DNA variation in the gene ELTD1 is associated with tick burden in cattle

L. R. Porto Neto*,†, R. J. Bunch*, B. E. Harrison* and W. Barendse*

*Cooperative Research Centre for Beef Genetic Technologies, CSIRO Livestock Industries, Queensland Bioscience Precinct, St. Lucia, QLD 4067, Australia. †The University of Queensland, School of Animal Studies, Gatton, QLD 4343, Australia

Summary

Ticks and tick-born diseases are major constraints on cattle production in tropical and subtropical regions in the world. Previously, we identified single nucleotide polymorphisms (SNPs) associated with tick resistance on bovine chromosome 3 at approximately 70 Mb. In this study, we genotyped a dairy \( n = 1133 \) and a beef \( n = 774 \) sample to confirm the association of the intronic SNP rs29019303 and its gene \( \text{ELTD1} \) with tick burden. We genotyped 18 additional SNPs in a region of 181 kb and found that rs29019303 was significantly \( P < 0.05 \) associated with tick burden in both samples with the same favourable allele. A second SNP in this same genomic region was also significantly associated with tick burden in each sample. The associations using haplotypes were stronger than for single markers, including a haplotype of nine tag SNPs that was highly significantly \( P = 0.0008 \) associated with tick counts in the dairy animals. This haplotype and two others were significant after Bonferroni correction for multiple testing. The estimated size of the effects was close to 0.9% of the residual variance in both samples tested.

Keywords Boophilus, Cattle, DNA, Genome-wide association, parasite, QLT.

Genetic variants associated with tick burdens have been previously identified in cattle. The first association was to the serum amylase locus on bovine chromosome (BTA) 3 (Ashton et al. 1968). Other studies showed that class I antigens of the bovine major histocompatibility complex (MHC), typed with a microlymphocytotoxicity test, were associated with tick load in cattle (Stear et al. 1989, 1990). Microsatellites in the Bovine lymphocyte antigens were also used in a candidate gene study and were shown to be associated with tick burden (Acosta-Rodrı́quez et al. 2005; Martinez et al. 2006; Untalan et al. 2007), confirming the location of genetic variation affecting the host resistance to ticks on bovine chromosome 23 (BTA23). More recently, microsatellites were used to scan all cattle chromosomes for quantitative trait loci for tick burden, and positive associations were found on 6 of them (Gasparin et al. 2007; Regitano et al. 2008; Machado et al. 2010). Finally, a genome-wide association study (GWAS) was performed using single nucleotide polymorphisms (SNPs), which found scores of SNPs significantly \( P < 0.001 \) associated with tick burden (Barendse 2007; Turner et al. 2010). BTA3 was found to contain a QTL for tick burden in the GWAS, albeit not at the serum amylase locus, but the microsatellite studies did not find evidence for linkage to BTA3.

The aim of this study was to confirm the associations of SNPs with tick phenotypes on BTA3 and to estimate the size of the effect in a large sample. We tested 19 SNPs from BTA3, including two SNPs from the GWAS (rs29019302 and rs29019303), the latter of which was highly significantly \( P < 0.001 \) associated with tick burden (Barendse 2007). These two SNPs are located in the second intron of the \( \text{ELTD1} \) (\( \text{EGF, latrophilin and seven transmembrane domain containing 1} \)) gene (Fig. 1). To confirm these associations and so verify the location of the QTL, we tested these two SNPs as well as an additional 17 SNPs from the Affymetrix 10K SNP array or the bovine assembly SNP database located within 181 kb around the \( \text{ELTD1} \) gene (Table S1).

Sample 1, consisting of a subset of the cattle of the taurine dairy tick experiment (DTE \( n = 1133 \)), had been collected in part of Queensland in northern Australia where ticks are endemic (tick zone) and were counted over two or more seasons (Turner et al. 2010). None of the animals in sample 1 had been used in the GWAS (Barendse 2007; Turner et al. 2010).
Sample 2, consisting of Brahman and Tropical Composite beef cattle [Cooperative Research Centre (CRC) \( n = 774 \)], also from the same tick zone, had ticks scored in one season (Prayaga et al. 2009). Tick counts represent actual counts of ticks in the size range 4.5–8.0 mm that were then subsequently log-transformed to approximate normality (Turner et al. 2010). Tick scores are a visual appraisal of the number of ticks of a size >4.5 mm into 6

Figure 1 The ELTD1 region at BTA3. (a) Single nucleotide polymorphism (SNP) and haplotype association analyses with tick phenotypes in the dairy tick experiment (DTE) and Cooperative Research Centre (CRC) samples. (b) and (c) Linkage disequilibrium (LD) \((r^2)\) between SNPs from DTE (b) and CRC (c) samples. (d) The ELTD1 gene location and structure aligned with the association and LD plots. (e) Expanded region of the BTA3 showing the location of closer coding regions.
grades: 0 is no ticks, 1 is ≤10 ticks, 2 is 11–30 ticks, 3 is 31–80 ticks, 4 is 81–150 ticks and 5 is >150 ticks, and the data did not require transformation to normality (Prayaga et al. 2009).

Tick burdens were analysed using mixed models with the ASReml software (Gilmour et al. 2002). The models took the general form of trait ~ mean + fixed effects + animal + error, where animal and error were fitted as random effects. For the DTE sample, the fixed effects of property, season and breed type were modelled as main fixed effects, where season was included the identity of the counter, where an animal was included if it had been measured twice, and where all available pedigree information was included in the model. For the CRC sample, the trait tick score was modelled with the fixed effects of breed, herd of origin, cohort, calving month and their first degree interactions. Three generations of pedigree information were available, and parentage was confirmed by DNA parentage testing. Neither of these models included SNP or haplotype as a covariate. Residual trait values for each animal extracted from these models were regressed on the number of copies of a SNP allele or haplotype.

Genotypes were obtained using SNPlex assays (Applied Biosystems) following the manufacturers’ instructions, and genotypes were scored by two individuals before merging with the residual phenotypes. Departures from Hardy–Weinberg equilibrium (HWE) were evaluated using PLINK (Purcell et al. 2007) within breed. Allele frequency differences were evaluated by comparing genotypic distributions between breeds (Weir 1996). The linkage disequilibrium (LD) between markers and the identification of Tag SNPs for the region was analysed using HAPLOVIEW 4.1 (Barrett et al. 2005). For association analyses, haplotypes were obtained using PHASE 2.1.1 (Stephens et al. 2001; Stephens & Scheet 2005) applying the algorithm five times for each set of SNP’s on animals with full genotypic data. We analysed haplotypes that included all nine of the Tag SNPs for the region (Tag haplotypes) and also haplotypes formed by a ‘sliding window’ of three adjacent SNPs, sliding one SNP at a time (3-locus haplotypes). The association between tick burden and SNP was evaluated through the regression of the tick residual phenotype on the number of copies of a particular allele or haplotype. In the haplotype analysis, each animal was scored for the number of copies of each haplotype that it possessed. Then the residual phenotype was regressed on the number of copies of each haplotype, one haplotype at a time. The significance threshold was adjusted for the number of haplotypes by Bonferroni correction of the significance threshold (Pritchard & Rosenberg 1999). The proportion of residual variance explained by a SNP was calculated from the square of the correlation between the phenotype residuals and the number of copies of each allele or haplotype.

Three of the 19 SNPs were monomorphic in all populations sampled, and a fourth was monomorphic in CRC and had a minor allele frequency <0.01 in the DTE sample. There were 128 comparisons of 16 SNPs by eight breed types, of which eight, or 6.3%, showed significant departures from HWE. One of the SNPs, rs43338299, was significantly out of HWE in most breed samples (P < 0.05). Fourteen of the 16 segregating SNPs had significant (P < 0.05) differences in allelic frequencies between the DTE, Brahman and the Tropical Composite sample (Table S1). The LD between SNPs in the region differs between samples (Fig. 1b,c). In the DTE sample, four haplotype blocks were identified, showing high r² between SNPs; the longest one, comprising seven SNPs, spans 64 kb. The same haplotype blocks were not found in the CRC sample, even though four haplotype blocks were also identified for this sample.

Three SNPs and several haplotypes including these SNPs were significantly associated with tick burden. The SNP rs29019303, which had been significantly (P < 0.001) associated with tick burden in the GWAS (Barendse 2007), was significantly (P < 0.05) associated with tick burden in both confirmation samples with the same favourable allele (Table 1 and Fig. 1a). In addition, two other SNPs were also significantly associated (P < 0.05) with tick burden: rs43338299 in the DTE and rs43338300 in the CRC sample. These last two markers are 1.7 kb apart and showed high LD between them in both samples (DTE r² = 0.74 and CRC r² = 0.92), being part of the same haplotype block. There were two significant (P < 0.01) Tag haplotypes, one in each sample (Table 1 and Fig. 1a). The Tag haplotype ‘h10’ was highly significantly (P = 0.0008) associated with tick burden in the DTE sample and this was still significant after Bonferroni correction of the significance threshold. The Tag haplotype ‘h1’ was highly significantly (P = 0.0085) associated with tick burden in the CRC sample, but this was not significant after Bonferroni correction of the significance threshold. The ‘h1’ haplotype shares seven alleles with the ‘h10’ haplotype. There were two 3-locus haplotypes (SNP-14,15,16-h1 and SNP-15,16,17-h1) significantly (P < 0.05) associated with tick burden in the DTE experiment after Bonferroni correction. SNP 15 and SNP 16 in these haplotypes are rs43338299 and rs43338300, respectively, and these are individually significantly associated with tick burden. These two haplotypes as well as a third 3-locus haplotype were significantly (P < 0.05) associated with tick burden in the CRC sample, although none of these were significant after Bonferroni correction in that sample (Table 1).

The QTL accounts for between 0.5% and 1.0% of the residual variance estimated in two relatively large samples. Using single markers, the estimate R² ranged between 0.5% and 0.7% and using haplotypes R² ranged between 0.6% and 1.0%. This range of effect size is similar to the estimates found by our group in another association study of genetic variation in the integrin alpha 11 (ITGA11) gene.
with tick burden (Porto Neto et al. 2010). Single nucleotide polymorphism effects in this size range mean that further confirmatory studies will require large numbers of individuals to be able to detect the effect, and these SNPs will need to form part of a panel of DNA markers for predicting tick burden. These effects are smaller than those found in Machado et al. (2010). This could be because of the method of estimation of the proportion of the variance explained. Nevertheless, QTL linkage studies will not detect all possible QTL, and these have to be above a certain size to be detected (e.g. Barendse et al. 2008; Davis et al. 2008).

Table 1 Markers and haplotypes most significantly associated with tick burden.

<table>
<thead>
<tr>
<th>SNP/Hap</th>
<th>Code</th>
<th>Sample</th>
<th>n0</th>
<th>n1</th>
<th>n2</th>
<th>R^2</th>
<th>s^4</th>
<th>SE^5</th>
<th>P^6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single marker</td>
<td>rs29019303</td>
<td>DTE</td>
<td>167</td>
<td>382</td>
<td>240</td>
<td>0.005</td>
<td>-0.099</td>
<td>0.050</td>
<td>0.0480</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRC</td>
<td>45</td>
<td>283</td>
<td>346</td>
<td>0.007</td>
<td>-0.135</td>
<td>0.064</td>
<td>0.0349</td>
</tr>
<tr>
<td>rs43338299</td>
<td></td>
<td>DTE</td>
<td>596</td>
<td>70</td>
<td>111</td>
<td>0.006</td>
<td>-0.111</td>
<td>0.050</td>
<td>0.0273</td>
</tr>
<tr>
<td>rs43338300</td>
<td></td>
<td>CRC</td>
<td>5</td>
<td>60</td>
<td>573</td>
<td>0.007</td>
<td>-0.251</td>
<td>0.116</td>
<td>0.0312</td>
</tr>
<tr>
<td>Tag haplotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tag h1</td>
<td></td>
<td>CRC (20)</td>
<td>703</td>
<td>69</td>
<td>1</td>
<td>0.009</td>
<td>0.322</td>
<td>0.122</td>
<td>0.0085</td>
</tr>
<tr>
<td>Tag h10</td>
<td></td>
<td>DTE (27)</td>
<td>1006</td>
<td>121</td>
<td>5</td>
<td>0.010</td>
<td>0.298</td>
<td>0.089</td>
<td>0.0085*</td>
</tr>
<tr>
<td>3-locus haplotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13,14,15-h1</td>
<td>110 - -</td>
<td>CRC (4)</td>
<td>596</td>
<td>61</td>
<td>4</td>
<td>0.008</td>
<td>0.281</td>
<td>0.119</td>
<td>0.0185</td>
</tr>
</tbody>
</table>
| 14,15,16-h1  | - 101 -    | DTE (4) | 633| 188| 72 | 0.008| -0.144| 0.053| 0.0067*
| 15,16,17-h1  | - 001      | CRC (5) | 662| 78 | 12 | 0.007| 0.213| 0.094| 0.0240|
|              |            | CRC (4) | 661| 75 | 15 | 0.006| 0.185| 0.091| 0.0427|

DTE, dairy tick experiment; CRC, Cooperative Research Centre and number of haplotypes reconstructed.

*Significant after Bonferroni correction.

1Single nucleotide polymorphisms (SNP) or haplotype id. The haplotype id is the reference number for the SNP included in the haplotype and haplotype number. Tag haplotypes include rs43024753, rs29019302, rs29019303, rs43338248, rs43338255, rs43338319, rs43338318, rs43338300, rs43338299. 13 – rs43338319, 14 – rs43338318, 15 – rs43338300, 16 – rs43338299, 17 – rs43338784.

2Single marker analyses: n0, number of animals homozygote for the allele closer to A in the alphabet, n1, number of heterozygote animals, n2, number of animals homozygote for the alternative allele. Haplotype analyses: n0, n1, n2, number of animals with zero, one or two copies of the haplotype.

3Proportion of residual variance explained.

4Haplotype substitution effect in phenotypic standard deviations.

5Standard error.

6Uncorrected P-values.

Approximately 200 kb distal to ELTD1 is a predicted coding sequence that shows homology to the IFI44 gene (similar to interferon-induced protein 44), a component of the adaptive immune system (Fig. 1e). However, this predicted coding sequence has not yet been shown to be transcribed in cattle, and because of its similarity to an additional, adjacent predicted coding sequence, IFI44L, several SNPs in the region have complex patterns of inheritance. Some of these, like INRA-143, have not even been assigned to the Btau4.0 genome assembly (Matukumalli et al. 2009). However, IFI44 in humans appears to be responsive to viral infection (Lin et al. 2007; Bochkov et al. 2010), so it is not clear whether future research should focus only on that gene. Further study of this region to identify causative mutations will require the analysis of several of the genes, including a determination that the association with ELTD1 is either causal or merely because of LD to the QTL.

Acknowledgements

We thank A. Reverter and R. J. Hawken for comments on the manuscript. LRPN was supported by an Endeavour International Postgraduate Research Scholarship, a University of Queensland International Student Living...
Allowance and a Cooperative Research Centre for Beef Genetic Technologies (Beef CRC) scholarship. This work was co-funded through the Beef CRC and CSIRO-LI (WB).

References


**Supporting information**

Additional supporting information may be found in the online version of this article.

*Table S1* Single nucleotide polymorphism marker descriptions.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.