

A novel statistical model to identify biomarkers in 2D proteomic gels

S.Wu*, M.Black**, K.Atkinson[†], R.North^{††}, A.Rodrigo*

*Bioinformatics Institute, [†]School of Biological Sciences and ^{††}School of Medicine, University of Auckland, **University of Otago, New Zealand

Proteomic technologies are used to identify differentially expressed plasma proteins that may serve as biomarkers to predict disease. In this study, our aim is to use 2D-gel electrophoresis to identify sets of proteins in early pregnancy plasma that are associated with the subsequent development of pre-eclampsia, a severe hypertensive complication of pregnancy. However, due to technical issues, traditional statistical methods lack the power to detect significant changes in protein abundance between women with and without diseases.

We have developed a novel statistical model of 2D-gel data that incorporates both the probability that a spot is expressed and the conditional probability of expression intensity. The model also takes account of threshold detection levels. Using this model, we have gone on to develop two approaches to identifying spots implicated in differences between women with and without pre-eclampsia. These approaches use either a Likelihood Ratio Test or a Bayesian MCMC procedure to identify significant spots. In this talk, I present our model and discuss the relative merits of the two approaches we have developed.