Bioinformatics assignment. Professor M. J. Stear 12th February 2015

The assignment is to determine the association between IgE responses and individual SNP in the DRB1 locus.

In this assignment, you are going to explore the association between polymorphism in the ovine MHC and the IgE response to nematode infection. Nematodes are a serious problem in sheep husbandry costing British farmers about £100,000,000 per year. IgE plays a key role in regulating the establishment and survival of nematodes. These are data from real sheep and will need some ‘cleaning’ before analysis. All sheep are of the Texel breed and come from the same farm.

I have placed five files on moodle. The texelhaplotypes file contains the sheep identity, and the names of two MHC haplotypes. The Texel DRB1 haplotypes files relates the haplotype number to the DRB1 sequence of exon2 The TexeligE.txt file contains the measured activity of IgE against the third larval stage (L3) of the nematode *Teladorsagia circumcincta* and a Box-Cox transformation. The IgE response against third stage larvae is the most important one. The IgE response does not follow a normal distribution and a suitable (Box-Cox) transformation is the fourth root (i.e. raise IgE to the power 0.25). The fourth file is ‘Additional Sheep Information’ . Typically IgE responses are influenced by sex, year and date of birth as well as background genes inherited from the father and mother (sire and dam). The fifth file (Texel DRB1 haplotypes.txt) contains the accession numbers that allow you to access the DNA and protein sequences.

You are free to use any programmes and procedures you wish. I suggest that you first create a database of SNP for each of the alleles. Then determine the gene frequencies and linkage disequilibrium of these SNP. Then explore the association between SNP and IgE response in a mixed model fitting year, sex and date of birth as fixed effects with sire and dam as random effects.