

Biomathematics and Statistics Scotland

Rob Kempton & Jim McNicol

Biomathematics & Statistics Scotland (BioSS) is devoted to the application of statistics and mathematics in the biological sciences. Its principal remit is to support the SERAD programme of biological research, which is carried out within the SABRI, SAC and RBGE. This is achieved through a dispersed group of statisticians, mathematics and computing experts based at BioSS centres in Edinburgh, Dundee, Aberdeen and Ayr. In 1999, new research funding allowed further expansion of work in veterinary epidemiology, particularly in understanding the dynamics of zoonotic diseases including E. coli 0157. Links were further strengthened with Scottish universities through establishment of a formal link with the Department of Statistics and Modelling Science, University of Strathclyde, focused principally on our new West of Scotland Unit, and the appointment of two research students with the University of Aberdeen. The SCRI group continues to support the rapidly expanding SERAD-funded programme in genomics. Recent activities of the group in this and other areas are described below.

Work in statistical genetics progressed with the development of methodology for estimating a linkage map from dominant and codominant markers in tetraploid species such as potato. This has been implemented in a computer program and will form the theoretical basis for QTL mapping in tetraploid species. Methodology for multi-trait QTL mapping in diploid species has also been developed, and applied to map QTLs affecting six yield characters in barley. The importance of this work was recognised by an invitation to review statistical methods for linkage analysis at a meeting organised by BBSRC and the British Council on 'New horizons for marker technologies and their application for cereal genomics.'

The release of new molecular sequence analysis software (EMBOSS) and new, Web-based, bioinformatics services has considerably affected the way that SCRI scientists analyse their molecular sequence data and the consequent support required from BioSS. Updated training courses are planned for 2000/01 when the EMBOSS graphical user interface becomes

available. Considerable progress has been made in developing improved hidden Markov models to detect evidence of mosaic sequences (due to recombination) in DNA multiple alignments, as part of a BBSRC-funded project under their Bioinformatics initiative. The method has been programmed in MATLAB. Collaboration continues with the Applied Statistics Department, University of Reading, on distance-based approaches to detecting recombination in large datasets. Among several collaborative projects with SCRI scientists involving phylogenetic methods, an analysis of protein sequence data was used to illustrate enzyme activity and down-regulation in transgenic plants.

Sampling designs have been devised for soil surveys in the Micronet project and other Soil-Plant Dynamics programmes. These involve establishing random transects from a set of systematically arranged origins and provide a compromise between providing an overall assessment of the site and quantifying spatial dependence by the variogram.