MiRror: a combinatorial analysis web tool for ensembles of microRNAs and their targets

Yitzhak Friedman\textsuperscript{1,†}, Guy Naamati\textsuperscript{2,†} and Michal Linial\textsuperscript{1,*}

\textsuperscript{1}Department of Biological Chemistry, Sudarsky Center for Computational Biology and \textsuperscript{2}School of Computer Science and Engineering, Hebrew University of Jerusalem, 91904, Israel

ABSTRACT

Summary: The miRror application provides insights on microRNA (miRNA) regulation. It is based on the notion of a combinatorial regulation by an ensemble of miRNAs or genes. miRror integrates predictions from a dozen of miRNA resources that are based on complementary algorithms into a unified statistical framework. For miRNAs set as input, the online tool provides a ranked list of targets, based on set of resources selected by the user, according to their significance of being coordinately regulated. Symmetrically, a set of genes can be used as input to suggest a set of miRNAs. The user can restrict the analysis for the preferred tissue or cell line. miRror is suitable for analyzing results from miRNAs profiling, proteomics and gene expression arrays.

Availability: http://www.proto.cs.huji.ac.il/mirror

Received on February 11, 2010; revised on June 2, 2010; accepted on June 3, 2010

1 INTRODUCTION

Recent miRNA discovery techniques confirmed the presence of hundreds of miRNAs in healthy and diseased tissues. Currently, approximately 900 miRNAs are reported in human and approximately 600 in mouse. It was shown that miRNA binding sites are not confined solely to 3′-untranslated regions (UTRs), but also are found in coding regions and introns (Kim, 2005). Furthermore, multiple features were shown to be imperative in the regulation by an ensemble of miRNAs or genes. miRror integrates predictions from a dozen of miRNAs with over approximately 600 in mouse. It was shown that miRNA binding sites are not confined solely to 3′-untranslated regions (UTRs), but also are found in coding regions and introns (Kim, 2005). Furthermore, multiple features were shown to be imperative in the regulation by miRNAs (Sethupathy et al., 2008; Nielsen et al., 2007). This led to a growing number of resources and algorithms that predict the pairing of miRNAs and their targets (Mendes et al., 2009). Currently, miRBase (Griffiths-Jones et al., 2008) is the most exhaustive collection of miRNAs with over approximately 15,000 mature miRNA sequences. miRBase collection (Enright et al., 2003; Griffiths-Jones et al., 2008; John et al., 2004) is based on MicroCosm algorithm. Additional exhaustive miRNA targets predictors include TargetScan (Lewis et al., 2003) and PiTar (Krek et al., 2005). Other large-scale resources for microRNA–target pairing include PITA (Kertesz et al., 2007), microRNA.org that is based on miRanda algorithm (Betel et al., 2008; Enright et al., 2003; John et al., 2004), DIANA-microT 3.0 (Marcogliese et al., 2009), miRDB (Wang, 2008) and EIMMO-MirZ (Hausser et al., 2009). For a subset of miRNAs–target pairs, in vivo experimental validation is available in TarBase database (Papadopoulos et al., 2009).

We have developed miRror, an interactive tool for analyzing experimental results under the notion of coordination in miRNAs regulation. A coordinated regulation by several miRNAs was proposed (see discussion (Ivanovska and Cleary, 2008; Krek et al., 2005)). For a set of miRNAs, miRror outputs a ranked list of gene targets according to their likelihood to be targeted by the miRNA ensemble and vice versa for a set of regulated genes. miRror converts all predictor resources results into a unified platform by incorporating a statistical measure according to the hypergeometrical distribution. For example, in miRNAs overexpressing experiments in human cells (Lim et al., 2005), 174 genes (miR-124) and 65 genes (miR-1) were downregulated and were used as input for miRror. The original miRNAs were identified as the top candidates with a P-value <3e-11 and with 91–93% of the genes marked as targets for the relevant miRNA. Similar results were obtained by using only 50% of the gene lists, supporting the robustness of the miRror platform.

2 SYSTEM OVERVIEW

miRror is a platform by which an experimentalist can gain biological insights on sets of molecular entities. Such sets are the product of large-scale experiments such as miRNA profiling, mass spectrometry proteomics and gene expression data. miRror operates at a dual mode where ensemble of (i) miRNAs and (ii) gene targets/proteins are used as the input.

2.1 Prediction tools resources integration

Database files from multiple miRNA–target prediction resources were separated to human, mouse and fly. The resources used include TargetScan, MicroCosm implemented in miRBase, PiTar, DIANA-MicroT, PITA, EIMMO-MirZ, miRanda-based microRNA.org, TargetRank, miRDB and TarBase. miRror combines the resources following a conversion of the miRNAs and gene targets identifiers. RefSeq and UniProtKB identifiers are indicated as the primary entries. TarBase used as a validation resource rather than a prediction database.

2.2 miRtegrate analysis

miRtegrate protocol is the core of miRror (Fig. 1). It calculates the probability of matches between miRNAs in the user list and the best set of targets derived from miRNA–target predictions as reported by each of the selected resources. For each gene, we
The miRror results are a ranked list of best predicted molecules. FlyAtlas (Chintapalli et al., 2007) calculations. Tissues are annotated by the gene expression data from different tissues and cell lines can be applied for the miRtegrate on several advanced options. Several background distributions such as different tissues and cell lines can be applied for the miRtegrate by the summation of predictions. As some miRNAs are predicted to pair with hundreds of targets, a reduction to a small but significant set of targets is valuable for the experimentalist. miRror application towards a set of genes as input is powerful in proposing potential miRNAs ensemble. Thus, the online tool can be used to generate hypotheses on the role of a specific genes or specific minimal set of miRNAs in any cellular settings.

**ACKNOWLEDGEMENTS**

We thank the ProtoNet team and Solange Karsenty for design and operation of the miRror website.

**REFERENCES**


3 CONCLUSIONS

We present miRror as a novel integration scheme for a dozen of miRNA-target prediction resources. miRror provides a double-sided view for ensembles of miRNAs or genes. Based on the notion of combinational regulation adopted by miRror, the reported list of targets for a miRNAs ensemble is a small subset of the inclusive list that is produced by the summation of predictions. As some miRNAs are predicted to pair with hundreds of targets, a reduction to a small but significant set of targets is valuable for the experimentalist. miRror application towards a set of genes as input is powerful in proposing potential miRNAs ensemble. Thus, the online tool can be used to generate hypotheses on the role of a specific genes or specific minimal set of miRNAs in any cellular settings.