J. Dairy Sci. 96:6059–6069
http://dx.doi.org/10.3168/jds.2013-6824


# Identification of quantitative trait loci underlying milk traits in Spanish dairy sheep using linkage plus combined linkage disequilibrium and linkage analysis approaches

**E. Garcia-Gámez,[1] B. Gutiérrez-Gil,[1] A. Suarez-Vega, L. F. de la Fuente, and J. J. Arranz[2]**
Departamento de Producción Animal, Facultad de Veterinaria, Universidad de León, 24071 León, Spain

## ABSTRACT

In this study, 2 procedures were used to analyze a data set from a whole-genome scan, one based on linkage analysis information and the other combing linkage disequilibrium and linkage analysis (LDLA), to determine the quantitative trait loci (QTL) influencing milk production traits in sheep. A total of 1,696 animals from 16 half-sib families were genotyped using the OvineSNP50 BeadChip (Illumina Inc., San Diego, CA) and analysis was performed using a daughter design. Moreover, the same data set has been previously investigated through a genome-wide association (GWA) analysis and a comparison of results from the 3 methods has been possible. The linkage analysis and LDLA methodologies yielded different results, although some significantly associated regions were common to both procedures. The linkage analysis detected 3 overlapping genome-wise significant QTL on sheep chromosome (OAR) 2 influencing milk yield, protein yield, and fat yield, whereas 34 genome-wise significant QTL regions were detected using the LDLA approach. The most significant QTL for protein and fat percentages was detected on OAR3, which was reported in a previous GWA analysis. Both the linkage analysis and LDLA identified many other chromosome-wise significant associations across different sheep autosomes. Additional analyses were performed on OAR2 and OAR3 to determine the possible causality of the most significant polymorphisms identified for these genetic effects by the previously reported GWA analysis. For OAR3, the analyses demonstrated additional genetic proof of the causality previously suggested by our group for a single nucleotide polymorphism located in the α-lactalbumin gene (*LALBA*). In summary, although the results shown here suggest that in commercial dairy populations, the LDLA method exhibits a higher efficiency to map QTL than the simple linkage analysis or linkage disequilibrium methods, we believe that comparing the 3 analysis methods is the best approach to obtain a global picture of all identifiable QTL segregating in the population at both family-based and population-based levels.
**Key words:** sheep, milk trait, quantitative trait loci, combined linkage disequilibrium and linkage analysis method

## INTRODUCTION

For several years, mapping the QTL of economic interest in livestock species using genetic markers was based on pedigree or family information (Georges, 2007). In commercial dairy populations, paternal half-sib pedigrees, generated using AI, have been used to investigate the co-segregation of genetic markers and QTL in large half-sib families (daughter or granddaughter designs; Weller et al., 1990; Georges et al., 1995). More recently, due to the development of next-generation sequencing technologies, the genomes of many livestock species have been sequenced. These advances have been accompanied by the availability of cost-efficient, high-throughput genotyping platforms that allow the investigation of the genomes at a high resolution to identify genetic variants for traits of economic and biological interest.

In dairy cattle, since the availability of SNP array panels, genome-wide association (**GWA**) analyses have replaced traditional linkage mapping studies (Mai et al., 2010; Schopen et al., 2011). These GWA studies, which are ideally conducted in unrelated individuals, and the increased marker density offered by the high-throughput genotyping platforms allow the analysis of population-wide linkage disequilibrium (**LD**). Because LD generally extends over much shorter distances than the genetic linkage analyzed in classical QTL detection studies based on linkage analysis, these LD-based analyses allow the identification of reduced confidence intervals for the QTL location compared with estimates based on classical linkage analyses (Weller and Ron, 2011).

Received March 18, 2013.
Accepted May 22, 2013.
[1]Both authors contributed equally to this work.
[2]Corresponding author: jjarrs@unileon.es

Nevertheless, the power of GWA studies depends on the extent of the LD across the genome (Andersson, 2009). Investigations in cattle have demonstrated that the LD varies along the genome and that a certain degree of LD exists between markers on different chromosomes (Farnir et al., 2002; McRae et al., 2002). For the populations structured as half-sib families, several studies have suggested that the combination of LD and linkage analysis information (**LDLA**) enhances the feasibility of fine-mapping QTL regions (Meuwissen et al., 2001, 2002; Farnir et al., 2002). Instead of using single SNP information, the LDLA approach is based on marker haplotypes and the construction of an identity-by-descent matrix. In a half-sib design, the linkage information is provided by the paternal haplotypes of the progeny, whereas both the paternal and maternal haplotypes of the progeny provide the LD information. To implement LDLA successfully in this type of experimental design, the maternal haplotypes should be a good representation of the whole-population haplotypes (Meuwissen et al., 2002). Additionally, an optimal number of half-sib families should be included in the experiment. In a simulation study, Roldan et al. (2012) demonstrated that the analysis of 15 to 50 medium-sized half-sib families (between 20 and 65 daughters per sire) produced more accurate QTL mapping results than a small number of large families (5 sires with 200 progeny each) or a large number of small half-sib families (100 sires with 10 progeny each).

In sheep, linkage-based genome scans used to identify QTL underlying milk production traits have been performed using microsatellite data in various commercial (Barillet et al., 2006; Gutiérrez-Gil et al., 2009) and experimental populations (Raadsma et al., 2009; Mateescu and Thonney, 2010). The QTL detected by these classical linkage-mapping experiments are generally population specific and exhibit low significance levels and large confidence intervals (Carta et al., 2009). Based on the availability of the OvineSNP50 BeadChip (Illumina Inc., San Diego, CA), our group recently published a GWA study for milk production traits in a commercial Spanish Churra sheep population (García-Gámez et al., 2012a). In the GWA study, in addition to 14 chromosome-wise significant associations, we identified 2 overlapping experiment-wise significant associations on sheep chromosome (**OAR**) 3, which were coincident with a previously identified QTL for milk protein percentage in Churra sheep (Gutiérrez-Gil et al., 2009). For this QTL, the increased marker density in the GWA study allowed for the identification of the first putative causative mutation (quantitative trait nucleotide, **QTN**) for milk traits in dairy sheep (García-Gámez et al., 2012a). However, the analysis of the SNP array panel data set in Churra sheep demonstrated that the LD magnitude is much shorter than observed in cattle and several other sheep breeds (García-Gámez et al., 2012c), which suggests that in this population, based on the number of markers provided by the OvineSNP50 BeadChip, the exclusive use of the LD-based analysis may demonstrate several limitations for the detection of causal mutations.

Based on this observation, the same data set used by García-Gámez et al. (2012a) was analyzed in this study to exploit the half-sib family structure of this commercial population for the identification of QTL. The results of the SNP-based whole genome scan subjected to 2 different analysis methods (the linkage analysis and LDLA approaches) were analyzed to identify the QTL influencing milk production traits. The results were compared with the GWA study reported previously (García-Gámez et al., 2012a), allowing the comparison of the 3 analysis methods and the assessment of the information that linkage-based analyses can produce in this type of commercial populations versus the use of single population-based information.

## MATERIALS AND METHODS

### *Resource Population, Phenotypes, and Marker Map*

The Spanish Churra commercial population analyzed in this study has been described in García-Gámez et al. (2012a). In total, 1,696 animals distributed in 16 half-sib families, including 1,680 ewes and the corresponding 16 rams, were commercially genotyped using the OvineSNP50 BeadChip at AROS Applied Biotechnology AS (Aarhus, Denmark) and Laboratoire d'Analyses Génétiques pour les Espèces Animales (LABOGENA; Jouy-en-Josas, France). After a quality control analysis per animal (call rate >90%) and per SNP (call rate >95%; minor allele frequency >0.05; correspondence with Hardy-Weinberg equilibrium: $P > 0.00001$), 1,681 animals, including the 16 rams and 43,784 autosomal SNP, were retained for the analysis.

The phenotypes included in the analysis were the yield deviations (**YD**) corresponding to the milk production traits routinely collected by the National Association of Churra Breeders as follows: milk yield (**MY**), protein percentage (**PP**), fat percentage (**FP**), protein yield (**PY**), and fat yield (**FY**). The YD estimates were calculated following a multivariate animal repeatability model in which the raw phenotypes were corrected for the environmental effects of herd test day, birth order, age of the ewe at parturition (as a covariate nested within birth order), number of born lambs, number of weeks of milk production of the ewe, and the ewe's permanent environmental effect.