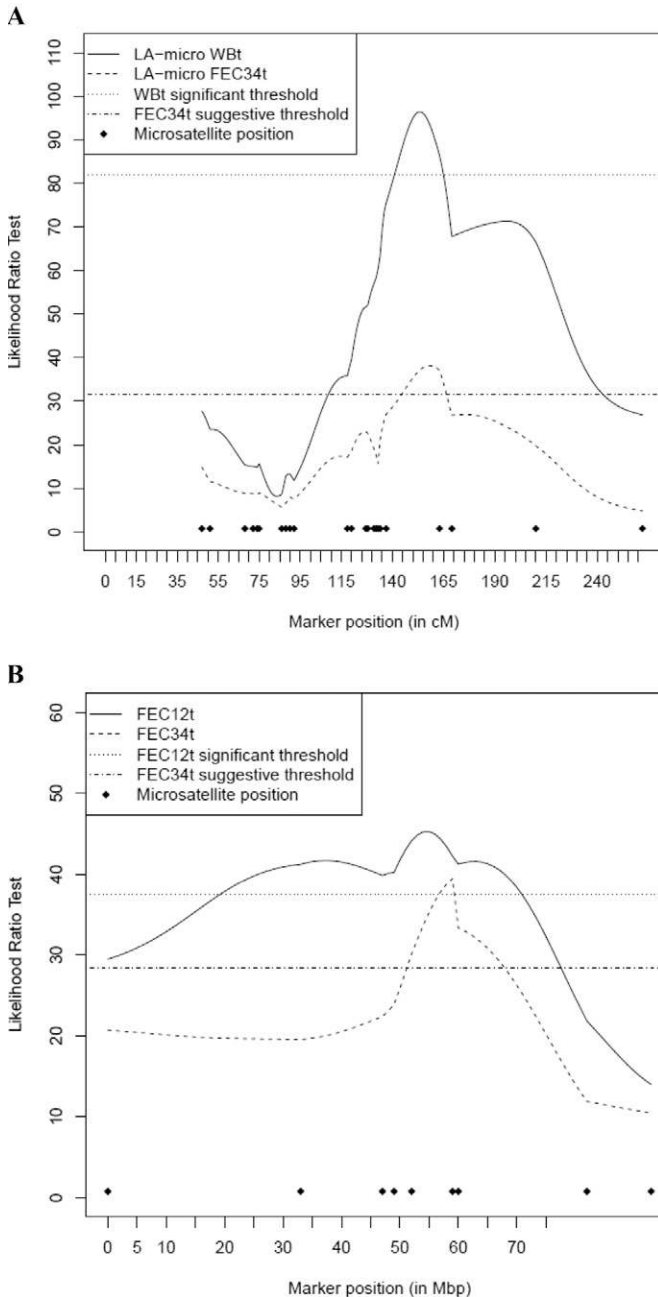
A

B

Figure 1. Likelihood profiles of linkage analyses on (A) OAR3 for worm burden (WBt) and fecal egg count at 2nd infection (FEC34t) and on (B) OAR12 for fecal egg count at first and second infection. Marker positions are indicated with a diamond symbol, and estimated thresholds are indicated with horizontal dotted and dashed-dotted lines. In Panel A, the likelihood ratio test statistics of a linkage analysis performed on fecal egg count at 2nd infection (LA-micro FEC34t, dashed line) and worm burden (LA-micro WBt, solid line) measured on the 332 animals from the selective genotyping design is plotted against microsatellites positions (in cM) along OAR3. In Panel B, the likelihood ratio test statistics of linkage analyses performed on fecal egg count at 1st infection (solid line) and 2nd infection (dashed line) measured on the 332 animals from the selective genotyping design is plotted against microsatellites positions (in cM) along OAR12.

A

B

Figure 2. Comparison of the likelihood ratio test profiles obtained on OAR12 for fecal egg count at (A) first and (B) second infection with the linkage, association, or joint linkage and association analyses with SNP data. In Panel A, the likelihood ratio test statistics of linkage (solid line), association (circles), and joint linkage and association analyses (triangles) that a QTL affects fecal egg count at first infection in the SNP-genotyped backcross population is plotted along OAR12. The estimated suggestive threshold of the linkage analysis is indicated by a dotted horizontal line. In Panel B, the likelihood ratio test statistics of linkage (solid line), association (circles), and joint linkage and association analyses (triangles) that a QTL affects fecal egg count at second infection in the SNP-genotyped backcross population is plotted along OAR12. The estimated suggestive threshold of the linkage analysis is indicated by a dotted horizontal line.

An Extensive Genome Scan for Resistance to *H. contortus*

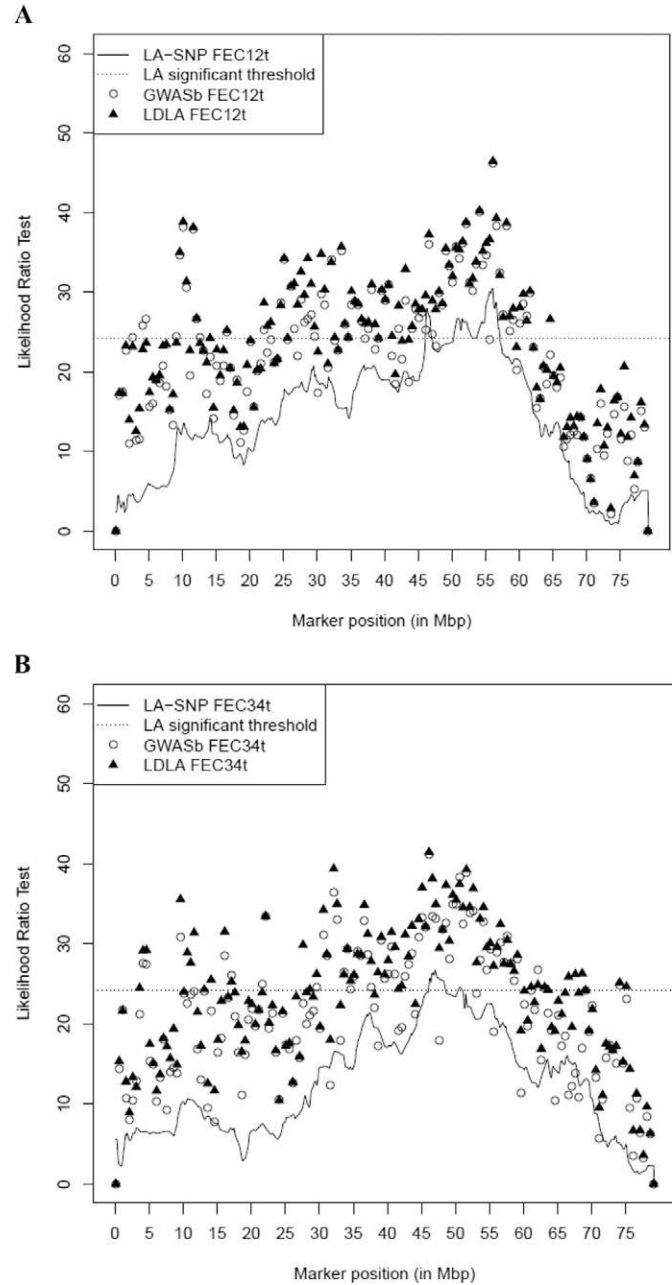In total, LA mapped 23 different QTL regions on 20 chromosomes, and LD-based analysis found 1% genome-wise significant QTL on 22 chromosomes. Results from linkage analysis are in good agreement with other QTL mapping studies in which relatively few significant QTL were detected (Crawford et al., 2006; Davies et al., 2006; Gutierrez-Gil et al., 2009; Dominik et

Table 6. QTL detected in at least one linkage disequilibrium (LD)-based analysis at the 5% genome-wise significance (GWS) threshold

OAR	Trait[1]	Flanking SNP[2]	Pos.[3]	P_{GWAS}[3,4]	P_{GWASb}	P_{LDLAa}	p_{LDLAb}
1	PCV1c	OAR1_45014772.1 – OAR1_45507542.1	43.6			*	
2	Peps1	OAR2_10601656.1 – OAR2_10799411.1	11.7			**	**
4	Len	OAR4_37940876.1 – OAR4_38205790.1	35.2			**	*
4	pHt	OAR4_86339728.1 – OAR4_86628139.1	80.7			**	
5	FEC12t	OAR5_92638062.1 – OAR5_92975517.1	85.1		*	**	
5	IgGst	OAR5_67605574.1 – OAR5_67883800_X.1	62.1			*	
		OAR5_100699982.1 – DU183841_402.1	93.1		*		**
5	PCV2c	OAR5_98137778.1 – OAR5_98330992.1	90.6			**	
5	PCVt_a	OAR5_94568486.1 – OAR5_94826280_X.1	87.1		**	**	**
		s36267.1 – OAR5_96703012.1	89.1	*			
5	Peps1	s06286.1 – s29268.1	94.6			**	
5	Peps2	OAR5_114751184.1 – s38472.1	106.1			**	**
7	FEC12t	OAR7_36947114.1 – OAR7_37429300.1	32.9			**	
7	Len	OAR7_15034944.1 – s39389.1	14.9		**	**	**
		s62332.1 – OAR7_17669851.1	16.9			**	**
8	PCVt_a	OAR8_49745739.1 – OAR8_50006699.1	46.5			**	
9	IgGst	OAR9_85325486.1 – s48117.1	81.0			**	
9	Peps1	OAR9_88974726.1 – OAR9_89299027.1	84.5			**	
12	FEC12t	OAR12_62226914.1 – s68186.1	56.1		*	**	*
12	FEC34t	OAR12_51099743.1 – OAR12_51339516.1	46.1	*	*	**	*
12	FECt_a	OAR12_36077499.1 – s41448.1	32.1	*			
		s23035.1 – OAR12_56589339.1	51.1		**	**	**
12	Peps2	OAR12_40508365.1 – OAR12_41603863.1	36.6			**	
13	FEC34t	s05259.1 – s09612.1	70.7		**	**	**
		s05603.1 – s43133.1	72.2	*		*	
14	IgGmt	OAR14_48832510.1 – s30682.1	46.6			**	
14	PCVt_a	OAR14_1357526.1 – s24656.1	1.1			**	
14	Peps1	OAR14_32272439.1 – OAR14_32706322.1	31.1			**	
15	PCV2c	OAR15_40719719.1 – OAR15_40926306.1	39.2			**	**
15	pHt	s66581.1 – s54590.1	65.2			**	
16	FEC34t	s04660.1 – OAR16_26684182.1	24.7			*	
16	WBt	OAR16_16777017.1 – OAR16_17276041.1	15.7			**	
		s33208.1 – s31223.1	16.7		**	*	*
17	Peps2	OAR17_67650184.1 – s10326.1	62.6			**	
20	FEC12t	OAR20_24357620.1 – OAR20_24626067.1	23.2	*		*	*
20	FEC34t	s69570.1 – OAR20_32868803.1	29.2			*	
20	PCVt_a	s04766.1 – OAR20_44320222.1	40.2			**	
21	IgGmt	s27845.1 – OAR21_14592163.1	12.7			**	
21	Len	OAR21_22572149.1 – OAR21_23049634.1	20.2			**	
21	Peps1	s54156.1 – s26955.1	36.7			**	
		s44626.1 – s61819.1	37.2		**		
21	Peps2	s44626.1 – s61819.1	37.2	**	*	*	*
		s26955.1 – s39524.1	37.7		**	**	
22	Len	OAR22_23894546.1 – OAR22_24105777.1	20.3			**	
23	FEC34t	OAR23_61434545.1 – OAR23_61932991.1	58.3			**	
23	FECt_a	s72843.1 – OAR23_16996616.1	15.8			*	
23	WBt	s13806.1 – s71481.1	48.3	*		**	**
25	PCV2c	OAR25_43636572.1 – s50428.1	40.1			*	
26	IgGmt	OAR26_21857857.1 – OAR26_22456940.1	18.7			**	

*5% genome-wise threshold.

**1% genome-wise threshold.

[1]FEC, fecal egg count, 12 and 34 indicate FEC after 1st and 2nd challenge, respectively; PCV, packed-cell volume, 1 and 2 indicate PCV after 1st and 2nd challenge, respectively, and c indicates values corrected with PCV0 fitted as a covariable; WB, worm burden; SexR, sex ratio in adult worm population; Len, female worm length; pH, abomasal pH; IgGm, IgG in abomasal mucus; IgGs, IgG in serum; Peps, pepsinogen; t, fourth root transformation of the variable; _a, within-animal physiological variation accounted for.

[2]SNP flanking the haplotypes with significant association.

[3]QTL positions, given in Mbp.

[4]GWAS and GWASb stand for association analysis; joint linkage and association analyses are denoted LDLA and LDLAb; the b indicates that haplotypes breed origin was taken into account in the analysis.
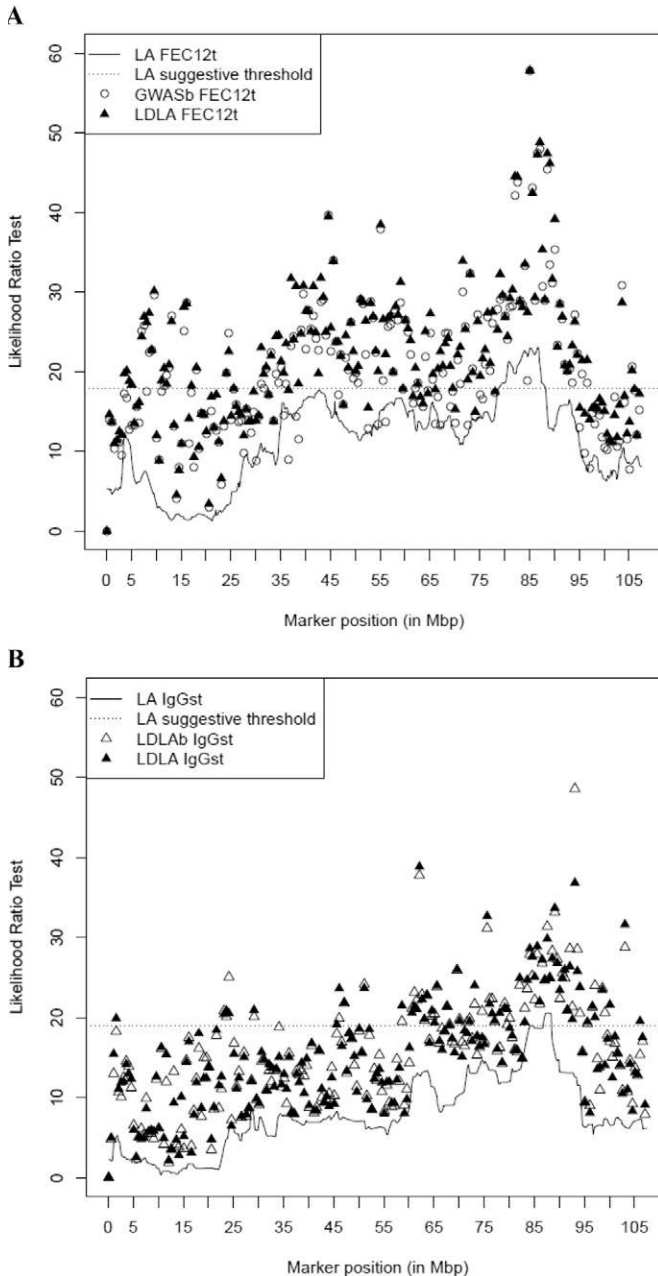
A

B

Figure 3. Likelihood ratio test profiles for fecal egg count at 1st infection and the serum IgG concentration plotted against SNP position on Chromosome 5 for linkage analyses and linkage disequilibrium (LD)-based analyses. In Panel A, the likelihood ratio test statistic that a QTL affects fecal egg count at first infection (FEC12t) is plotted against SNP position (in Mbp) along OAR5. The linkage analysis (solid line), association analysis (circle), and joint linkage and association analysis (filled triangle) are represented. In Panel B, likelihood ratio test profiles of the linkage analysis (solid line) and the joint linkage and association analysis with (LDLAb, filled triangle) or without (LDLA, open triangle) considering the breed origin of the haplotypes are plotted for IgG. The estimated suggestive threshold of the linkage analysis is indicated by a dotted horizontal line.

al., 2010), whereas in some studies only suggestive QTL could be detected (Beh et al., 2002; Beraldi et al., 2007). Using ovine SNP data, Kemper et al. (2011) showed that resistance to GIN followed the same infinitesimal model as other complex traits (Hayes and Goddard, 2001; Cole et al., 2009). Indeed they found numerous loci (99 SNP
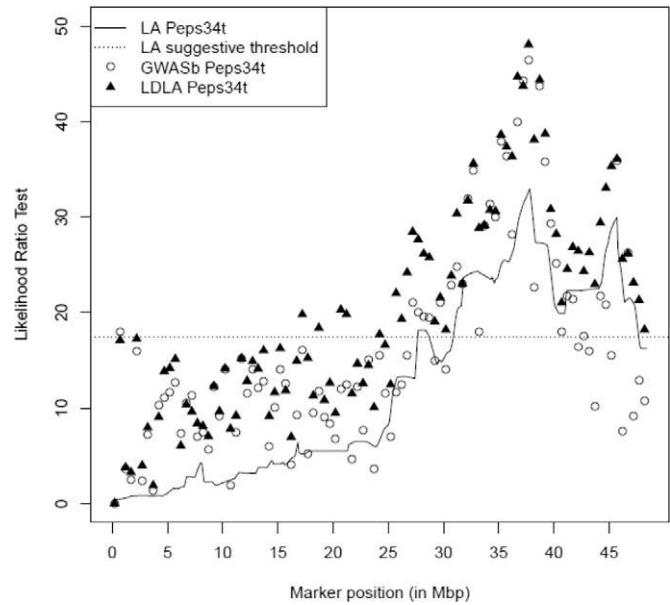

Figure 4. Likelihood ratio test profiles of the linkage analysis with SNP data and both association and joint association and linkage (LDLA) analyses on OAR21 for the pepsinogen concentration measured at 2nd infection. The likelihood ratio test statistic that a QTL affects pepsinogen concentration at 2nd infection is plotted against SNP position (in Mbp) along OAR21 for the linkage analysis (solid line), association analysis (circle), and joint linkage and association analysis (triangle). The estimated suggestive threshold of the linkage analysis is indicated by a dotted horizontal line.

and 65 SNP significantly affecting *T. colubriformis* and *H. contortus* FEC), each exhibiting small to moderate effects (from 0.02% to 0.48% of the phenotypic variation). These effects are in strong contrast to the FEC-associated QTL detected by LA-SNP analysis that explained between 1.7 and 7.5% of the phenotypic variation. By contrast, the QTL effects estimated by LA-SNP analysis were in good agreement with other linkage analyses
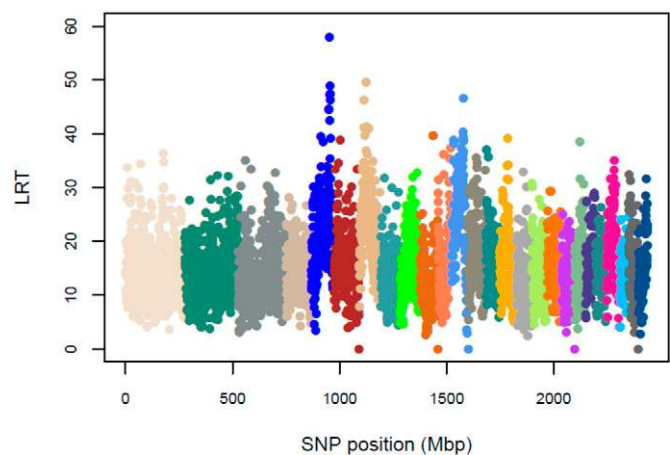

Figure 5. Manhattan plot of the likelihood ratio test values obtained for fecal egg count at 1st infection with the joint association and linkage analysis (LDLA). The likelihood ratio test that a QTL affects FEC at the 1st infection in the backcross population obtained with the LDLA analysis is plotted against the corresponding SNP haplotype position (in Mbp) along the genome. Each plot corresponds to a haplotype of 4 consecutive SNP and each color corresponds to a chromosome from 1 to 26. See online version to view figure in color.
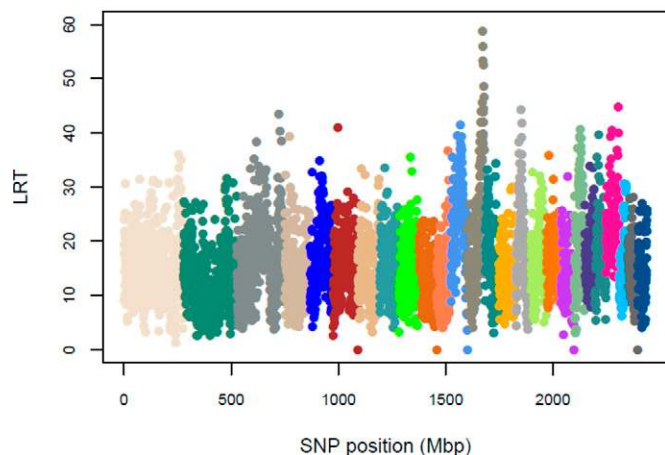
Figure 6. Manhattan plot of the likelihood ratio test values obtained for fecal egg count at 2nd infection with the joint association and linkage analysis (LDLA). The likelihood ratio test that a QTL affects fecal egg count at the 2nd infection in the backcross population obtained with the LDLA analysis is plotted against the corresponding SNP haplotype position (in Mbp) along the genome. Each plot corresponds to a haplotype of 4 consecutive SNP and each color corresponds to a chromosome from 1 to 26. See online version to view figure in color.

(Gutierrez-Gil et al., 2009; Marshall et al., 2009), but somewhat less than the 6 to 13% reported by linkage analysis in other pure breed studies (Davies et al., 2006; Dominik et al., 2010; Matika et al., 2011).

Comparing Linkage Analysis with Microsatellites and SNP

Both the increase in the size of the population studied (from 332 selected extreme BC1 to the 1,000 SNP genotyped BC1 and BC2 individuals) and the denser molecular information led to the detection of 3 times more QTL than in the initial selective genotyping design. The effects of switching from microsatellite to SNP genotypes in terms of power and precision of the linkage analysis was investigated (Supplementary Data 5) as follows. Additional linkage analyses with either microsatellites or SNP genotypes were performed as described in Materials and Methods, using phenotypes of lambs genotyped for both microsatellites and SNP markers (i.e., individuals with extreme FEC_a from 4 families; n = 268). These additional findings showed that the dense genome-wide SNP coverage resulted in the detection of new QTL on chromosomes (i.e., OAR7, 9, 14, 15, and 21) with rather small microsatellite density (on average, 1 microsatellite every 35 cM), . In addition, some of the most consistent QTL regions mapped on OAR5, 12, and 23 with the complete dataset were also found with this reduced dataset whatever the type of markers (Supplementary Data 5). However, the QTL region on OAR3 achieved a *P* value of only 0.073 with SNP data (Supplementary Figure 2) but was found suggestive with microsatellites. Other QTL (on OAR4, 6, 8, 17, 18, 19, and 25) were found with mi-

crosatellite markers in the reduced population only. In the case of OAR3, the observed discrepancy might therefore originate in the molecular data itself. For the 7 other QTL that were found with microsatellites in the reduced population, the sampling of progeny may have resulted in an overestimation of the QTL effects known as the Beavis effect (Beavis, 1994). This kind of result was also obtained by Melchinger et al. (1998), who detected some QTL in a small plant population that could not be found in a greater population. In linkage analysis, the precision of mapping is mostly driven by the numbers of observed meioses and the QTL effect (Darvasi et al., 1993). In addition, a relatively large microsatellite density had been achieved on OAR5, 12, and 23 (1 microsatellite every 6 to 12 cM), and SNP genotyping was not expected to provide a greatly improved precision. However, the increased marker density provided by SNP markers slightly narrowed down the confidence intervals of the QTL found in common with microsatellites and SNP on OAR5, 12, and 23 (between –13 Mbp to –47.7 Mbp from microsatellites to SNP markers, data not shown).

SNP Unlock the Potential of Existing LD

The LD extent in sheep, estimated with either microsatellites (McRae et al., 2002; Meadows et al., 2008) or SNP markers (Kemper et al., 2011) is small, and the r^2 values estimated in MBB and RMN breeds followed the same pattern. Thus, a dense SNP map was necessary to take advantage of existing LD for mapping purposes. In addition, our population was a mixture of 2 breeds as illustrated by the rapid decrease of LD phase persistence between MBB and RMN breeds with genomic distances (Supplementary Figure 1C). This is consistent with the results of de Roos et al. (2008) in Australian Angus and New Zealand Holstein cows. Moreover, even at short distances (<10 Kb,) some correlations between *r* exhibited opposite signs.

Therefore, 2 types of analyses were performed to take this small LD phase persistence between the 2 pure breeds into account. The GWAS and LDLA did not consider the breed origin of the haplotypes and, hence, focused either on QTL with the same effects in the 2 breeds (both MBB and RMN haplotypes were considered to be the same) or QTL segregating in the RMN breed only (because the frequency in the founders of any eventual MBB-specific haplotypes would be too small to be considered in the analysis). By contrast, GWASb and LDLAb aimed at detecting breed specific QTL in accordance with the hypothesis that a different genetic background was responsible for the resistance of MBB to *H. contortus* infection (Yazwinski et al., 1980; Courtney et al., 1985; Aumont et al., 2003). This hypothesis was confirmed for the LA-SNP QTL located on OAR5, OAR12, and OAR13 because the alleles originating from the MBB breed were linked to smaller FEC (the

difference between the RMN grandmaternal chromosome effect and the MBB grandpaternal chromosome effect being positive). Reinforcing these findings, the GWASb model identified 6 QTL on OAR5, 7, 12, 13, 16, and 21 that could not be detected in the GWAS model. In addition, FEC-related QTL and the QTL for female worm length on OAR7 were related to identical-by-state alleles with opposite effects in each breed (data not shown). However, OAR16 exhibited an opposite pattern and suggested that some part of the RMN genome has a better ability to reduce FEC. This finding corroborates the findings of Terefe et al. (2007), who demonstrated that RMN lambs controlled the parasite more efficiently upon reinfection in comparison with the first infection.

Interestingly, a MBB-specific haplotype mapped at 93.1 Mbp on OAR5 exerted the most favorable effect on IgGst. This QTL was detected by GWASb and LDLAb analyses only, hence illustrating the benefits of taking into account the breed origin of the haplotypes. Indeed, LDLA analysis resulted in a lesser LRT than LDLAb (10 points less) and shifted the maximal LRT value 30 Mbp away at 62.1 Mbp (see Figure 3B). It might be possible that 2 different regions, one carrying an old polymorphism common to both breeds (at 62.1 Mbp) and the other a much more recent mutation (at 93.1 Mbp), may affect IgGst. The older mutation might tend to fixation and, hence, show smaller variations between the possible alleles than the more recent mutation.

Contrasting with this QTL, every other GWASb-specific QTL was also detected by LDLA analysis. The LDLA was expected to have more power than GWAS analyses because it benefits both from the precision of association and the robustness of linkage (Meuwissen and Goddard, 2007). Interestingly, LDLA and LDLAb showed similar LRT profiles (except on OAR5 for IgGst). Both methods are based on the same model: a within-family QTL effect that provides the most important features of the differences between MBB and RMN breeds, and paternal and maternal haplotypic effects. The additional information provided by association, therefore, comes from the RMN haplotypes. Indeed, only 4 F_1 sires were available, thus providing 4 different MBB haplotypes segregating at the most, and this did not contribute much to the likelihood. In turn, this leads to few differences in the test statistics of the 2 models, with more degrees of freedom for LDLAb. Subsequently, less QTL are declared significant with LDLAb than with LDLA.

Consistent Findings in Three Regions

Both LA- and LD-based analysis resulted in mapping many QTL involved in resistance to *H. contortus*. However, only a few regions exhibited consistent findings, suggesting the actual presence of a QTL: several correlated traits affected by a limited region (OAR5, 12, and 21) and large significance (OAR12 and 21) or sim-

ply a large significance on OAR13.

It is interesting to note that a limited region of OAR12 (between 46 and 56 Mbp) was associated with each of the FEC traits, whatever the infection rank by *H. contortus*, and was confirmed in the BC2 population. The QTL affecting FEC were also mapped on OAR12 in a Merino flock (Beh et al., 2002) and in the free-living Soay sheep population (Beraldi et al., 2007). In the latter, the QTL was mapped between CSSM3 and MCMA52 located at 40.4 Mbp and 60.8 Mbp, respectively (Archibald et al., 2010), which is in good agreement with the results reported herein. Taken together these findings suggest that OAR12 is a key player affecting the excretion of *H. contortus* eggs. However, LD in our population did not help refine the location of this QTL. The GWASb confirmed that MBB haplotypes had the most favorable effect, whereas haplotypes linked to susceptibility segregated in the RMN breed (data not shown). It is therefore difficult to determine whether only one gene underlies this QTL region or whether several genes are present.

By contrast, both OAR5 and OAR13 benefited from LD information; the LRT statistic was affected by the addition of RMN haplotypes to the model, hence greatly enhancing the original signals observed by LA analysis. On OAR5, several peaks were present between 85 and 93 Mbp and were consistently associated with FEC, IgGst, or PCV. Here again, it is difficult to determine whether a unique gene with pleiotropic effects is present. By contrast, the 2-Mbp region between 70.2 and 72.2 Mbp on OAR13 was significantly associated with FEC34t in every analysis. This is a particularly interesting finding because the region is small. Given the QTL effect on FEC34t and the relatively large correlations with other traits (e.g., WBt and PCV2), one would have expected these latter traits to be affected by this region.

First Report of a Significant QTL Affecting Pepsinogen Concentration

The BC2 lambs were measured for pepsinogenemia that is associated with abomasal mucosa disruption during hemonchosis (Simpson, 2000). Interestingly, at reinfection, smaller FEC values were observed in lambs that had greater pepsinogenemia as described by Terefe et al. (2007), who compared the immune responses to *H. contortus* infection in MBB and RMN lambs. Corroborating these findings, a QTL region was mapped on OAR21 for both Peps34t and FECt_a. In addition, multiple associations with different LD-based analyses highlighted a 1-Mbp region that was significantly associated with pepsinogen concentration upon both first and second infections. As far as we know, this is the first significant QTL associated with pepsinogen concentration during GIN infection in sheep. Gutierrez-Gil et al. (2009) measured this trait in a Churra ewe granddaugh-

ter design but did not detect any QTL that segregated for this trait. This difference in results may be explained by our experimental design because infection, as well as the time point of pepsinogen measurement, was strictly controlled. In addition, the GWASb model showed that MBB haplotypes were not found in the RMN breed and were associated with greater pepsinogen concentrations. Interestingly, the bovine *pepsinogen A* locus (*PGA5*) has been positioned between 37,514,983 and 37,525,605 bp (Archibald et al., 2010) and hence constitutes the most obvious functional candidate gene. Several other genes related to the immune response are located in the vicinity of this gene: CD5 known to participate in the selection and activation of both B and T cells and CD6, which is involved in the inhibition of macrophage apoptosis. Additionally, Dominik et al. (2010) found a QTL for the variation in blood eosinophil counts following a 1st infection by *T. colubriformis* in the same region: between CSRD72 and BMS1948 located, respectively, at 40.3 and 51.2 Mbp on OAR21 (Archibald et al., 2010). Even if worm genera are different, eosinophils are known to be key players in the rejection of worms (Meeusen and Balic, 2000; Terefe et al., 2009). In addition, Terefe et al. (2007) reported greater eosinophil counts and greater pepsinogen concentrations in MBB purebred animals, whatever the time of infection by *H. contortus*. Together, these independent findings suggest that a relatively small region on OAR21 might affect blood eosinophils, pepsinogen concentration, and FEC during GIN infection.

Conclusion

In summary, this is one of the most extensive descriptions of *H. contortus* infection in sheep and one of the first SNP-based QTL detection studies for *H. contortus* in sheep.

Many QTL of small to moderate effects were found, and 4 main regions on OAR5, 12, 13, and 21 were identified. The LD phase persistence between the 2 breeds constitutive of the backcross population was small, and some haplotypes were breed-specific. Therefore, a specific model was implemented to cluster haplotypes according to their breed origin, while looking for haplotype-trait associations. This model fitted better reality and outperformed the simple association model that did not take breed origin into account. The chromosomes coming from the MBB breed were usually associated with the most favorable effects. The QTL mapped on OAR5 and OAR12 were found with microsatellite and SNP markers, the confidence intervals being narrower in the latter case.

We report the first QTL affecting pepsinogen concentration, with the region of greatest significance mapping precisely to the *PGA5* gene on OAR21. The OAR5 and OAR13 showed QTL with large effects or pleiotropic effects or both that could not be matched to any known functional candidate genes. The OAR12 remains an interesting can-

didate because a 10-Mbp region was consistently involved upon both first and second infections. Further work would be required to dissect the underlying genetic mechanisms.

LITERATURE CITED

Amarante, A. F. T., P. A. Bricarello, R. A. Rocha, and S. M. Gennari. 2004. Resistance of Santa Ines, Suffolk and Ile de France sheep to naturally acquired gastrointestinal nematode infections. Vet. Parasitol. 120:91–106.

Archibald, A. L., N. E. Cockett, B. P. Dalrymple, T. Faraut, J. W. Kijas, J. F. Maddox, J. C. McEwan, V. Hutton Oddy, H. W. Raadsma, C. Wade, J. Wang, W. Wang, and X. Xun. 2010. The sheep genome reference sequence: A work in progress. Anim. Genet. 41:449–453.

Aumont, G., L. Gruner, and G. Hostache. 2003. Comparison of the resistance to sympatric and allopatric isolates of Haemonchus contortus of Black Belly sheep in Guadeloupe (FWI) and of INRA 401 sheep in France. Vet. Parasitol. 116:139–150.

Beavis, W. D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. Pages 250–256 in The Proceedings of the Forty-Ninth Annual Corn & Sorghum Industry Research Conference. American Seed Trade Assoc., Washington, DC.

Beh, K. J., D. J. Hulme, M. J. Callaghan, Z. Leish, I. Lenane, R. G. Windon, and J. F. Maddox. 2002. A genome scan for quantitative trait loci affecting resistance to Trichostrongylus colubriformis in sheep. Animal Genetics 33:97–106.

Beraldi, D., A. F. McRae, J. Gratten, J. G. Pilkington, J. Slate, P. M. Visscher, and J. M. Pemberton. 2007. Quantitative trait loci (QTL) mapping of resistance to strongyles and coccidia in the free-living Soay sheep (*Ovis aries*). Int. J. Parasitol. 37:121–129.

Bishop, S. C., and C. A. Morris. 2007. Genetics of disease resistance in sheep and goats. Small Ruminant Res. 70:48–59.

Bovenhuis, H., and R. J. Spelman. 2000. Selective genotyping to detect quantitative trait loci for multiple traits in outbred populations. J. Dairy Sci. 83:173–180.

Cezar, A. S., G. Toscan, G. Camillo, L. A. Sangioni, H. O. Ribas, and F. S. Vogel. 2010. Multiple resistance of gastrointestinal nematodes to nine different drugs in a sheep flock in southern Brazil. Vet. Parasitol. 173:157–160.

Churchill, G. A., and R. W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963–971.

Cole, J. B., P. M. VanRaden, J. R. O'Connell, C. P. Van Tassell, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and G. R. Wiggans. 2009. Distribution and location of genetic effects for dairy traits. J Dairy Sci. 92:2931–2946.

Coop, R. L., W. D. Smith, K. W. Angus, R. B. Graham, S. E. Wright, and F. Jackson. 1985. Effect of *Ostertagia ostertagi* on lamb performance and cross resistance to *O. circumcincta*. Res. Vet. Sci. 39:200–206.

Courtney, C. H., C. F. Parker, K. E. McClure, and R. P. Herd. 1985. Resistance of nonlambing exotic and domestic ewes to naturally acquired gastrointestinal nematodes. Int. J. Parasitol. 15:239–243.

Crawford, A. M., K. A. Paterson, K. G. Dodds, C. Diez Tascon, P. A. Williamson, M. Roberts Thomson, S. A. Bisset, A. E. Beattie, G. J. Greer, R. S. Green, R. Wheeler, R. J. Shaw, K. Knowler, J. C. McEwan. 2006. Discovery of quantitative trait loci for resistance to parasitic nematode infection in sheep: I. Analysis of outcross pedigrees. BMC Genomics 7:178.

Darvasi, A., A. Weinreb, V. Minke, J. I. Weller, and M. Soller. 1993. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics 134:943–951.

Davies, G., M. J. Stear, M. Benothman, O. Abuagob, A. Kerr, S. Mitchell, and S. C. Bishop. 2006. Quantitative trait loci associated with

parasitic infection in Scottish blackface sheep. Heredity 96:252–258.

Dominik, S., P. W. Hunt, J. McNally, A. Murrell, A. Hall, and I. W. Purvis. 2010. Detection of quantitative trait loci for internal parasite resistance in sheep. I. Linkage analysis in a Romney × Merino sheep backcross population. Parasitology 137:1275–1282.

Dorny, P., and J. Vercruysse. 1998. Evaluation of a micro method for the routine determination of serum pepsinogen in cattle. Res. Vet. Sci. 65:259–262.

Elsen, J. M., B. Mangin, B. Goffinet, D. Boichard, and P. Le Roy. 1999. Alternative models for QTL detection in livestock. I. General introduction. Genet. Sel. Evol. 31:213.

Good, B., J. P. Hanrahan, B. A. Crowley, and G. Mulcahy. 2006. Texel sheep are more resistant to natural nematode challenge than Suffolk sheep based on faecal egg count and nematode burden. Vet. Parasitol. 136:317–327.

Green, P., K. Falls, and S. Crooks. 1990. Documentation for CRI-MAP. v. 2.4. Washington Univ. School of Medicine, St. Louis, MO.

Gutierrez-Gil, B., J. Perez, L. Alvarez, M. Martinez-Valladares, L. F. de la Fuente, Y. Bayon, A. Meana, F. San Primitivo, F. A. Rojo-Vazquez, and J. J. Arranz. 2009. Quantitative trait loci for resistance to trichostrongylid infection in Spanish Churra sheep. Genet. Sel. Evol. 41:46.

Hayes, B., and M. E. Goddard. 2001. The distribution of the effects of genes affecting quantitative traits in livestock. Genet. Sel. Evol. 33:209–229.

Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38:226–231.

Hoglund, J., K. Gustafsson, B. L. Ljungstrom, A. Engstrom, A. Donnan and P. Skuce. 2009. Anthelmintic resistance in Swedish sheep flocks based on a comparison of the results from the faecal egg count reduction test and resistant allele frequencies of the beta-tubulin gene. Vet Parasitol. 161:60–68.

Howell, S. B., J. M. Burke, J. E. Miller, T. H. Terrill, E. Valencia, M. J. Williams, L. H. Williamson, A. M. Zajac, and R. M. Kaplan. 2008. Prevalence of anthelmintic resistance on sheep and goat farms in the southeastern United States. J. Am. Vet. Med. Assoc. 233:1913–1919.

Kemper, K. E., D. L. Emery, S. C. Bishop, H. Oddy, B. J. Hayes, S. Dominik, J. M. Henshall, and M. E. Goddard. 2011. The distribution of SNP marker effects for faecal worm egg count in sheep, and the feasibility of using these markers to predict genetic merit for resistance to worm infections. Genet Res. 93:203–219.

Knott, S. A., L. Marklund, C. S. Haley, K. Andersson, W. Davies, H. Ellegren, I. Fredholm, I. Hansson, B. Hoyheim, K. Lundström, M. Moller, and L. Andersson. 1998. Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and large white pigs. Genetics 149:1069–1080.

Lacroux, C., T. H. Nguyen, O. Andreoletti, F. Prevot, C. Grisez, J. P. Bergeaud, L. Gruner, J. C. Brunel, D. Francois, P. Dorchies, P. Jacquiet. 2006. Haemonchus contortus (Nematoda: Trichostrongylidae) infection in lambs elicits an unequivocal Th2 immune response. Vet. Res. 37:607–622.

Lander, E. S., and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199.

Lander, E., and L. Kruglyak. 1995. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. Nat Genet. 11:241–247.

Legarra, A., and R. L. Fernando. 2009. Linear models for joint association and linkage QTL mapping. Genet. Sel. Evol. 41:43.

Lynch, M., and B. Walsh. 1998. Genetics Analysis of Quantitative Traits. Sinauer Associates, Sunderland, MA.

Mandonnet, N., M. Bachand, M. Mahieu, R. Arquet, F. Baudron, L. Abinne-Molza, H. Varo, and G. Aumont. 2005. Impact on productivity of peri-parturient rise in fecal egg counts in Creole goats in the humid tropics. Vet. Parasitol. 134:249–259.

Mandonnet, N., V. Ducrocq, R. Arquet, and G. Aumont. 2003. Mortality of Creole kids during infection with gastrointestinal strongyles: A survival analysis. J. Anim. Sci. 81:2401–2408.

Marshall, K., J. F. Maddox, S. H. Lee, Y. Zhang, L. Kahn, H.-U. Graser, C. Gondro, S. W. Walkden-Brown, and J. H. J. van der Werf. 2009. Genetic mapping of quantitative trait loci for resistance to Haemonchus contortus in sheep. Anim. Genet. 40:262–272.

Matika, O., R. Pong-Wong, J. A. Wooliams, and S. C. Bishop. 2011. Confirmation of two quantitative trait loci regions for nematode resistance in commercial British terminal sire breeds. Animal 5:1149–1156.

McRae, A. F., J. C. McEwan, K. G. Dodds, T. Wilson, A. M. Crawford, and J. Slate. 2002. Linkage disequilibrium in domestic sheep. Genetics 160:1113–1122.

Meadows, J. R., E. K. Chan, and J. W. Kijas. 2008. Linkage disequilibrium compared between five populations of domestic sheep. BMC Genet. 9:61.

Meeusen, E. N., and A. Balic. 2000. Do eosinophils have a role in the killing of helminth parasites? Parasitol. Today 16:95–101.

Melchinger, A. E., H. F. Utz, and C. C. Schön. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403.

Meuwissen, T. H. E., and M. E. Goddard. 2007. Multipoint identity-by-descent prediction using dense markers to map quantitative trait loci and estimate effective population size. Genetics 176:2551–2560.

Ooijen, J. 1992. Accuracy of mapping quantitative trait loci in autogamous species. Theor. Appl. Genet. 84:803–811.

Pérez-Enciso, M., and L. Varona. 2000. Quantitative trait loci mapping in F2 crosses between outbred lines. Genetics 155:391–405.

Piepho, H. P. 2001. A quick method for computing approximate thresholds for quantitative trait loci detection. Genetics 157:425–432.

Raynaud, J. P. 1970. Etude de l'efficacité d'une technique de coproscopie quantitative pour le diagnostic de routine et le contrôle des infestations parasitaires des bovins, ovins, équins et porcins. Ann. Parasitol. 45:321–342.

de Roos, A. P., B. J. Hayes, R. J. Spelman, and M. E. Goddard. 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. Genetics 179:1503–1512.

Sargison, N. D., F. Jackson, D. J. Bartley, D. J. Wilson, L. J. Stenhouse, and C. D. Penny. 2007. Observations on the emergence of multiple anthelmintic resistance in sheep flocks in the south-east of Scotland. Vet. Parasitol. 145:65–76.

Simpson, H. V. 2000. Pathophysiology of abomasal parasitism: is the host or parasite responsible? Vet J. 160:177–191.

Stear, M. J., M. Parkand, S. C. Bishop. 1996. The key components of resistance to Ostertagia circumcincta in lambs. Parasitol. Today 12:438–441.

Terefe, G., C. Lacroux, O. Andreoletti, C. Grisez, F. Prevot, J. P. Bergeaud, J. Penicaud, V. Rouillon, L. Gruner, J. C. Brunel, A. Legarra, J. M. Elsen, J. Bouix, R. Rupp, and C. R. Moreno. 2007. Immune response to Haemonchus contortus infection in susceptible (INRA 401) and resistant (Barbados Black Belly) breeds of lambs. Parasite Immunol. 29:415–424.

Terefe, G., C. Lacroux, F. Prévot, C. Grisez, J. P. Bergeaud, C. Bleuart, P. Dorchies, G. Foucras, and P. Jacquiet. 2009. Eosinophils in Haemonchus contortus-infected resistant and susceptible breeds of sheep: Abomasal tissue recruitment and in vitro functional state. Vet. Parasitol. 165:161–164.

Traversa, D., B. Paoletti, D. Otranto, and J. Miller. 2007. First report of multiple drug resistance in trichostrongyles affecting sheep under field conditions in Italy. Parasitol Res. 101:1713–1716.

Yazwinski, T. A., L. Goode, D. J. Moncol, G. W. Morgan, and A. C. Linnerud. 1980. *Haemonchus contortus* resistance in straightbred and crossbred Barbados Blackbelly sheep. J Anim. Sci. 51:279–284.

APPENDIX 1: EXPRESSION OF THE TRUE CORRELATION OF VARIABLES MEASURED ON A SELECTIVE SET OF DATA

Assuming Variable A is measured on the complete data set. Because Variables B and C are only measured on a selected Group G from A, the aim of this appendix is to estimate the true correlations between A, B, and C using observations in the Group G: A_G, B_G, and C_G.

In general, the correlation between Variables A and B is:

$$r(A,B) = \beta(A,B)\frac{\sigma_A}{\sigma_B}$$

where σ_A and σ_B are the SD of Variables A and B and $\beta(A,B)$ is the regression coefficient of A on B.

Because the regression coefficient is not affected by the data selection:

$$r(A,B) = r(A_G, B_G)\frac{\sigma_{B_G}.\sigma_A}{\sigma_{A_G}\sigma_B}$$

The unknown parameter in this equation is σ_B. $\sigma_B^2 = \sigma_{eB}^2 + \beta^2(A,B)\sigma_A^2$, σ_{eB} is the residual variance of the regression of B on A. Because this variance is not affected by the selection of A: $\sigma_{B_G}^2 = \sigma_{eB}^2 + \beta^2(A_G,B_G)\sigma_{A_G}^2$, then $\sigma_B^2 = \sigma_{B_G}^2 + \hat{a}^2(A_G,B_G)(\sigma_A^2 - \sigma_{A_G}^2)$.

Finally, the correlation between A and B is:

$$r(A,B) = r(A_S, B_S).\frac{\sigma_{B_G}.\sigma_A}{\sigma_{A_G}\sqrt{\left[\sigma_{B_G}^2 + \beta^2(A_G,B_G).(\sigma_A^2 - \sigma_{A_G}^2)\right]}}$$

The determination of the correlation between Variables B and C is complicated by the fact that both variables are measured on the selected Group G.

$$r(B,C) = \frac{\text{cov}(B,C)}{\sigma_B\sigma_C},$$

Assuming $D = B + C$, then $2\,\text{cov}(B,C) = \sigma_D^2 - (\sigma_B^2 + \sigma_C^2)$ and $\beta(A,D) = \beta(A,B) + \beta(A,C)$.

$$2\,\text{cov}(B,C) = \sigma_{D_G}^2 -$$
$$\left\{\beta^2(A_G,D_G) - \left[\sigma_{B_G}^2 - \beta^2(A_G,B_G) + \sigma_{C_G}^2 - \beta^2(A_G,C_G)\right]\right\} \times$$
$$(\sigma_A^2 - \sigma_{A_G}^2)$$

$$\text{cov}(B,C) = \text{cov}(B_G,C_G) - \beta(A_G,B_G)\beta(A_G,C_G)(\sigma_A^2 - \sigma_{A_G}^2)$$

$$\text{cov}(B,C) = r(B_G,C_G)\sigma_{B_G}\sigma_{C_G} - \beta(A,B_G)\beta(A,C_G)(\sigma_A^2 - \sigma_{A_G}^2)$$

$$r(B,C) = \frac{r(B_G,C_G)\sigma_{B_G}\sigma_{C_G} - \beta(A,B_G)\beta(A,C_G)(\sigma_A^2 - \sigma_{A_G}^2)}{\sigma_B\sigma_C}$$

$$r(B,C) = \frac{r(B_G,C_G)\sigma_{B_G}\sigma_{C_G} - \beta(A,B_G)\beta(A,C_G)(\sigma_A^2 - \sigma_{A_G}^2)}{\sqrt{\left[\sigma_{B_G}^2 - \beta(A,B)(\sigma_A^2 - \sigma_{A_G}^2)\right]\left[\sigma_{C_G}^2 - \beta(A,C)(\sigma_A^2 - \sigma_{A_G}^2)\right]}}$$

APPENDIX 2: TRUE QTL EFFECT UNDER SELECTIVE GENOTYPING

As previously, A is measured on the complete data set, and marker genotyping and B phenotyping were performed only on the Group G: the proportion p of the low and high tails of the distribution of A. In this context, the true estimations of the QTL effect in a sire QTL design for A and B were proposed by Bovenhuis and Spelman (2000).

QTL Effect on A under Selective Genotyping on A

Assuming a QTL is segregating in the population, the true additive QTL effect is noted a, and the estimated QTL effect in the selective genotyping group G is a_G. Consequently, the estimated QTL effect by Qtlmap software can be corrected following the proposition of by Bovenhuis and Spelman (2000):

$$a = \frac{a_G}{\left(1 + Z_{1-p/2}i_{p/2}\right)}$$

where $z_{1-p/2}$ is the deviation of the truncation point from the mean corresponding to p/2; and $i_{p/2}$ is the selection intensity corresponding to p/2.

QTL Effect on B under Selective Genotyping on A

Assuming the true additive QTL effect on B is noted b, and the estimated QTL effect in the selective genotyping group G is b_G. Because the regression coefficient is not affected by the data selection, Bovenhuis and Spelman (2000) showed that:

$$b = b_G - \beta(A,B)\left(\frac{Z_{1-p/2}i_{p/2}a_G}{1 + Z_{1-p/2}i_{p/2}}\right)$$

$\beta(A,B)$ is the regression coefficient of A on B.