ARVIN : Identifying Risk Noncoding Variants Using Disease-relevant Gene Regulatory Networks

Long Gao, Yasin Uzun, Kai Tan

October 31, 2017

Contents

1 Introcution

Identifying causalnoncoding variants remains a daunting task. Becausenoncoding variantsexert their effectsin the context of a gene regulatory network (GRN), we hypothesize that explicit use of disease-relevant GRNcan significantly improve the inference accuracy of noncoding risk variants. We describe Annotation of Regulatory Variants using Integrated Networks (ARVIN), a general computational framework for predictingcausal noncoding variants. For each disease, ARVIN first constructs a GRN using multi-dimensional omics data oncell/tissue-type relevant to the disease. ARVIN then uses a set of novel regulatory network-basedfeatures, combined with sequence-based features to make predictions.Using known causal variants ingene promoters and enhancersin a number of diseases, we show ARVIN outperforms state-of-the-art methods thatuse sequence-based features alone.