Professor David Balding
University College London

Title: Fine mapping of causal HLA variants using penalised regression

Abstract: Rheumatoid arthritis (RA) has been associated with markers in the HLA region, most notably with a group of *HLA-DRB1* alleles termed the shared epitope. There is evidence of additional risk loci but identification of them has been hampered by the extent of linkage disequilibrium. We address this problem using penalised regression, which imposes a penalty in the likelihood function that shrinks the estimated regression coefficients towards zero. The penalty can be interpreted as a Bayesian prior favouring effect sizes close to zero, but for computational reasons we restrict attention to point estimates obtained by maximising the penalised likelihood (i.e. the maximum *a posteriori* or MAP estimates). Ridge regression is one well-established penalised likelihood method that employs a Gaussian prior/penalty while the LASSO uses a Laplace (double exponential, DE) prior with a mode at zero so that the MAP estimates may also be zero. Whereas for ridge regression all MAP estimates are non-zero, the LASSO can generate sparse solutions with few non-zero estimates. We also investigate a generalisation of the LASSO called the HyperLasso (HLASSO), which uses the normal-exponential-gamma prior in place of the DE and can give a sharper mode and flatter tails than the DE. We analysed data from the Genetics Of Rheumatoid Arthritis (GoRA) study conducted by GlaxoSmithKline, which using genotype imputation we combined with the Wellcome Trust Case Control Consortium (WTCCC) SNPs from the HLA region. The North American Rheumatoid Arthritis Consortium (NARAC) study was used as a replicate study in which to validate our findings.