GENETICS, SOFTWARE

CAREFUL CONTROL OF STRATIFICATION

16.12.2009

A careful control of population stratification removes most “significant” SNP associations – shown again at the Genetic Analysis Workshop 16

> Here we use stratification scores to achieve a unified and coherent control for confounding. We used the first 10 principal components, calculated genome-wide using a set of 81,500 loci that had been selected to have low pair-wise linkage disequilibrium, as risk factors in a logistic model to calculate the stratification score. We then divided the data into five strata based on quantiles of the stratification score. The VIF of these stratified data is 1.04, indicating substantial control of stratification. However, after control for stratification, we find that there are no significant loci associated with rheumatoid arthritis outside of the HLA region.

I did not know that

> the chip data contained several sources of distortion, including greatly elevated noise levels and artifact-prone SNP regions

Further control of confounding will remove even more associations (HLA maybe sacrosant) – putting us finally back into the 70ies, yea, yea.