Network And Pathway-Based Analysis Of Candidate Chromosomal Regions For *Mycobacterium Avium Ssp. Paratuberculosis* Antibody Response In Cattle

*G.Minozzi**, F. Strozzi*, L. Buggiotti^, A. Stella†, M. Luini‡ and J.L. Williams*

Introduction

Mycobacterium avium subspecies paratuberculosis (MAP) causes paratuberculosis (ParaTB) or Johne's disease in cattle, a chronic granulomatous gastroenteritis characterized by diarrhoea, weight loss, drop in milk production and ultimately death (Olsen I et al. (2002)). Johne's disease is not treatable and vaccine efficacy is still controversial. There is evidence for host genetic component to susceptibility to ParaTB, with estimated heritability between 0.06 and 0.1 (Koets et al. (2000), Mortensen et al. (2004), Gonda et al. (2006), Hinger et al. (2008)). The aim of this work was to perform a case-control study using a high density SNP panel (Illumina BovineSNP50 BeadChip) to localize genes having an impact on Johne's disease susceptibility and to unravel the gene networks underlining the mechanisms of susceptibility to Johne's disease.

Network and pathway oriented analysis is a significant advance beyond the single gene function analysis, and involves assessment of the joint impact of several genes acting together at the larger functional level. In this work the role of the genes that are close to the significant SNPs found by single-SNP analysis was investigated, but in addition an analysis was performed to identify sets of genes that act together and are involved in the pathology of this complex disease. Finally, the overall aim was not only to investigate the role of the genes that are close to the significant SNPs found by single-SNP analysis, but also to identify a bigger subset of genes that together are involved in the pathology of the complex disease.

Material and methods

Animals. Samples were collected from routine Johnes disease screening of Holstein cattle carried out between 2007 and 2008 in Italy, in herds with high occurrence of Johne's diseases. Samples were classified based on the serum antibodies produced in response to MAP infection using the ID-screen® test (Id.Vet Montpellier, France). In total 966

* PTP, Lodi, Italy
† IBBA, CNR, Lodi, Italy
^*present address Università della Tuscia, Viterbo, Italy*
‡ IZSLER , Lodi, Italy

individuals were chosen from 119 farms. Of these samples 483 were MAP antibody positive (cases) and 483 MAP antibody negative (controls): case-control samples were age, herd, Elisa test-day matched.

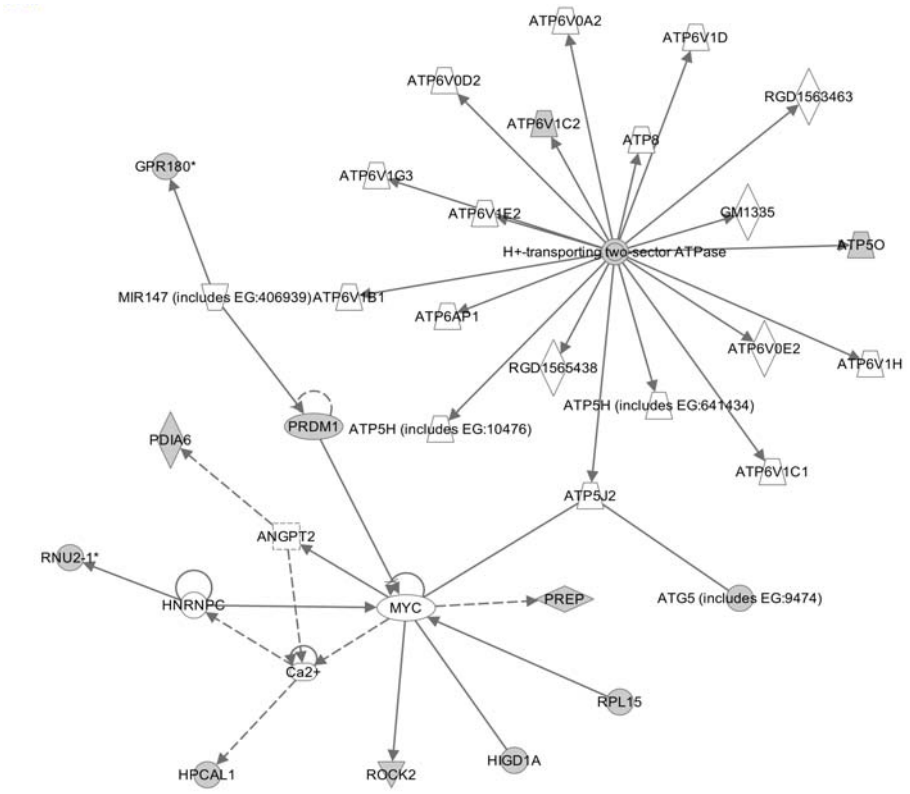
Association analyses. Genome-wide association analysis was performed using the GenABEL package in R *(*Aulchenko et al. (2007)) with the GRAMMAR-CG approach (Amin N et al (2007), Aulchenko et al. (2007)).

Gene analysis in candidate regions and network - pathway analysis. Genes within 1 Mb upstream and downstream from the 6 SNP most significantly associated with disease were retrieved from the Ensembl database. Location and gene annotation were based on the Btau_4.0 genome assembly and on Ensembl release 56 (http://www.ensembl.org).
The canonical transcript for each gene was considered for the analysis. A bioinformatic pipeline, was created to process all the cow genes present in the candidate regions. Blast was used to perform a homology search against the Human RefSeq and the Uniprot databases to find a possible annotation for novel cow genes and to retrieve homologous human transcript IDs, suitable for the pathways and network analysis.
Data were analyzed through the use of Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com), using the human RefSeq ID retrieved by the homology search on the cow genes. Each gene identifier was then mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base. Networks of these genes were then algorithmically generated by the software based on their connectivity.

Results and discussion

The genome wide association analysis identified several chromosomal regions associated with disease status: a region on chromosome 12 with high significance ($P < 5x10^{-6}$), while further regions on chromosome 9, 11, and 12 were identified with moderate significance ($P < 5x10^{-5}$). The association on Chromosome 9 was also seen in a genome-wide association study for MAP, using a different phenotype: presence of bacteria in cells of the ileum (Settles et al. (2009)), and the significant SNP on chromosome 12 fell within a QTL for Johnes disease susceptibility (Gonda et al. (2007)), providing additional evidence that genes within these regions are involved in response to mycobacterium infections.
The annotation and pathway analysis identified 93 genes within the candidate regions, for 82 of these genes an annotation was retrieved using the bioinformatics pipeline on the Human Refseq database. Interestingly, 17 novel unannotated cow genes in the chromosome 12 candidate region, showed high similarity with the same Human RefSeq, the ATP-binding cassette sub-family C (CFTR/MRP) member 4, with different levels of sequence identity. These are most likely genes belonging to the same gene family.
The pathway-based analysis showed that the 18 out of the 82 annotated genes were involved in 50 different pathways from chemokine signaling, to leukocyte extravasation signaling and even to urea cycle. Although the precise mechanism of response to disease is unknown, it may involve maturation of the immune system in terms of balance between T-cell subsets and the specific tissue distribution of immune cells. Furthermore several physiological changes are involved during the progression of the disease. These changes may be affected by a broad spectrum of pathways including those highlighted by this analysis.

Figure 1. Network involved in Molecular Transport, Energy production and nucleic acid metabolism, genes or gene products are represented as nodes and the biological relationship between two nodes is represented as an edge



However, although the genes identified belong to several different pathways, these can be described in only few networks. The most relevant network involves molecular transport, energy production and nucleic acid metabolism (Figure 1). The entire network includes 25 genes, 12 of which are present in the candidate regions and include the PRMD1 gene which is a transcription repressor that acts on the beta interferon gene expression and affects maturation of B-lymphocytes to antibody secreting cells.

The PRDM1 gene has been shown to play major roles in regulating the differentiation of B and T lymphocytes in human (Martins et al. (2006)). Moreover, PRDM1 interacts with multiple chromatin-modifying enzymes to induce transcriptional repression at the IFN-beta promoter (Ren et al. (1999)). This gene represents a strong candidate to be investigated for Johne's disease susceptibility.

Conclusion

The results show that the pathway based analysis of regions with significant markers identified from GWAS may identify biologically meaningful pathways associated with, in this case, the response to disease. Study of these pathways may help to identify strong candidate genes that may be tested further eg. in functional studies. This information may be then useful in developing diagnostic tests or therapeutic approach to control the disease.

Acknowledgments The authors thank Yuri Aulchenko, Dan Bradley, Emma Finley, and Donagh Berry for their valuable contributions and suggestions. Rossana Capoferri, Caterina Cambulli and Simona Panelli are acknowledged for managing the samples. Funding for GM and LB was provided by the EC MacroSys project ref number FP7-KBBE-2007-1-1-2-211602, JLW and FS acknowledge the Cariplo Foundation for research support.

References

Amin N., van Duijn C.M., Aulchenko Y.S. (2007) *PLoS ONE*., 5;2:e1274.

Aulchenko Y.S., Ripke S., Isaacs A., et al. 2007 *Bioinformatics.,* 23:1294-6.

Aulchenko YS, de Koning DJ, Haley C. (2007) *Genetics.* 177: 577-85.

Gonda M.G., Chang Y.M., Shook G.E.,et al. (2006). *J. Dairy Sci.* 89: 1804–12.

Gonda M.G., Kirkpatrick B.W., Shook G.E. et al. (2007). *Anim. Genet*. 38: 389–96.

Hinger M., Brandt H., Erhardt G. (2008). *J. Dairy Sci*., 91: 3237–44.

Ingenuity® Systems, www.ingenuity.com

Koets A.P., Adugna G., Janss L.L.G.,et al.(2000). *J. of Dairy Sci*., 83: 2702–8.

Martins G.A., Cimmino L., Shapiro-Shelef M., et al. (2006)., *Nat. Immunol*.7:457–465.

Mortensen H., Nielsen S.S., and Berg P. (2004). *J. Dairy Sci*., 87:2108–13.

Olsen I., Sigurğardóttir G., Djønne B. (2002). *Vet Q,*. 24:12-28.

Ren, B., Chee K.J., Kim T.H. et al. (1999). *Genes Dev*., 13:125–137.

Settles M., Zanella R., McKay S.D., et al. (2009). *Anim Genet*., 40:655-62.