Identifying epitopes in allergic reactions through hierarchical Bayesian modelling

S. Arima, A. Cuteri, V. Pecora, L. Tardella

Egg allergy represents one of the most common IgE-mediated food allergies: the EGG PROJECT led by the researchers of Department of Allergology, Università Cattolica Del Sacro Cuore (Roma) studies allergenic epitopes of ovalbumin (OVA) in patients with hen's egg allergy using the peptide microarray analysis, a novel method that can provide useful information on the nature of specific allergies. We aim at **1.** identifying which epitopes are recognized by allergic patients and **2.** comparing the epitopes identified before and after the treatment for those patients which successfully (*successes*) and unsuccessfully (*failures*) concluded the desensitization therapy. All allergic patients underwent oral desensitization to egg: a microarray immunoassay has been performed with sera from each patients with IgE-mediated egg allergy. A library of 125 peptides, consisting each of 15 amino acids overlapping by 12, has been printed in two set of triplicates onto SuperEpoxy glass slides. Since the peptide microarray is a very recent technology, there is currently no generally accepted approach to guide data analysis to detect epitope regions and make comparisons between alternative biological conditions. In this work we propose an alternative comprehensive model for peptide Signal to Noise Ratio (SNR) relying on a hierarchical structure to take into account relevant features which affect the measured expressions. The autocorrelation functions of the SNRs clearly suggest to discard the hypothesis of independence of measurements over consecutive peptides, mainly due to the overlapping amino acids in neighboring peptides. We propose a hierarchical Bayesian model which taking into account the interior spot dependence, addressed the scientific questions at stake. Borrowing from the econometric literature, the mean effect of the peptide is modeled as an autoregressive conditional root model and the transition between the state 1 (when the peptide belongs to an epitope region) and 0 (when the peptide does not belong to an epitope region) is modeled with a two-state (0 or 1) Markov process which allows to take into account the presence/absence of a peak and its spreading over the neighboring peptides. Using both real and simulated data, the proposed model has been compared with other models relying on different assumptions (e.g. the independence of the SNRs); in order to select the order of the autoregressive component, Bayesian model selection procedures have been performed.

S. Arima, Dip. di studi geoeconomici, linguistici, statistici e storici per l'analisi regionale, SAPIENZA Università di Roma, Italy (e-mail serena.arima@uniroma1.it).

A.Cuteri, Dip. di statistica, probabilità e statistiche applicate, SAPIENZA Università di Roma, Italy (email antonella.cuteri@uniroma1.it)

V. Pecora, Dip. di Allergologia, Università Cattolica Del Sacro Cuore, Policlinico A. Gemelli, Rome, Italy (e-mail valentina.pecora@hotmail.it)

L.Tardella, Dip. di statistica, probabilità e statistiche applicate, SAPIENZA Università di Roma, Italy (email luca.tardella@uniroma1.it)