Proc. Assoc. Advmt. Anim. Breed. Genet. 18:283-286

ON THE EXPRESSION PROFILE OF CANDIDATE GENES CONFERRING RESISTANCE TO GASTRO-INTESTINAL NEMATODES IN SHEEP

Shivashankar H Nagaraj [1], Antonio Reverter [1], Moira Menzies [1], Nick Andronicos[2], and Aaron Ingham [1]

[1]CSIRO Livestock Industries, 306 Carmody Road, St. Lucia, Qld 4067, Australia
[2]CSIRO Livestock Industries, Armidale, NSW 2350, Australia

SUMMARY

We report on the integration of a number of RT-PCR expression datasets totalling 102 genes across 49 experimental conditions of relevance to gastro-intestinal nematodes in sheep. The experimental conditions include three challenge trials, six tissues, five genotypes and two parasites surveyed at multiple time points. After mining gene interaction data for these 102 genes, we have generated an interaction network that enabled us to study these genes in a biologically relevant, system context. Following a systematic investigation of this network, we have identified a number of genes encoding extracellular matrix proteins that may be useful biomarkers enabling identification of parasite resistant sheep.

INTRODUCTION

Selective breeding of sheep that are genetically resistant to nematode infection may be used to overcome problems associated with anthelmintic resistance (Waller, 2006). This strategy would be greatly enhanced if accurate tests were available for attributing resistance. Differential expression of genes or their encoded protein products between resistant (RES) and susceptible (SUS) animals provides one option for such tests. For this reason, we have generated gene expression profiles in RES and SUS animals before and after nematode challenge, demonstrating the impact of time, parasite species, multiplicity of challenge and tissue.

In this study, we combine expression data from various studies from our laboratory and apply data mining techniques to publicly available human data to generate a gene network that may be used to determine how the performance of one gene might be informed by others in the network and in so doing determine an optimal selection of genes that are predictive of phenotype. We also intend to determine the location of our key candidates within a previously determined protein interaction network. We overlay a series of gene attributes into the network, including regulatory and extracellular component, allowing us to efficiently identify targets that are likely to be measureable in the blood, thereby making good candidates for biomarker assays.

MATERIALS AND METHODS

We use the gene expression data from the set of 76 candidate genes related to nematode resistance in sheep reported by Ingham *et al.* (2007). The set comprised of genes differentially expressed (DE) between RES and SUS animals. To these data, an additional set of 26 genes was incorporated from studies in further trials with more biological replicates and tissues. In total, 102 genes across 49 experimental conditions were represented in the combined data set including 2 parasites, 3 trials, 5 flocks, 6 tissues, and 16 time points. The experimental layout of the resulting data set was that of an incomplete block design with only two genes (GAPDH and RPL) represented across the 49 conditions.

The entire data set was normalized by fitting a mixed ANOVA model with threshold cycle in the PCR reaction as the measure of expression abundance as the dependent variable. Main design effects were treated as fixed effects, while gene and gene by condition interaction were treated as

random effects. PermutMatrix (http://www.lirmm.fr/~caraux/PermutMatrix/EN/index.html) was used to perform cluster analysis of the normalized expression data across rows (genes) and columns (conditions).

Finally, we downloaded a set of 55,606 true positive interactions among 7,197 human genes that were defined from functional studies (Franke *et al.*, 2006). This interactions dataset was built including 2,788 confirmed, direct, physical protein-protein interactions derived from the Biomolecular Interaction Network Database (BIND; http://bind.ca), 18,176 confirmed human protein interactions from the Human Protein Reference Database (HPRD; http://www.hprd.org/), 22,012 direct functional interactions from the Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg), and 16,295 interactions derived from Reactome (http://www.reactome.org).

RESULTS AND DISCUSSION

Figure 1 shows the result of the hierarchical cluster analysis of the expression profile of 102 genes across the 49 conditions. Expression results for 76 genes have been reported previously by Ingham *et al.* (2007). To this dataset we have added 6 additional genes studied in the same trial but not reported along with a further 20 genes analysed in third independent experiment. See supplementary data for details. To avoid biases in the generation of clusters, the average gene expression was imputed in the cells corresponding to genes not surveyed in a given condition. These are represented as black cells in Figure 1. The cluster of columns, corresponding to experimental conditions, reveals the tissues as being the biggest hierarchy in the clusters, followed by parasite and time point. The cluster of rows, corresponding to genes, reveals groups of genes enriched for specific functions. Across the rows, three main clusters are clearly distinguishable including toll-like receptors, cytokines, and proteases.

The mining of the interaction dataset resulted in a network with 703 nodes (genes) connected by 1,090 edges (interactions) that was visualized using Cytoscape (http://www.cytoscape.org). The entire view of the network is given in Figure 2. Interestingly, some of the clusters observed in Figure 1 and derived from gene expression data, retained their integrity in the network. For instance, TLR pathway members form a cluster in Figure 1 and a sub-network in Figure 2. The network file in Cytoscape format used to generate Figure 2 is available for download from our public website (http://www.livestockgenomics.csiro.au/courses/Shiv_AAABG09.html).

One limitation of using gene expression data as a biomarker is that the expression profile of the gene and subsequent location of its encoded product are often tissue specific. In this case, DE genes distinguishing RES and SUS animals in gut tissues might only be detectable in these tissues. Sampling gut tissue is highly invasive and therefore not appropriate for assaying in a routine manner. Instead, extracellular products are easily sampled in the blood. For this reason we performed a systematic investigation of this network and identified a number of genes encoding extracellular matrix protein candidates, associated with our DE candidates. The DE genes KCNJ15 and DYRK3 both distinguish RES and SUS animals in the gut, but these factors are intracellular. Through cluster analysis these genes were shown to interact with the extracellular factors IL16 and NT5E respectively, making these candidates potentially useful targets for biomarker development. Taken together, the results illustrate the benefits of integrating gene expression data, together with interaction networks to study genes involved in complex biological processes/signalling pathways associated with parasite resistance in sheep.

FUTURE DIRECTIONS AND CONCLUSIONS

Although gene expression data offers a list of DE genes, and interaction networks supply a list of dual connections, the nature of the information that can be drawn from either approach is relatively

limited. Transcript abundance alone reveals little about the mechanisms underlying the observed changes or details of different simultaneously occurring events.

In order to understand the genetic basis for ovine resistance to diseases, we need to gain insights into the dynamic regulatory nature of the immune response. Here we have shown how a network systems approach can be used to successfully inform the selection of potential biomarkers. Moving forward, we plan to capture the dynamism inherent in immune responses by

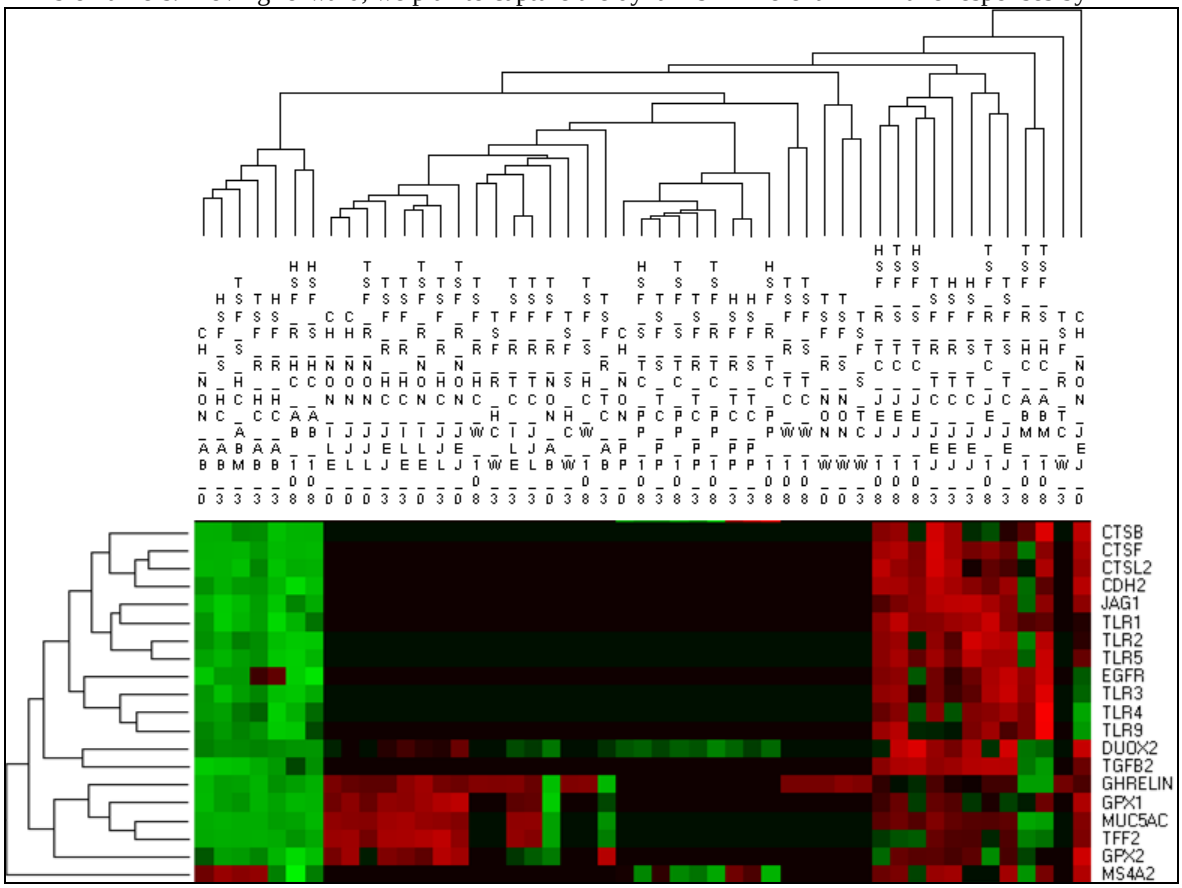

Figure 1. Hierarchical cluster analysis of 102 genes (rows) by 49 experimental conditions (columns) derived from parasite resistance studies with sheep.

A snapshot of representative cluster comprising TLR pathway genes has been shown; complete version of figure 1 containing all the clusters is available for download from our public website (http://www.livestockgenomics.csiro.au/courses/Shiv_AAABG09.html). The spectrum goes from green to red for low and high expression, respectively. Abbreviations used in column headings are as follows; Animals (TSF Trichostrongylus Selection Flock; HSF, Haemonchus Selection Flock; Chiswick, Outbred commercial sheep); Genetic Resistance (R, Resistant; S, Susceptible); Infection status (HC, Challenged with *Haemonchus contortus*; TC, Challenged with *Trichostrongylus colubriformis*; NONE, Sheep not challenged); Tissue sampled (ABOM, abomasum; ILE, ileum; JEJ, jejunum, WBC, white blood cells; JEJILE, junction of jejunum and ileum; PP, Peyer's Patch); Time of sample collection relevant to challenge (0, 3, or 108 days).

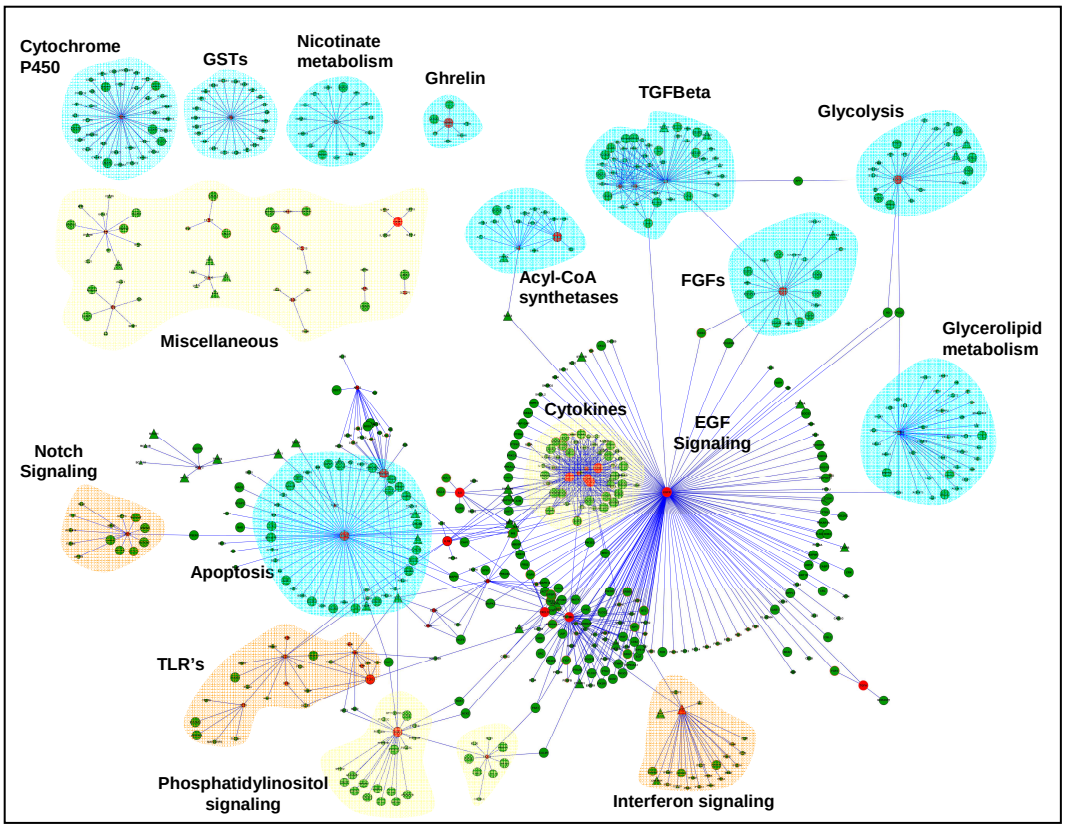

Figure 2: Network with 703 nodes (genes) connected by 1,090 edges (interactions). Visualised attributes included: genes surveyed in the RT-PCR assays were coloured red (otherwise green); transcription factors as triangles (otherwise, circles); big symbols were assigned to post-translational modifiers; red borders for secreted proteins; and sub-networks are highlighted as clouds.

studying gene co-expression networks overlaid with transcription regulation associated with regulatory events. These studies are crucial, given the current, serious resistance problems in parasites against most treatments, and residue problems in meat, milk and the environment.

ACKNOWLEDGEMENTS

SHN is grateful to CSIRO for the award of an OCE Post-Doctoral Fellowship. The financial support of the CSIRO Transformational Biology Initiative is gratefully acknowledged. The authors are grateful to SheepGenomics for financial support.

REFERENCES

Franke, L., van Bakel, H., Fokkens, L., de Jong, E.D., Egmont-Petersen, M. and Wijmenga, C. (2006) *Am. J. Hum. Genet.* **78**:1011.

Ingham, A., Reverter, A., Windon, R., Hunt, P. and Menzies M. (2007) *Int J Parasitol*. **38**:431

Waller, P.J. (2006) *Vet. Parasitol.* **139**:1