

Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models

Eric Frichot ¹, Sean D Schoville ¹, Guillaume Bouchard ², Olivier François ¹

¹ TIMC-IMAG UMR 5525, Université Joseph Fourier Grenoble, Centre National de la Recherche Scientifique, Grenoble, France

² Xerox Research Center Europe, Meylan, France

Abstract. Local adaptation through natural selection plays a central role in shaping the genetic variation of populations. A way to investigate signatures of local adaptation, especially when beneficial alleles have weak phenotypic effects, is to identify polymorphisms that exhibit high correlation with environmental variables. However the geographical basis of both environmental and genetic variation can confound interpretation of these associations, as they can also result from genetic drift at neutral loci.

Here we propose an integrated framework based on spatial statistics, population genetics and ecological modeling for scans for signatures of local adaptation from genomic data. We present a novel class of algorithms to detect correlations between environmental and genetic variation that take account background levels of population structure and spatial autocorrelation in allele frequencies generated by isolation-by-distance mechanisms. Our framework uses Latent Factor Mixed Models, a hierarchical Bayesian mixed model in which environmental variables are fixed effects and population structure is introduced as random effects.

We implement fast algorithms that simultaneously estimate scores and loadings for the genotypic matrix and effects of environmental variables. Comparing these new algorithms with related methods provides evidence that LFMM can efficiently estimate random effects due to population history and isolation-by-distance patterns when computing gene-environment correlations, and decrease the number of false-positive associations in genome scans. We then apply these models to plant and human genetic data, identifying several genes with functions related to development that exhibit strong correlations with climatic gradients.