

Promoter sequence analysis of differentially expressed genes in sheep following a nematode parasite challenge

Shivashankar Hiriyur Nagaraj^{1,2}
Shivashankar.Hiriyur-Nagaraj@csiro.au

Aaron Ingham¹
Aaron.Ingham@csiro.au

Antonio Reverter¹
Tony.Reverter-Gomez@csiro.au

- 1 CSIRO Livestock Industries, Queensland Bioscience Precinct, 306 Carmody Road, St. Lucia, Qld 4067, Australia.
- 2 Corresponding author

Abstract

We have previously reported a list of genes that were differentially expressed (DE) in gastrointestinal tissues collected from genetically resistant or susceptible sheep following challenge with parasitic nematodes. In a bid to define genes that may be co-regulated for either the resistant or susceptible lines the DE genes were clustered into groups that show similar expression patterns at various sample collection times. Due to the lack of suitable ovine sequence, the upstream region of orthologous bovine genes within each cluster was then searched for the presence of conserved sequence motifs. As a result, we found several overrepresented sequence motifs that show a high degree of sequence similarity to well defined transcription factor binding sites and as such are potential regulatory elements. Furthermore, three random datasets derived from bovine genes were subjected to similar analysis to test whether the overrepresentation is attributable to chance given the bovine genome composition. Motifs in randomly selected sequences and DE gene clusters were different from each other suggesting the identification of putative regulatory elements. This study serves as a starting point for understanding gene regulation in sheep following nematode challenge and may help to define the genetic basis of varying ability to resist infection with gastrointestinal nematodes.

Keywords: gene regulation, cis-regulatory elements, transcription factors, parasitic nematodes, differential expression.

1 Introduction

Gene regulation at the level of transcription is critical for cellular machinery and noncoding DNA sequences play an essential role in the spatiotemporal regulation of gene transcription. Transcription factor (TF) proteins communicate with their target genes by recognition of very short DNA signals, called the *cis*-regulatory elements.

We recently studied gene expression for a candidate set of genes in sheep following a nematode challenge to understand the genetic differences between resistant and susceptible individuals. Several DE genes were identified in resistant and susceptible animals including members from Toll-like receptors (TLR), free radical producing genes and NF- κ B signalling pathway [1]. In order to gain additional insight into gene regulation, we investigated promoter regions for gene clusters to explore consensus motifs in upstream regions common to DE genes. We used the *Bos taurus* genome as a reference genome, as upstream regions were not available for all of the genes in sheep. However, regulatory elements are evolutionarily important and thus conserved across species [2]. Here, we report the results from promoter sequence analysis for DE genes classified into four clusters including TLR, XDH, Cathepsin and DUOX; as well as the independent analysis for each DE gene.

2 Method and Results

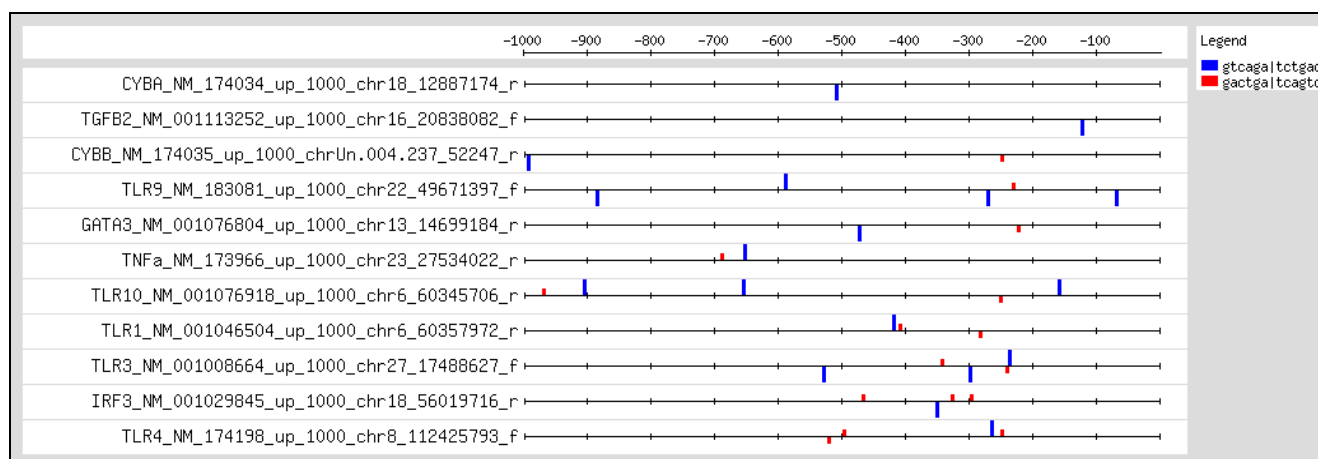
For all of the genes selected for promoter analysis, a region of 2000 bp upstream of the transcription start site for each of the corresponding bovine RefSeq genes was retrieved from University of California-Santa Cruz genome browser (<http://genome.ucsc.edu>). Similarly, three random datasets (20 genes each) containing upstream regions were generated. Promoter sequence analysis was conducted using RSAT [3] and Eukaryotic Promoter Database [4] containing entries for eukaryotic POL II promoters, for which the transcription start site has been determined experimentally.

Several overrepresented oligonucleotides were found in DE gene clusters and Table 1 shows the motifs with significant scores. Similarly, overrepresented oligonucleotides were generated from random sequence data (data not shown) and compared to motifs from the four gene clusters. In addition, feature maps were drawn for all the datasets to visually inspect the overrepresented motifs. Figure 1 shows one such feature map for the TLR cluster.

Table 1: Overrepresented nucleotides from the DE gene clusters (best consensus are shown)

Gene Cluster	Number of genes per cluster	Sequence	observed occurrences	expected occurrences	P-value
TLR Cluster	25	gtcaga	25	9.81	3.5e-05
XDH cluster	10	cggcgc	16	3.88	3.3e-06
XDH cluster	10	ccgcgc	18	5.29	1.1e-05
Cathepsin cluster	9	cccccg	17	4.42	4.2e-06

Figure 1: Feature map of TLR cluster showing overrepresented oligonucleotides in TLR cluster



Potential TF binding sites in upstream regions of genes were predicted for DE genes (resistant and susceptible) using a library of mononucleotide weight matrices derived from experimentally determined binding sites. We found several putative TF binding sites (Table 2).

Table 2: Selected hits from matrix search for transcription factor binding sites

Query Gene	Position (strand)	Sequence	Transcription Factor Name
Resistant: TLR2	150 (-)	aCACTTga	Nkx2-5 (homeobox gene)
Resistant: DUOX1	948 (-)	acaAACAAac	FOXD3(fork head)
Susceptible: IKBKB	686 (-)	GGAAAttccc	NF-kappaB
Susceptible: TGFB2	121 (+)	gatAACGGtc	v-Myb

3 Discussion

The observed occurrences were 3-4 times the expected occurrences of motifs in DE gene clusters and consensus motifs in DE gene clusters were discovered. For instance, the important TLR pathways members TLR1, TLR3, TLR9, TLR10, GATA3 and IRF3 share a common motif “gtcaga”, suggesting the involvement of one or more global regulators following sheep infection. Further research will include (1) comparisons to regulatory regions from other mammalian genomes; and (2) experimental validation using ChIP-chip and ChIP-Sequencing experiments.

Acknowledgments

The methodology and results presented here were generated using results of research supported by SheepGenomics, which is an initiative of Australian Wool Innovation Limited and Meat & Livestock Australia Limited.

References

- [1] Ingham A, Reverter A, Windon R, Hunt P, Menzies M. *Int J Parasitol.* 2008 Mar;38(3-4):431-42.
- [2] Ureta-Vidal A, Ettwiller L, Birney E. *Nat Rev Genet.* 2003 Apr; 4(4):251-62.
- [3] Thomas-Chollier M, Sand O, Turatsinze JV, van Helden J et. al. *Nucleic Acids Res.* 2008 Jul 1; 36:W119-W127.
- [4] Schmid CD, Perier R, Praz V, Bucher P. *Nucleic Acids Res.* 2006 Jan 1; 34:D82-85.