Poster L-38

Meta Prediction of Human microRNA Targets

Authors:
Alal Eran (*Children's Hospital Informatics Program, Harvard-MIT Division of Health Sciences and Technology*)
Alvin Kho (*Children's Hospital Informatics Program, Harvard-MIT Division of Health Sciences and Technology*)
Iris Eisenberg (*Howard Hughes Medical Institute, Children's Hospital Boston*)
Michal Galdzicki (*Children's Hospital Informatics Program, Harvard-MIT Division of Health Sciences and Technology*)
Kamila Naxerova (*Children's Hospital Informatics Program, Harvard-MIT Division of Health Sciences and Technology*)
Marco Ramoni (*Children's Hospital Informatics Program, Harvard-MIT Division of Health Sciences and Technology*)
Louis Kunkel (*Howard Hughes Medical Institute, Children's Hospital Boston*)
Isaac Kohane (*Children's Hospital Informatics Program, Harvard-MIT Division of Health Sciences and Technology*)

Short Abstract: We present MAMI - MetA Mir:target Inference - a meta-predictor of human microRNA targets based on integrating five widely-used contemporary predictors. We conduct the first head-to-head performance evaluation of these predictors on hundreds of recently validated microRNA-target pairs, and show that MAMI performs significantly better than any single prediction method.

Long Abstract:
MicroRNAs comprise 2% of all known human genes, but most of their in vivo functions remain unknown. Several observation-based microRNA target prediction methods have been developed, but their predictive power has not been evaluated in a general setting. The recent explosion in functional microRNA studies has enabled us to conduct the first large-scale performance evaluation of these methods on hundreds of recently validated miR:target pairs. We developed MAMI – MetA Mir:target Inference – a meta-predictor of human microRNA targets that is based on integrating the five leading prediction methods. In silico validations show MAMI to perform significantly better than any single prediction method that it integrates.

Introduction
MicroRNAs (miRs) are central gene expression regulators, functioning as post transcriptional suppressors. In metazoa, these short non coding RNAs induce translational repression by binding to 3' UTRs of target mRNAs.
Target prediction is challenging because the recognition of a target mRNA occurs in the context of a protein complex (RISC), that does not require perfect sequence complementarity, nor thermodynamic stability between the miR and its target.
Several computational methods have been developed to predict miR targets in human, all based on rules derived from empirical observations, in an attempt to characterize the miR

binding site mechanics. The most sophisticated ones are TargetScanS and its predecessor TargetScan, miRanda, DIANA-microT, mirTarget and picTar. They all rely on observations made on a few in vivo validated miR:target pairs and were never subject to large-scale performance evaluations.
Recent developments in functional assays dramatically increased the availability of validated microRNA targets, which can be used as a benchmark dataset for a first large-scale performance evaluation of the predictors.

Common Ground

To facilitate head-to-head evaluations we first ensured that all methods that report a confidence score for their predictions share the same confidence score distribution. We then mapped those scores to a common range.
Each method predicts microRNA targets for different pools of mRNAs, and uses different nomenclature to report its predictions. 3' UTRs were chosen as the basis for uniform identification as all methods limit their analysis of a potential target to its 3' UTR. All transcripts with a common 3' UTR were clustered under an HGNC identifier or Entrez gene symbol, if the former didn't exist. Clustering was determined according to sequence alignment of the union of transcripts considered by the different predictors as candidate targets with their genomic origin.

Head-to-head performance evaluation

We compiled a set of 434 recently validated human microRNA-target pairs using TarBase and a manual literature review: 410 are experimentally-proven miR:target pairs, and 24 were specifically shown to have no in vivo interaction.
Each of the 6 predictors was asked to find an onto function between subsets of the 45 unique miRs and 379 unique targets that comprise the validated set. A "yes" vote was recorded for each pair included in that function. A "no" vote was recorded if either the miR or the target were in the domain/range of the function but not mapped to one another. An "abstained" vote was recorded for all other cases. The performance of each predicted function was evaluated in terms of the following parameters: overall sensitivity, specificity, accuracy, Matthews Correlation Coefficient (MCC) and area under the ROC curve (AUC).
All predictors that report a confidence score demonstrated no correlation between the quality of a prediction and its confidence score. Consequently there was no tradeoff between the specificity and sensitivity across different cutoffs, leading to abnormal ROC curves and nonexistent AUC.
The 2003 TargetScan algorithm scored a negative MCC, indicating worse than random predictions, but its improved version TargetScanS performed better than any other method, with 60% accuracy (62% sensitivity, 43% specificity) compared to 31% accuracy of the runner-up miRtarget.
DIANA-microT proved to be the most reliable method with 0.1 MCC. This quality comes at the expense of abstaining in most votes (17% coverage) with 19% accuracy (100% specificity, 9% sensitivity).
picTar had a negative MCC. When confined to a specific range of prediction scores that yield a positive MCC, picTar performed with 100% specificity and 5% sensitivity (11% accuracy).
miRanda and mirTarget's performances were also characterized by high specificity and poor sensitivity, regardless of their reported confidence scores (88%, 80% specificity with

17%,29% sensitivity, respectively).

Improved meta prediction

Based on the performance analyses, we integrated the strong points of each predictor into MAMI– the first meta-predictor of human microRNA targets. MAMI relies on TargetScanS, miRanda, DIANA-microT, miRtarget and a subset of picTar's predictions.

To minimize major differences between the training and test settings, the validated dataset was classified into families according to the number of predictors participating in a vote for a given miR-target pair. 30% of each family was randomly selected for the test set. Uninformative pairs winning abstained votes across all predictors were excluded.

As the performance evaluation showed no correlation between the reported confidence scores and the actual prediction quality, confidence scores were ignored and the integration was based on the overall performance of each predictor. To have the "best of both worlds" (TargetScanS's high sensitivity and the others' high specificity), MAMI predicts a true miR:target interaction whenever any predictor votes yes, with a confidence score of the sum of MCCs of predictors voting yes for this miR:target. MAMI was trained to predict false interactions by finding the cutoff of the sum of MCCs of the no voters that will optimize all performance parameters. When sum(MCCs of "no" voters) > 0.085, MAMI predicts no interaction for a miR and target in question.

MAMI's average performance across 1000 test sets showed 73% accuracy with an MCC of 0.8 - significantly better than any single predictor. This improvement does not sacrifice the coverage (0.64). MAMI's confidence score is based on large-scale performance evaluations and is tunable to desired specificity and sensitivity levels. While current methods provide predictions for positive miR:target interactions only, MAMI can also predict when a miR does not bind to a target.

With the accumulation of more validated microRNA targets, the observation-based predictors constituting MAMI can be refined, further improving MAMI's predictive power.