SBL – Bayesian Lasso for Detecting Rare Genetic Variants Associated with Survival Phenotypes

Rare genetic variants are one of the key factors in understanding the etiology of common diseases. Although much focus on the literature has been on rare single nucleotide variants (rSNVs), rare haplotype variants (rHTVs) offer perhaps even greater biological relevance, as variants on the same chromosomes are often passed jointly. Several methods have been developed to detect rHTV effects on common diseases based on the Bayesian Lasso methodology, collectively referred to as LBL, for binary and quantitative traits; greater power has been demonstrated over a number of rSNV-based methods. To further extend the capability of LBL, we have developed a method for detecting rHTVs associated with survival traits, referred to as Survival Bayesian Lasso (SBL). The current implementation of SBL is based on the accelerated failure time framework. SBL utilizes Weibull, loglogistic, and lognormal distributions to accommodate various types of potential hazards and interpretation schemes. A selection procedure is implemented in SBL to choose the most appropriate distribution. While SBL mainly focuses on rHTVs main effects, it can also evaluate environmental covariates as well as their interactions with HTVs. We applied SBL to the TCGA breast cancer dataset and identified a risk HTV that resides in the tumor suppressor CDH1 gene. We also conducted extensive simulations to gauge the performance of SBL.