**Identifying candidate genes for the regulation of the response to *Trypanosoma congolense* infection**

**Introduction**

African cattle breeds differ significantly in their ability to survive low to moderate levels of challenge with *Trypanosoma congolense*. Similarly the survival times of inbred mouse strains vary substantially after infection. We have previously identified three regions of the mouse genome that regulate survival after infection in two crosses (A/J and Balb/c × C57BL/6) (Kemp et al 1997 Nature Genetics). These have been designated Tir1, Tir2 and Tir3 for Trypanosoma infection response loci on mouse chromosomes 17, 5 and 1 respectively. We have now used two strategies to reduce the size of these regions using microsatellite markers. Initial analysis of these regions identified candidate genes that show a correlation with survival time at the QTL loci. This was carried out by genotyping Introgressed congenic mice on C57BL/6 and A/J background. The introgressed regions were defined using sequences from the C57BL/6 and A/J alleles. We have now used a combination of marker genotyping and mapping strategies to identify candidate genes that may regulate the response to *Trypanosoma congolense*. The candidate genes identified have been compared to those previously identified and their expression profiles examined.

**Congenic Mice**

Congenic lines are created by crossing resistant C57BL/6 with the susceptible A/J mice. At each generation the offspring are genotyped to identify those animals that are carrying alleles from C57BL/6 in the target region of interest and these are then selected to be bred into the recipient genome. After seven generations of backcrossing to A/J, heterogeneous carriers of the C57BL/6 donor region of interest were intercrossed and homozygous carriers were selected as the congeneric line and designated TimCC in this study. Homozygous carriers of the A/J alleles were selected as the control line and designated TimAA. The creation of these mice lines makes it possible to study the effect of each locus in isolation from the other loci and their background.

**Mapping loci controlling the response to infection in 129J mice**

Identifying the regions controlling the response to infection in additional mouse strains make it possible to refine the list of candidate genes that might regulate survival time. F2 C57BL/6 × 129J were bred by crossing 129J and C57BL/6 mice to create and F1 generation and then by intercrossing their offspring to create an F2. 135 F2 C57BL/6 × 129J mice were genotyped with the Illumina 364 SNP mapping panel and the data was analysed with the IQLX package (Figure 3). The data confirmed the presence of QTL on chromosomes 1, 5 and 17 but there was no evidence of a QTL on chromosome 5 suggesting that 129J might carry the C57BL/6 allele at this locus. The locus on chromosome 17 was significantly associated with survival. The locus on distal chromosome 1 was not significant after correction for multiple testing, but since it was supported by multiple SNP markers and it coincided with previously identified loci on A/J and Balb/c mice, it may be real but the effect size of the allele may be insufficient to show a significant effect in this relatively small panel of mice.

**Haplotype analysis**

The availability of mapping data from three pairs of mouse strains makes it possible to look for correlation between haplotypes across the QTL regions and response to infection. The genome of inbred mouse strains is believed to be composed of mosaic of regions derived from three or four ancestral strains. Given that we have observed the same QTL in multiple pairs of mouse strains it is likely that the polymorphisms that make C57BL/6 more resistant are derived from one of these ancestral strains. If the ancestral strain from which a gene is inherited can be identified then it should be possible to predict if an animal is carrying a C57BL/6 allele that may influence survival at this locus. We have used introgression mapping to identify those animals that are carrying alleles from C57BL/6 in the target region of interest and these mice are then selected to be bred into the recipient genome. After seven generations of backcrossing to A/J, heterogeneous carriers of the C57BL/6 donor region of interest were intercrossed and homozygous carriers were selected as the congeneric line and designated TimCC in this study. Homozygous carriers of the A/J alleles were selected as the control line and designated TimAA. The creation of these mice lines makes it possible to study the effect of each locus in isolation from the other loci and their background.

**Effect of gene copy number on expression**

Single nucleotide polymorphisms that are accounted for by genetic variation. Regions of the mouse genome can be duplicated or deleted as is already known as copy number variants (CNV). CNV between C57BL/6 and A/J mice were detected using Agilent 244k whole genome CNH arrays. An amplification of the Gli1 gene within the Tim1 QTL on chromosome 17. This gene was not identified by the haplotype analysis but expression analysis indicated that the gene was more highly expressed in A/J mice than the C57BL/6. **Conclusions**

The combination of congeneric mice, additional mapping data and discovery of copy number variations within the QTL regions has made it possible to identify a short list of genes that might regulate the response to infection with *T. congolense*. This list is now sufficiently short to permit us to undertake detailed studies on the role of individual genes in the response to infection and hence determine whether they cause the difference in survival time after infection. The identification of these genes is expected to give an insight into the pathways that regulate the response to infection and may lead to new approaches to treatment.

**Acknowledgements**

We thank John Watamura, Moses Ogugo and Leonie Wardsworth for technical assistance. These studies were funded by the Wellcome Trust.