

Poster Title

A fast index structure for RNA-RNA interaction search

Authors

Cagri Aksay, Chi Kin Ho, Emre Karakoc, S. Cenk Sahinalp

Affiliation

School of Computing Science
Simon Fraser University
Vancouver, Canada

Poster Abstract

Non coding RNAs (ncRNAs) are known to regulate gene expression by establishing joint structures with their target mRNAs, inhibiting ribosomal access. A growing number of such ncRNAs and their known targets are available through newly developed public databases such as Rfam and NPinter. However, with the advent of novel sequencing technologies such as 454 which promises an exponential growth in the number of ncRNAs identified, fast methods are needed for searching a database for interacting candidates.

Here we propose an index structure for a collection of ncRNAs that enables the efficient searching for all potential interactions with a query mRNA. The index structure has two main components: (1) pRuNA, a sequence-based pruning technique which eliminates a significant fraction of the ncRNA collection, whilst retaining only the most likely candidates for forming a stable joint structure with the query mRNA; and (2) inteRNA, a fast technique for determining the joint secondary structure between two RNA molecules by examining interactions between different types of loop structures. Together, they provide a simple but effective index structure to quickly search an ncRNA collection for interacting sequences.

We tested our indexing technique using the complete ncRNA collection from the Rfam database, and used the NPinter database to query and validate known interactions. pRuNA was able to prune 86% of the ncRNA collection on average, whilst retaining the correct interacting RNA every time, which was confirmed using inteRNA. In the context of bacterial RNAs, pRuNA was able to correctly prune 91% of the collection.