BIOINFORMATICS ORIGINAL PAPER Vol. 27 no. 22 2011, pages 3093–3101
doi:10.1093/bioinformatics/btr534

Genome analysis Advance Access publication September 27, 2011

A wide repertoire of miRNA binding sites: prediction and functional implications

Naama Elefant, Yael Altuvia and Hanah Margalit*

Department of Microbiology and Molecular Genetics, IMRIC, Faculty of Medicine, The Hebrew University of Jerusalem, POB 12272, Jerusalem 91120, Israel

Associate Editor: Ivo Hofacker

ABSTRACT

Motivation: Over the past decade, deciphering the roles of microRNAs (miRNAs) has relied heavily upon the identification of their targets. Most of the targets that were computationally and experimentally characterized were evolutionarily conserved 'seed' targets, containing a perfect 6–8 nt match between the miRNA 5′-region and the messenger RNA (mRNA). Gradually, it has become evident that other types of miRNA binding can confer target regulation, but their characterization has been lagging behind.

Results: Here, we complement the putative evolutionarily-conserved seed-containing targets by a wide repertoire of putative targets exhibiting a variety of miRNA binding patterns, predicted by our algorithm RepTar. These include non-conserved sites, 'seed' binding sites with G:U-wobbles within the seed, '3′ compensatory' sites and 'centered' sites. Apart from the centered sites, we demonstrate the functionality of these sites and characterize the target profile of a miRNA by the types of binding sites predicted in its target 3′ UTRs. We find that different miRNAs have individual target profiles, with some more inclined to seed binding and others more inclined to binding through 3′ compensatory sites. This diversity in targeting patterns is also evident within several miRNA families (defined by common seed sequences), leading to divergence in the target sets of members of the same family. The prediction of non-conventional miRNA targets is also beneficial in the search for targets of the non-conserved viral miRNAs. Analyzing the cellular targets of viral miRNAs, we show that viral miRNAs use various binding patterns to exploit cellular miRNA binding sites and suggest roles for these targets in virus–host interactions.

Availability: All RepTar's predictions are available for simple and advanced querying at http://reptar.ekmd.huji.ac.il

Contact: hanahm@ekmd.huji.ac.il

Supplementary Information: Supplementary data are available at *Bioinformatics* online.

Received on March 29, 2011; revised on August 29, 2011; revised on September 6, 2011

1 INTRODUCTION

MicroRNAs (miRNAs) have emerged as key regulators of gene expression, exerting their regulatory function post-transcriptionally by base-pairing with the mRNA [reviewed in Ref. (Bartel, 2009)]. Over the past decade, hundreds of miRNA genes have been discovered in metazoa, plants and viruses. While it has been demonstrated that quite a few of them have important roles in central cellular processes [reviewed in Ref. (Wienholds and Plasterk, 2005)], understanding these roles relies heavily upon identification of their target genes. Consequently, many experimental and computational efforts have been made to identify miRNA targets. Most of the first discovered miRNA targets exhibited evolutionary conservation and full complementarity in their 3′-untranslated region (3′ UTR) to a short sequence of ∼7 nt in the 5′ region of the miRNA, termed seed [reviewed in Ref. (Bartel, 2009)]. Therefore, the main features of miRNA binding sites that were initially used to predict miRNA targets were the presence of a seed and evolutionary conservation of the site [reviewed in Ref. (Bartel, 2009)]. These features narrowed the vast number of possible base-pairings between mRNA 3′ UTRs and miRNAs, and directed most research in the field to seed-containing targets. However, two of the originally discovered miRNA targets in *Caenorhabditis elegans*, as well as additional targets identified along the years, lacked perfect seed pairing, but rather exhibited extensive base-pairing between the mRNA 3′ UTR and the miRNA 3′ region, termed '3′ compensatory' sites (Brennecke *et al.*, 2005; Reinhart *et al.*, 2000; Slack *et al.*, 2000; Stern-Ginossar *et al.*, 2007). Furthermore, the miRNA binding options were recently expanded even further with the identification of 'centered sites', functional miRNA binding sites that lack both perfect seed and 3′ compensatory pairing but exhibit extensive pairing between the mRNA 3′ UTR and the center region of the miRNA (Shin *et al.*, 2010). All these additional types of sites were far less studied; non-evolutionarily conserved targets can be predicted by several algorithms (Betel *et al.*, 2010; Friedman *et al.*, 2009; John *et al.*, 2004; Kertesz *et al.*, 2007; Maragkakis *et al.*, 2009; Marín and Vanícek, 2011; Miranda *et al.*, 2006; Muckstein *et al.*, 2006; Rehmsmeier *et al.*, 2004; Smith *et al.*, 2010), the 3′ compensatory targets are predicted by only few (Betel *et al.*, 2010; Friedman *et al.*, 2009) and no target prediction algorithm to date predicts centered sites (see Supplementary Material S1 and Supplementary Table S1). Consequently, much is still unknown regarding the functionality and the biological implications of these non-conventional targets. This shortage may be even more pronounced when targets of viral miRNAs are considered, as these miRNAs are not evolutionarily conserved and one of their first discovered cellular targets (based on RepTar's predictions) contained a 3′ compensatory site (Stern-Ginossar *et al.*, 2007).

Here, we provide a wide repertoire of targets obtained by our algorithm 'RepTar', which enables identification of targets with various binding types. The set of RepTar's predictions includes

*To whom correspondence should be addressed.

evolutionarily conserved and organism-specific non-conserved targets that have either multiple or single binding sites of various types: seed binding sites (that may include G:U wobbles within the seed), 3′ compensatory binding sites and centered sites.

We applied RepTar to human and mouse miRNAs as well as to miRNAs from several viruses and provide our predictions in a database that enables simple and advanced querying at http://reptar.ekmd.huji.ac.il (Elefant *et al*., 2011). Using experimental data we are able to demonstrate that most of the non-conventional sites mediate down-regulation of their targets, and that the combination of a 3′ compensatory binding site with a perfect seed site confers as efficient down-regulation as two perfect seed sites. Characterizing the viral and cellular miRNAs by their binding site types we suggest functional implications for differential targeting.

2 METHODS

2.1 Databases of miRNA and 3′ UTR sequences

Seven sets of miRNAs (including sequences of the 'star' miRNAs-miRNA*) were extracted from the miRBase version 15.0 (http://microrna.sanger.ac.uk/) (Griffiths-Jones, 2004; Griffiths-Jones *et al*., 2006; Griffiths-Jones *et al*., 2008): a dataset of all human miRNAs (1100 miRNAs), a dataset of all mouse miRNAs (717 miRNAs), three datasets of miRNAs of human-infecting viruses: human cytomegalovirus (17 miRNAs), Kaposi sarcoma-associated virus (25 miRNAs) and Epstein Barr virus (44 miRNAs) and two datasets of mouse-infecting viruses: mouse cytomegalovirus (29 miRNAs) and mouse gammaherpesvirus (10 miRNAs).

The set of human 3′ UTR sequences and related data were extracted from University of California, Santa Cruz (UCSC) human genome assembly (build 36.1, hg18), dated March 2006 (http://genome.ucsc.edu). Nearly 24 407 unique refGene 3′ UTR sequences ranging from the stop codon to the polyadenylation site and defined as non-intronic regions were retrieved. Each RefSeq accession was then attributed to the corresponding gene-symbol accession using the kgXref table. This process grouped together RefSeq accessions of alternatively spliced sequences. The sequences of these alternatively spliced 3′ UTR variants were aligned using fasta version 3.4 with Smith–Waterman alignment and variants with ∼80% sequence identity were represented once by the longest common sequence. This procedure resulted in a final set of 18 421 3′ UTR sequences.

The set of mouse 3′ UTR sequences was extracted from the UCSC mouse genome assembly (build 37.1, mm9), dated July 2007 (http://genome.ucsc.edu). In total, 19 613 unique RefSeq 3′ UTR sequences were retrieved and processed, as described above, resulting in a final set of 17 069 sequences.

2.2 Dataset of experimentally-determined direct targets

Human and mouse validated miRNA targets were retrieved from the TarBase database version V5 June 2008 (Papadopoulos *et al*., 2009). Only 'TRUE' targets with 'direct support' were retrieved. Of these, several targets were excluded because of absence of the miRNA/3′ UTR in our database or differences between the reported miRNA sequence and the one in our database. This resulted in 197 miRNA:target pairs. These are listed in Supplementary Table S2, along with their annotation as predicted or not predicted by RepTar.

Cellular targets of viral miRNAs were collected from several papers that were published at the time of the analysis (Supplementary Table S3). Of the 23 direct cellular targets, two targets were excluded from the analysis: one target did not appear in our 3′ UTR database and the other, major histocompatibility complex class I-related chain B (MICB), was discovered based on RepTar's prediction and therefore excluded (Stern-Ginossar *et al*., 2007). This resulted in 21 miRNA:target pairs for our assessment.

2.3 mRNA and protein expression datasets

Experimental data of the change in mRNA expression levels and protein expression levels were retrieved from Baek *et al*. (2008) and Selbach *et al.* (2008). Data from three experiments of over-expression of miR-1, miR-124 and miR-181a and one experiment of knock-down of mouse miRNA miR-223 (Baek *et al*., 2008), as well as data from five experiments of overexpression of let-7b, miR-1, miR-16, miR-30a and miR-155 (Selbach *et al*., 2008) were used. Experimental data of the change in mRNA expression levels were retrieved from Grimson *et al*. (2007) for nine experiments of miRNA overexpression of miR-7, miR-9, miR-122, miR-128, miR-132, miR-133a, miR-142-3p, miR-148b and miR-181a. Only probes whose values exceeded the median intensity before transfection were considered.

In all analyses based on the above experimental data, the KS test was applied to compare between groups of predicted targets and all other genes in the experiment (referred to as background).

2.4 Dataset of binding sites determined in PAR-CLIP experiments

We downloaded experimentally verified Ago-interaction regions in mRNAs from Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP) experiments in Human embryonic kidney 293 (HEK293) cells from (Hafner *et al.*, 2010). In total, 17 319 regions were downloaded and of these, 5532 mapped to our database of 3′ UTRs. Within these regions we searched for RepTar's predicted miRNA binding sites for the top 100 miRNAs expressed in HEK293 cells requiring that the site resides fully within the region and that the T to C transition be within the predicted binding site.

2.5 Specific binding patterns for individual miRNAs

The predicted targets were divided according to their binding patterns (seed only, 3′ compensatory only and mixed). We then searched for miRNAs that had a greater fraction of 3′ compensatory targets than seed targets. These miRNAs were required to have at least 50 predicted targets and a fraction of 3′ compensatory targets that exceeded the fraction of seed targets by at least 5%. This resulted in 84 miRNAs and 8 star miRNAs with a fraction of 3′ compensatory targets that was larger than their fraction of seed targets. We compared the nucleotide distribution and expression profile of these 92 miRNAs showing a larger fraction of 3′ compensatory targets with that of 937 miRNAs showing a higher fraction of seed targets. For the expression analysis, we compared the number of tissues where the miRNAs in each group were expressed, using the miRNA expression atlas of Landgraf *et al*. (2007).

2.6 Overlap of viral and human miRNA binding sites

For the analysis of overlap between predicted binding sites of viral and cellular miRNAs we considered as overlapping, sites that differed in their genomic start position by no more than 5 nt. The statistical significance of the overlap between the sites was assessed for each pair of viral and cellular miRNAs using the hypergeometric test with false discovery rate (FDR) correction for 0.05.

3 RESULTS AND DISCUSSION

3.1 The RepTar algorithm

3.1.1 Overview of the algorithm To identify conventional, as well as non-conventional binding sites in 3′ UTRs of genes, we developed a predictive algorithm, RepTar, which is not restricted to evolutionarily conserved seed sites. The main steps of the algorithm are described in Figure 1 and detailed in Supplementary Material S1. Briefly, it begins by searching for statistically significant repeating motifs in each 3′ UTR, which are then

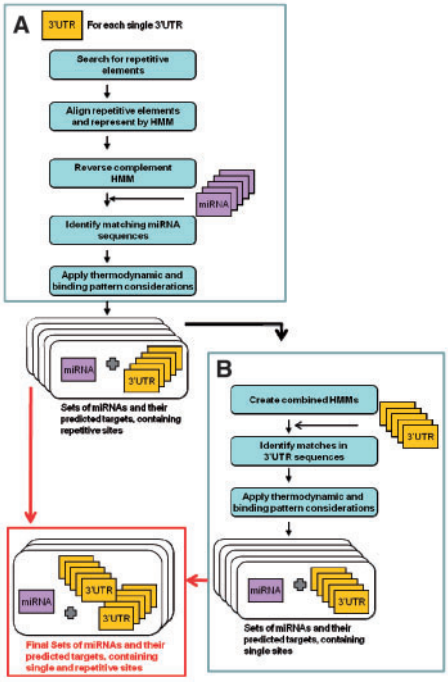

Fig. 1. Schematic overview of the RepTar algorithm. (**A**) First we identify targets containing repetitive binding sites for each miRNA, by characterizing repetitive elements in each 3′ UTR. (**B**) The sets of miRNAs and their putative 3′ UTR targets are then used in the next stage to identify additional targets containing single binding sites. The final set of RepTar's predictions consists of conventional and non-conventional targets identified in both stages.

represented by a profile Hidden-Markov-Model (HMM). These HMMs are reversed and complemented so that they represent the sequence that is reverse-complementary to the motif. The complementary HMMs are searched against a database of miRNA sequences, to identify pairs of complementary HMMs and miRNAs. To allow all types of miRNA binding sites, we specifically use the HMM representation that permits insertions and deletions in the alignment and therefore searches for perfectly, as well as imperfectly matching miRNAs. Using HMMs, we also award statistical significance to G:U pairing. Consequently, RepTar's initial search does not impose any restriction on the location of the match within the miRNA (whether it is in the 5′-end or 3′-end) or on the type of miRNA binding site.

Once statistically significant matching miRNAs are identified, several filtering criteria are applied, including the binding pattern and the thermodynamic stability of the miRNA:mRNA hybrid (Supplementary Material S1). Restrictions are imposed also at the gene level: for a gene to be considered a putative target gene, its 3′ UTR must contain at least two qualified binding sites. This requirement for repetitiveness provides a constricted and more reliable set of miRNA binding sites, without considering evolutionary conservation.

Finally, the constricted set of repetitive binding sites is used to define statistical models by which we identify targets with non-repetitive sites or with sites that lack a high scoring repetitive motif (Supplementary Material S1).

3.1.2 Application of RepTar RepTar was applied in search of targets of human and mouse miRNAs in human and mouse 3′ UTRs, respectively. It was also applied in search of cellular targets of several viral miRNAs: Human 3′ UTR sequences were searched for matches to miRNAs of human cytomegalovirus (HCMV), Kaposi sarcoma-associated herpes virus (KSHV) and Epstein Barr virus (EBV); Mouse 3′ UTR sequences were searched for matches to miRNAs of mouse cytomegalovirus (MCMV) and mouse gammaherpesvirus (MGHV) (for more details see Section 2).

3.1.3 Evaluation of RepTar To evaluate the RepTar algorithm, we assessed its sensitivity, precision and functionality in several ways. First, we turned to experimentally determined direct targets by small-scale experiments extracted from the TarBase database (Papadopoulos *et al*., 2009). RepTar successfully predicted 142 out of 197 reported direct targets of human and mouse miRNAs, defining a sensitivity of 72% on this dataset ($P \leq 2.5e\text{-}64$ by a hypergeometric test) (Supplementary Table S2). Of note, the direct targets of the TarBase database are comprised mainly of validated predictions of existing algorithms and therefore includes mainly conventional binding sites. Similarly, to validate RepTar's predictions of cellular targets of viral miRNAs we compiled a set of 21 experimentally validated direct cellular targets of viral miRNAs (Supplementary Table S3). RepTar correctly predicted 15/21 cellular targets of viral miRNAs, constituting a sensitivity of 71%, similar to that obtained for the human and mouse miRNAs.

Next we turned to large databases of targets obtained from large-scale experiments. We began by assessing the precision and sensitivity of RepTar by conducting the same analysis as was recently performed by Alexiou *et al*. (2009) on the same data of measured protein level changes after overexpression of five miRNAs (Selbach *et al*., 2008). In the study of Alexiou *et al*., the sensitivity and precision of miRNA target prediction algorithms was assessed relative to a set of genes defined as target genes based on the down-regulation of their protein levels below an assigned threshold ($\log_2$-fold change < -0.2). Following these definitions, RepTar's precision and sensitivity on this data were 25 and 23% respectively, demonstrating a lower precision than the less permissive miRNA target prediction algorithms but a higher sensitivity than available algorithms [see Ref. (Alexiou *et al*., 2009; Elefant *et al*., 2011)] (Supplementary Material S1). RepTar's ability to predict a wide variety of binding sites is expected to create higher false positive rates and render it less precise but in turn it is expected to award it higher sensitivity. Indeed, when filtering RepTar's predictions to include only perfect seed sites (full Watson–crick binding of positions 2–7 or 2–8) we obtain for the same data a higher precision of 39% but a lower sensitivity of 14%.

We then used the large-scale experiments to assess RepTar's predictions in terms of biological functionality, this time using the whole spectrum of changes in gene expression, without setting pre-defined thresholds of down-regulation (or up-regulation). We used both the data described above, as well as data from two additional large-scale experiments (Baek *et al*., 2008; Grimson *et al*., 2007) and evaluated the change in gene expression of the genes in the group of RepTar's predicted targets compared with all other genes in the data. In the experiment by Grimson *et al*. (2007) mRNA expression levels were measured upon overexpression of nine different human miRNAs. In the experiments of Baek *et al*. (2008) and Selbach *et al*. (2008) both changes in mRNA and protein

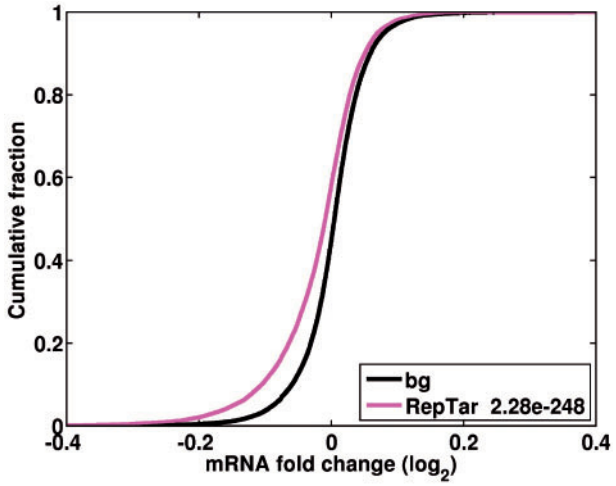

Fig. 2. Assessment of RepTar's predictions using large-scale mRNA expression data. RepTar's predictions were evaluated using pooled data from nine large-scale experiments measuring changes in mRNA expression under overexpression of a miRNA (Grimson *et al.*, 2007). The $\log_2$-fold change values of RepTar's predicted targets (RepTar) were compared with all other genes (bg), by a one-sided KS test. The predicted targets were statistically significantly down-regulated compared with the background genes. The *P*-value appears in the box.

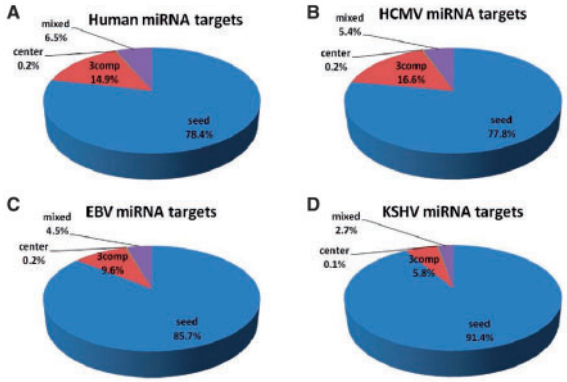

Fig. 3. Distribution of miRNA target genes. (**A**) Distribution of target types based on miRNA:mRNA pairs predicted for all human miRNAs: 'seed' targets—targets containing only seed binding sites, 'comp' targets—targets containing only 3′ compensatory binding sites, 'centered' targets—targets containing only centered sites and 'mixed' targets—containing more than one type of binding site. (**B–D**) Distribution of target types among cellular targets of viral miRNAs (binding site types are as in A).

expression levels were measured upon overexpression (and in two cases suppression) of three human and one mouse miRNAs and of five human miRNAs, respectively. For each of these experiments, we show that RepTar's predicted targets are statistically significantly down-regulated (or up-regulated in the case of the suppressed miRNAs) compared with all other genes on the array [KS test, $P < 0.05$ (Fig. 2 and Supplementary Figs. S1 and S2)]. In the same manner, we compared the change in cellular mRNA expression levels upon transfection of endothelial cells and B-cells with the cluster of KSHV miRNAs (Sébastien Pfeffer and Juergen Haas, personal communication). RepTar's predicted targets are statistically significantly down-regulated compared with the background genes [one-sided Kolmogorov–Smirnov (KS) test, $P \leq 0.05$, data not shown], confirming the validity and biological relevance of these viral miRNA target predictions as well.

Thus, when considering the full set of RepTar's predictions we find that these predicted targets exhibit statistically significant down-regulation in the presence of the corresponding miRNA. The set of RepTar's predictions is composed of targets that are bound through different binding patterns (e.g. seed, 3′ compensatory) and are characterized by various features (e.g. G:U pairing in the seed).

We next turned to characterize and assess the functionality of the various types of binding sites.

3.2 Various miRNA binding types

3.2.1 Distribution of miRNA binding types The most abundant type among RepTar's predicted binding sites is the seed site. Nearly 81% of RepTar's predicted sites are seed sites, while 18.6% are 3′ compensatory and <0.5% are centered sites. This is consistent with the general consensus that most miRNA binding sites are seed sites. While seed sites require pairing of 6-, 7- or 8-mers in the 3′ UTR, 3′ compensatory and centered sites require longer stretches of matches and can be expected to be less abundant than the seed sites. Many of our predicted target genes have several predicted binding sites within the sequence. When we assign each target gene a binding type according to the predicted binding sites that it contains, we find that 78% of the predicted human targets are 'seed targets' (contain only seed sites), 15% are '3′ compensatory targets' (contain only 3′ compensatory sites), <0.5% are 'centered targets' (contain only centered sites) and 6.5% are 'mixed targets' containing several binding sites of different types within their 3′ UTR (Fig. 3A).

3.2.2 Biological validity of the various predicted binding types To study the function of the various types of RepTar's predicted targets we used the experimental data of the change in mRNA expression levels described above (Grimson *et al.*, 2007). We divided the predicted targets by type and compared the expression fold change of the predicted targets in each group with that of the background genes (all genes not predicted as targets). In addition to the basic categorization of targets into types (Fig. 3), we studied additional categories of groups of targets: conserved target group (targets containing at least one binding site that showed evolutionary conservation), non-conserved target group (targets containing only non-evolutionarily conserved sites), perfect seed group (targets that contained only seed sites with at least one perfect seed) and seed with G:U wobble group (targets that contained only seed sites with G:U wobbles). The small number of targets included in the centered group was insufficient for achieving statistical significance and therefore, was not included in this analysis. For all other types of targets we find that each group, classified by the binding type, is statistically significantly down-regulated in the experimental data (Fig. 4A–C). These include evolutionarily conserved targets, non-conserved targets (which are statistically significantly down-regulated but less down-regulated than the conserved sites), seed targets, seed targets with G:U wobbles (which are statistically significantly down-regulated but less down-regulated than the Watson–Crick seed targets), and 3′ compensatory targets.

The functionality of the different binding sites has been gradually acknowledged. Non-conserved sites have been shown to be

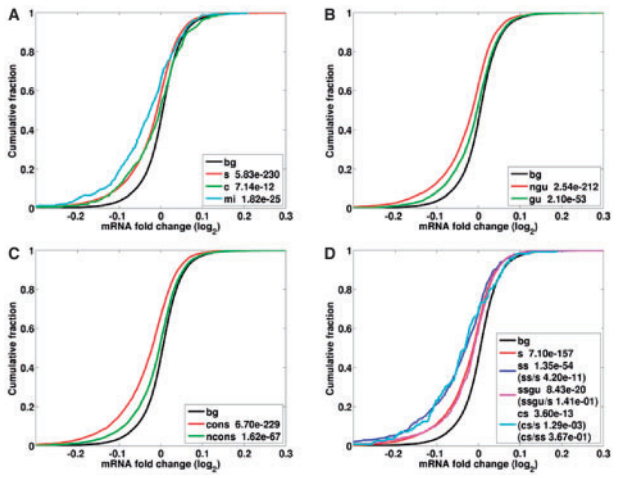

Fig. 4. Functionality of non-conventional targets. Log_2-fold change in mRNA expression levels of different groups of predicted target genes was compared that of the background genes (bg). Analysis as described in Figure 2A. (**A**) RepTar's predicted targets were divided into three groups according to their types: seed (s), 3′ compensatory (c) and mixed (m). All groups show a statistically significant down-regulation, with mixed targets showing the most significant down-regulation. (**B**) RepTar's predicted targets were divided into targets containing perfect seed matches with no G:U base-pairing (ngu) and targets containing seed matches with G:U base-pairing (gu). Both the 'ngu' and the 'gu' groups were statistically significantly down-regulated, although the 'gu' group is less down-regulated than the 'ngu' group. (**C**) RepTar's predicted targets were divided into evolutionarily conserved targets (cons) and targets that are not evolutionarily conserved (ncons). Both the 'ncons' and the 'cons' groups were statistically significantly down-regulated, although the 'ncons' group is less down-regulated than the 'cons' group. (**D**) RepTar's predicted targets were divided into those containing one perfect seed site (s), two perfect seed sites (ss), one perfect seed site and one seed site with G:U pairing (ssgu) and targets containing one 3′ compensatory site and one perfect seed site (cs). The log_2-fold change in mRNA expression levels of each group was compared with the 'bg' group using one-sided KS and to the 's' group using two-sided KS test. In addition, the log_2 fold change in mRNA expression levels of the 'cs' group was compared with the 'ss' group using two-sided KS test. All groups were statistically significantly down-regulated compared with the 'bg'. As expected, the 'ss' group was statistically significantly more down-regulated than the 's' group. In contrast, the 'ssgu' group was not statistically significantly more down-regulated than the 's' group and surprisingly, the 'cs' group did not differ statistically significantly from the 'ss' group, suggesting an important contribution of the 3′ compensatory site. (The first *P*-values are for the comparison with the 'bg', the following parentheses contain the two groups being compared and the *P*-value from the two-sided KS test comparing between the groups). *P*-values of all comparisons with 'bg' and with 's' are statistically significant with Bonferroni correction.

functional (Farh *et al.*, 2005) and are accounted for in many prediction algorithms. The functionality of seed sites with G:U wobbles was addressed in several papers (Didiano and Hobert, 2006; Doench and Sharp, 2004; Miranda *et al.*, 2006; Tay *et al.*, 2008), resulting in contradicting findings, and only some of the prediction algorithms account for these sites. The 3′ compensatory sites are still neglected by most prediction algorithms (Supplementary Table S1, Supplementary Material S1). Our results show that seed sites with G:U base-pairing within the seed are effectively down-regulated, providing additional support to their functionality. Likewise for the 3′ compensatory binding sites, initially discovered in *C.elegans* (Reinhart *et al.*, 2000; Slack *et al.*, 2000) and then reported in *Drosophila melanogaster* (Brennecke *et al.*, 2005). Our results show that targets with such non-conventional binding patterns can undergo functional down-regulation by the corresponding miRNAs.

A substantial portion of our predicted targets are mixed targets, containing more than one type of binding site (see also Supplementary Material S1). Interestingly, conducting the same analysis on this group of genes, we find that the mixed targets show the highest level of down-regulation (Fig. 4A). This high level of down-regulation may be due to the mere presence of a seed site, to the presence of multiple binding sites, or specifically to the combination of different types of binding sites. To assess whether this beneficial targeting is due to the combination of different types of sites, we compared between the down-regulation of groups of targets with different combinations of sites (Fig. 4D). As only few centered sites are found, these are not included in the analysis and we address mainly the combinations of different seed and 3′ compensatory sites. First, we compared targets with two perfect seed sites (not allowing G:U pairing) with targets with one perfect seed site. As expected, targets with two seed sites show a higher level of down-regulation compared with targets with a single seed site [as previously shown by Grimson *et al.* (2007)]. We then considered targets with one perfect seed and one seed with G:U pairing and compared them with targets with a single perfect seed site. In this case, we found that the down-regulation of the targets with two sites did not differ statistically significantly from the targets with a single perfect seed site, suggesting that the addition of a seed-G:U site does not contribute substantially to the down-regulation of a target already containing a perfect seed site. Finally, we compared mixed targets containing one perfect seed site and one 3′ compensatory site with targets with a single perfect seed site and with targets with two perfect seed sites. These mixed targets were statistically significantly more down-regulated than the single seed targets and, surprisingly, showed a down-regulation as substantial as the double seed targets (Fig. 4D). This was even more surprising in light of the results of comparing targets with single sites of different types, which showed that 3′ compensatory sites are statistically significantly more down-regulated than seed-G:U sites but not as efficient as perfect seed sites (Supplementary Fig. S3). Therefore, we suggest that it is the combination of 3′ compensatory and seed sites that has a synergistic effect and mediates down-regulation as effectively as the combination of two perfect seed sites.

We further validated RepTar's predicted binding sites and their biological functionality, using data of Ago-interaction regions in mRNAs from PAR-CLIP experiments in HEK293 cells (Hafner *et al.*, 2010). For the top 100 miRNAs expressed in HEK293 cells, we find that RepTar predicts miRNA binding sites in 36% of the experimentally verified regions in 3′ UTRs, similar to the portion predicted by Hafner *et al.* (40% of the regions). Within these RepTar predicted sites, 69% are perfect seed sites, 28% are seed sites with G:U wobbles and 3% are 3′ compensatory sites. The overall fraction of RepTar predicted 3′ compensatory targets for these 100 miRNAs is 9%. In the experimental data on changes in protein expression levels used above (Selbach *et al.*, 2008), 3% of the down-regulated proteins are predicted to be 3′ compensatory targets, whereas the overall fraction of RepTar predicted 3′ compensatory targets in this dataset is 5%. Some of RepTar's predictions are expected to be false positives, and it is possible that there is a higher false positive rate

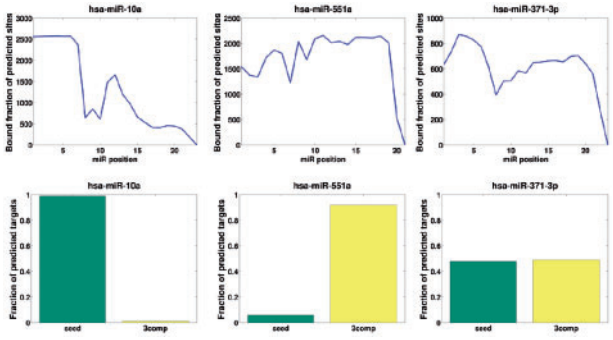

Fig. 5. miRNA binding patterns. (**A**) Binding patterns of individual miRNAs: for each position, the fraction of binding sites in which the position is bound is determined. miRNAs may show a preference for binding mostly through their seed positions, implying a preference for seed binding sites, (left panel), they may show excessive binding at non-seed binding positions implying a preference for binding 3′ compensatory sites (middle panel), or show comparable tendencies to both (right panel). (**B**) Target profile of individual miRNAs: for each miRNA we compute the distribution of targets by their different types. The examples demonstrate that miRNAs differ in their target profiles: most of the targets of hsa-miR-10a are seed targets, those of hsa-miR-551a are mostly 3′ compensatory targets, and hsa-miR-371-3p binds equally both types of targets. There is a high compatibility between the binding patterns of the miRNAs (A) and their target profiles (B).

among the predicted 3′ compensatory targets. Nonetheless, these results indicate that a portion of the predictions are indeed functional, and specifically, that the non-conventional sites predicted by RepTar can be functional.

3.2.3 Binding types of specific miRNAs Next, we turned to see if any of the miRNAs show a preference to a specific type of binding site, i.e. if predicted targets for a certain miRNA are enriched with one type of site over another (we refer to the distribution of miRNA targets classified by their binding site types as the miRNA's 'target profile'). For this analysis, we considered the seed targets and 3′ compensatory targets, which make up 93% of all predicted targets. Whereas most miRNAs show a target profile enriched with seed targets, some have equal representation of seed targets and 3′ compensatory targets and for 92 miRNAs, we find an enrichment of 3′ compensatory targets (Fig. 5, see Section 2). For several of these miRNAs, the 3′ compensatory fraction exceeds 80%. Interestingly, while conducting the same analysis on mouse miRNAs, we found a correlation between the target profiles of the human miRNAs and their orthologous mouse miRNAs (Spearman correlation, $r = 0.59$; $P < 0.0001$, Supplementary Fig. S4), where many of the orthologous mouse miRNAs show the same 3′ compensatory tendency. This suggests that individual miRNAs may have different targeting preferences and that some of these preferences may be conserved between human and mouse.

The diversity in target profiles may stem from differences in sequence composition between the miRNAs. Indeed, the 92 3′ compensatory-inclined miRNAs are richer in G nucleotides compared with the seed-inclined miRNAs ($P < 0.0001$ for whole miRNAs and $P \leq 0.0001$ for seed regions by chi-squared tests). This higher G composition is not region-specific, as the 5′- and 3′-ends show a similar distribution of nucleotides. As 3′ UTR sequences in human and mouse are generally A/U rich, it would be expected that miRNAs showing an abundance of extensively matching miRNA sites, such as are the 3′ compensatory sites, would be A/U rich as well. It is therefore interesting that these G-rich miRNAs are 3′ compensatory-inclined.

3.2.4 Functional implications of miRNA target profiles The differences in targeting profiles may have functional implications. One possible function is to allow differential regulation of members of the same miRNA family. Members of miRNA families are often similarly expressed, share the same seed and differ only in 3′ sequence. These miRNAs would share the same seed-mediated targets but could differ in other types of targets. Indeed, we find that many of the 3′ compensatory-inclined miRNAs are members of miRNA families where the other members are seed-inclined miRNAs (Table 1). For instance, in the let-7 family, all miRNAs are seed-inclined (with >80% of their targets being seed targets) except for miRNA-202 that exhibits 56% 3′ compensatory targets, or the miR-320 family exhibiting two seed-inclined miRNAs and two 3′ compensatory-inclined miRNAs. Thus, for some miRNA families, the use of differential targeting patterns by the members of the family can cause divergence in the target sets of the miRNAs in the family awarding individual functions for each member.

Several other functions may be proposed for these target profiles. The difference in binding sites may confer differences in the type of regulation, just as sites with near-perfect complementarity are thought to mediate degradation rather than translational repression. It is known that factors such as concentration of the mRNA molecules, mRNA half life and auxiliary RNA binding proteins affect miRNA targeting (Arvey *et al*., 2010; Jacobsen *et al*., 2010; Larsson *et al*., 2010). When considered in the context of the cell or tissue being investigated, it would be interesting to see if these factors contribute differently to the various types of targets. A hint to the possibility that these miRNAs may serve slightly different functions can be seen in the differential expression that they exhibit. 3′ compensatory-inclined miRNAs show lower expression levels and are expressed in fewer tissues compared with seed-inclined miRNAs ($P \leq 0.009$ by a two-sided Wilcoxon test) [based on the data in Ref. (Landgraf *et al*., 2007), see Section 2].

3.3 Cellular targets of herpesvirus miRNAs

3.3.1 Target profiles of cellular targets of herpesvirus miRNAs An additional group of targets provided by RepTar comprises cellular targets of viral miRNAs. As viral miRNAs are generally poorly evolutionarily conserved (Cai *et al*., 2005; Schafer *et al*., 2007) and as a 3′ compensatory target of a viral miRNA has been proven functional (Stern-Ginossar *et al*., 2007), the prediction of both conventional and non-conventional targets of viral miRNAs encompasses a greater set of potential cellular targets of viral miRNAs.

Consistent with the profile characteristics of the human miRNAs, we find that most viral miRNA cellular targets are seed targets, a non-negligible fraction are 3′ compensatory targets, a small fraction are mixed targets and few are centered targets (Fig. 3B–D). Investigating possible functional implications for the miRNA binding types of viral miRNAs we find that the group of 3′ compensatory targets of HCMV miRNAs are enriched in genes that regulate apoptosis, and genes that are involved in lymphocyte activation and differentiation

Table 1. Diversity of target profiles within miRNA family members

Human miRNA	Sequence	Profile (%)		Human miRNA	Sequence	Profile (%)	
		seed	comp			seed	comp
miR-99astar	CAAGCTCGCTTCTATGGGTCTG	16	**82**	miR-1322	GATGATGCTGCTGATGCTG	24	**65**
miR-99bstar	CAAGCTCGTGTCTGTGGGTCCG	**66**	27	miR-1272	GATGATGATGGCAGCAAATTCTGAAA	**99**	**0**
miR-150star	CTGGTACAGGCCTGGGGGACAG	19	**75**	miR-563	AGGTTGACATACGTTTCCC	27	**69**
miR-1263	ATGGTACCCTGGCATACTGAGT	**65**	32	miR-380star	TGGTTGACCATAGAACATGCGC	**100**	0
miR-202	AGAGGTATAGGGCATGGGAA	26	**56**	miR-4295	CAGTGCAATGTTTTCCTT	24	**63**
let-7a	TGAGGTAGTAGGTTGTATAGTT	**96**	2	miR-130a	CAGTGCAATGTTAAAAGGGCAT	**98**	1
let-7b	TGAGGTAGTAGGTTGTGTGGTT	**83**	8	miR-130b	CAGTGCAATGATGAAAGGGCAT	**98**	1
let-7c	TGAGGTAGTAGGTTGTATGGTT	**91**	4	miR-301a	CAGTGCAATAGTATTGTCAAAGC	**81**	14
let-7d	AGAGGTAGTAGGTTGCATAGTT	**96**	1	miR-301b	CAGTGCAATGATATTGTCAAAGC	**98**	1
let-7e	TGAGGTAGGAGGTTGTATAGTT	**86**	5	miR-454	TAGTGCAATATTGCTTATAGGGT	**71**	17
let-7f	TGAGGTAGTAGATTGTATAGTT	**96**	2				
let-7g	TGAGGTAGTAGTTTGTACAGTT	**93**	3				
let-7i	TGAGGTAGTAGTTTGTGCTGTT	**100**					
miR-98	TGAGGTAGTAAGTTGTATTGTT	**100**					
miR-320c	AAAAGCTGGGTTGAGAGGGT	41	**50**	miR-518a-3p	GAAAGCGCTTCCCTTTGCTGGA	15	**83**
miR-320d	AAAAGCTGGGTTGAGAGGA	29	**59**	miR-518c	CAAAGCGCTTCTCTTTAGAGTGT	40	**55**
miR-320a	AAAAGCTGGGTTGAGAGGGCGA	**77**	18	miR-518f	GAAAGCGCTTCTCTTTAGAGG	5	**94**
miR-320b	AAAAGCTGGGTTGAGAGGGCAA	**83**	12	miR-518d-3p	CAAAGCGCTTCCCTTTGGAGC	**81**	16
miR-378b	ACTGGACTTGGAGGCAGAA	3	**94**	miR-1825	TCCAGTGCCCTCCTCTCC	16	**74**
miR-378	ACTGGACTTGGAGTCAGAAGG	**67**	23	miR-199a-5p	CCCAGTGTTCAGACTACCTGTTC	**98**	2
miR-378c	ACTGGACTTGGAGTCAGAAGAGTGG	**99**	1	miR-199b-5p	CCCAGTGTTTAGACTATCTGTTC	**88**	8
miR-422a	ACTGGACTTAGGGTCAGAAGGC	**84**	11				
miR-500b	AATCCTTGCTACCTGGGT	29	**58**	miR-4265	CTGTGGGCTCAGCTCTGGG	33	**46**
miR-362-5p	AATCCTTGGAACCTAGGTGTGAGT	**92**	6	miR-4296	ATGTGGGCTCAGGCTCA	**51**	27
miR-501-5p	AATCCTTTGTCCCTGGGTGAGA	**81**	14	miR-4322	CTGTGGGCTCAGCGCGTGGGG	**58**	25
miR-1254	AGCCTGGAAGCTGGAGCCTGCAGT	34	**53**	miR-3131	TCGAGGACTGGTGGAAGGGCCTT	38	**54**
miR-661	TGCCTGGGTCTCTGGCCTGCGCGT	**95**	2	miR-151-5p	TCGAGGAGCTCACAGTCTAGT	**99**	1
miR-3116	TGCCTGGAACATAGTAGGGACT	**87**	8	miR-611	GCGAGGACCCCTCGGGGTCTGAC	**92**	5
miR-3137	TCTGTAGCCTGGGAGCAATGGGGT	25	**63**	miR-4298	CTGGGACAGGAGGAGGAGGCAG	24	**61**
miR-20bstar	ACTGTAGTATGGGCACTTCCAG	**98**	1	miR-1302	TTGGGACATACTTATGCTAAA	**100**	0
miR-596	AAGCCTGCCCGGCTCCTCGGG	29	**62**	miR-4299	GCTGGTGACATGAGAGGC	20	**67**
miR-4293	CAGCCTGACAGGAACAG	**93**	4	miR-548q	GCTGGTGCAAAAGTAATGGCGG	**99**	1

Listed are several miRNA families and the profile distribution of the predicted human targets for the members of the families. Each box represents a family and the first member (or members) is a 3′ compensatory-inclined miRNA, whereas, other members are seed-inclined (as indicated by their percentage of miRNA targets in each profile). Nucleotides underlined represent the seed of the miRNAs. Bold percentages indicate the dominant profile.

(by gene ontology annotation, $P \leq 0.003$, $P \leq 0.003$, $P \leq 0.02$, respectively). HCMV is known for its manipulative functions on its host cell, which include evasion of immune mediators and preventing cell apoptosis (Mocarski, 2002, 2004). Furthermore, the first HCMV miRNA cellular target to be discovered (based on RepTar's predictions) was the MICB gene, an NKG2D ligand mediating cell death by natural killer cells, which contained a non-evolutionary conserved, 3′ compensatory binding site of the virus miRNA UL112. Therefore, it is possible that for HCMV miRNAs, the 3′ compensatory targeting is a means for immune system and apoptosis regulation.

Another interesting binding type is the 'near-perfect complementarity' site, a binding pattern where extensive complementarity exists in both the seed and the 3′ region. This type of site is thought to mediate mRNA degradation rather than translational repression [reviewed in Ref. (Wienholds and Plasterk, 2005)]. Relatively, few such targets are predicted for human, mouse and viral miRNAs. Near-perfect complementarity to viral miRNAs has been reported in viral mRNA targets, mostly in genes encoded on the opposite strand of the viral miRNA [reviewed in Ref. (Gottwein and Cullen, 2008)]. However, such near-perfect complementarity has not yet been identified in

cellular target genes. RepTar predicts several such cellular targets of the viral miRNAs and this group contains dozens of genes that can limit viral replication, as well as cytokines, apoptosis-inducing genes and tumor suppressors, which may contribute to the known tumorigenicity of some of the herpes viruses (Supplementary Table S4). Among these are the DNA fragmentation factor alpha subunit (DFFA) and ataxia telangiectasia mutated (ATM) genes, both containing near-perfect complementarity binding sites for HCMV, as well as EBV miRNAs. The DFFA gene is a substrate of caspase-3 and triggers DNA fragmentation during apoptosis. The ATM gene encodes a protein that is considered one of the master regulators of the cell cycle and is known to regulate the tumor suppressor genes BRCA1 and P53. It would be to the virus' advantage to target such genes that endanger its survival and such near-perfect complementarity may mediate strong down-regulation or mRNA degradation.

3.3.2 Viral miRNAs may use non-conventional binding to exploit cellular miRNA binding sites In order to target cellular genes it is quite possible that viral miRNAs may attempt to use cellular miRNA binding sites [such targeting was demonstrated in Ref. (Gottwein *et al*., 2007; Nachmani *et al*., 2010; Stern-Ginossar *et al*., 2008)]. While several viral miRNAs show sequence similarity to human miRNAs (Gottwein *et al*., 2007; Stern-Ginossar *et al*., 2008) most viral miRNAs differ in sequence from human miRNAs. Therefore, in order to exploit cellular binding sites, the viral miRNA may bind the same area but with a slight shift in binding positions and/or use a different binding patterns in order to match the cellular miRNA's binding site. As RepTar's initial search for both cellular and viral miRNA binding sites stems from the identification of repeating regions in the 3′ UTR, it would not be surprising if predicted viral miRNA binding sites would overlap predicted cellular binding sites. Indeed, we find many overlapping cellular and viral sites, but, interestingly, we find that a quarter to a third of these overlapping sites are predicted to be bound by the cellular and viral miRNAs through different binding patterns (see for instance the predictions in the apoptosis related gene DIDO in Fig. 6). In addition, we find several individual viral miRNAs that show a higher than expected overlap between their predicted targets and those of a cellular miRNA ($P < 0.05$ by a hypergeomertic test with FDR correction for 0.05, see Table 2). For these highly overlapping miRNAs, a certain degree of sequence similarity is observed as well. One example is the KSHV miRNA miR-K12-11, known to function as an orthologue of the human miRNA miR-155 and showing a substantial overlap in predicted binding sites (Gottwein *et al*., 2007). Another example is the KSHV miRNA miR-K12-7 for which its predicted sites highly overlap the predicted sites of the human miRNA miR-216a. The human miRNA miR-216 was shown to be a negative regulator of apoptosis (Cheng *et al*., 2005) and it is possible that the KSHV miR-K12-7 has evolved to exploit the binding sites of hsa-miR-216 acquiring its anti-apoptotic function in order to increase the virus' survival within the host cell.

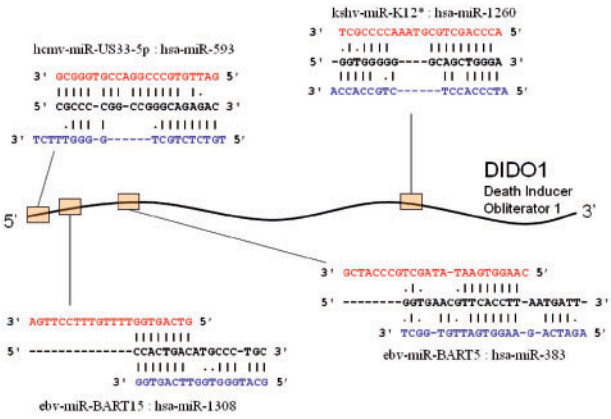

Fig. 6. Viruses may exploit binding sites of cellular miRNAs using a different binding type. An illustration of the DIDO (Death Inducer Obliterator 1) 3′ UTR and several predicted binding sites within this sequence. This illustration demonstrates an overlap between a viral 3′ compensatory site and a cellular seed site (first box from the 5′-end) and three overlapping sites of viral seed sites with cellular 3′ compensatory sites (second to forth boxes from the 5′-end). 3′ UTR sequences appear in black, viral miRNA sequences are the upper sequence and appear in red and cellular miRNA sequences are the lower sequences and appear in blue. Of note, the predictions of different binding sites for miRNAs from all three herpes viruses (HCMV, KSHV and EBV) in this gene may suggest convergent evolution of these viruses to target this important gene involved in apoptosis regulation.

Table 2. Viral and human miRNAs with overlapping binding sites

miRNAs[a]	Sequence Alignment of Human and Viral miRNAs[b]
hsa-miR-216a	UAAUCUCA----GCUGGCAACUGUGA
kshv-miR-K12-7	UGAUCCCAUGUUGCUGGC-GCU----
hsa-miR-1324	CCAGACAGAAUUCUAUGCACU–UUC-
ebv-miR-BART11-5p	UCAGACAG–UUUGGUGCGCUAGUUG
hsa-miR-155	UUAAUGCUAA--UCGUGAUAGGGGU
kshv-miR-K12-11	UUAAUGCUUAGCCUGUGUCCGA---
hsa-miR-29a	UAGCACCAUCUGAAAUCGGUUA----
ebv-miR-BART1-3p	UAGCACCGCU----AUCCACUAUGUC
hsa-miR-2114	UAGUCCCUUCCU–UGAAGCGGUC
kshv-miR-K12-7	UGAUCCCAUGUUGCUGGCGC---
hsa-miR-3065-3p	UCAGCACCAGGAUAUUGUUGGAG--
kshv-miR-K12-1*	GCAGCACCUG---UUUCCUGCAACC
hsa-miR-141*	CAUCUUCCAGUAC-AGUGUUGGA
kshv-miR-K12-2*	GAUCUUCCAGGGCUAGAGCUG--
hsa-miR-499-3p	AACAUCA------CAGCAAGUCUGUGCU
ebv-miR-BART4*	CACAUCACGUAGGCACCAGGUGU-----

[a]Listed are pairs of human (upper) and viral (lower) miRNAs with an overlap of at least 25% of their predicted sites and for which this overlap was statistically significant (hypergeometric test with FDR correction for 0.05).

[b]The miRNA sequences were aligned using the the european molecular biology open software suite (EMBOSS) procedure needle. Default parameters were used, except for the substitution matrix in which substitutions between A and G and between C and U were scored favorably (1 instead of −4).

4 CONCLUSIONS

In summary, we expand the repertoire of seed sites with additional non-conventional sites, which have been somewhat overlooked. While the prediction of this wide repertoire of sites is expected to produce a lower specificity with false positive predictions as well, the assessment of these targets against experimental data strongly supports their functionality. Furthermore, the proposed functions for the different types of sites emphasize their contribution toward

more effective down-regulation and underscore the importance of predicting such sites. Thus, by providing this comprehensive set of targets, we hope to promote further investigation of the roles and types of regulation mediated by both conventional and non-conventional miRNA targets.

ACKNOWLEDGEMENTS

We thank Sébastien Pfeffer and Juergen Haas for sharing with us unpublished experimental data. We are grateful to Naomi Habib, Zohar Itzhaki and Ariel Jaimovich for their thoughtful comments on the manuscript.

Funding: This study was supported by grants from the Israeli Science Foundation administered by the Israel Academy of Sciences and Humanities; US-Israel Bi-National Science Foundation, Israeli Cancer Research Fund (to H.M.) and Azrieli Fellowship (to N.E.).

Conflict of Interest: none declared.

REFERENCES

Alexiou,P. *et al.* (2009) Lost in translation: an assessment and perspective for computational microRNA target identification. *Bioinformatics*, **25**, 3049–3055.

Arvey,A. *et al.* (2010) Target mRNA abundance dilutes microRNA and siRNA activity. *Mol. Syst. Biol.*, **6**, 363.

Baek,D. *et al.* (2008) The impact of microRNAs on protein output. *Nature*, **455**, 64–71.

Bartel,D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215–233.

Betel,D. *et al.* (2010) Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.*, **11**, R90.

Brennecke,J. *et al.* (2005) Principles of microRNA-target recognition, *PLoS Biol.*, **3**, e85.

Cai,X. *et al.* (2005) Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells. *Proc. Natl Acad. Sci. USA*, **102**, 5570–5575.

Cheng,J. *et al.* (2005) Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science*, **308**, 1149–1154.

Didiano,D. and Hobert,O. (2006) Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat. Struct. Mol. Biol.*, **13**, 849–851.

Doench,J.G. and Sharp,P.A. (2004) Specificity of microRNA target selection in translational repression. *Genes Dev.*, **18**, 504–511.

Elefant,N. *et al.* (2011) RepTar: a database of predicted cellular targets of host and viral miRNAs. *Nucleic Acids Res.*, **39**, D188–D194.

Farh,K.K. *et al.* (2005) The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science*, **310**, 1817–1821.

Friedman,R.C. *et al.* (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.*, **19**, 92–105.

Gottwein,E. and Cullen,B.R. (2008) Viral and cellular microRNAs as determinants of viral pathogenesis and immunity. *Cell Host Microbe*, **3**, 375–387.

Gottwein,E. *et al.* (2007) A viral microRNA functions as an orthologue of cellular miR-155. *Nature*, **450**, 1096–1099.

Griffiths-Jones,S. (2004) The microRNA Registry, *Nucleic Acids Res*, **32**, D109–D111.

Griffiths-Jones,S. *et al.* (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.*, **34**, D140–D144.

Griffiths-Jones,S. *et al.* (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res.*, **36**, D154–D158.

Grimson,A. *et al.* (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell*, **27**, 91–105.

Hafner,M. *et al.* (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*, **141**, 129–141.

Jacobsen, A. *et al.* (2010) Signatures of RNA binding proteins globally coupled to effective microRNA target sites. *Genome Res.*, **20**, 1010–1019.

John,B. *et al.* (2004) Human MicroRNA targets. *PLoS Biol.*, **2**, e363.

Kertesz,M. *et al.* (2007) The role of site accessibility in microRNA target recognition. *Nat. Genet.*, **39**, 1278–1284.

Landgraf,P. *et al.* (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*, **129**, 1401–1414.

Larsson,E. *et al.* (2010) mRNA turnover rate limits siRNA and microRNA efficacy. *Mol. Syst. Biol.*, **6**, 433.

Maragkakis,M. *et al.* (2009) DIANA-microT web server: elucidating microRNA functions through target prediction. *Nucleic Acids Res.*, **37**, W273–W276.

Marín,R.M. and Vanícek,J. (2011) Efficient use of accessibility in microRNA target prediction. *Nucleic Acids Res*., **39,** 19–29.

Miranda,K.C. *et al.* (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell*, **126**, 1203–1217.

Mocarski,E.S., Jr. (2002) Immunomodulation by cytomegaloviruses: manipulative strategies beyond evasion. *Trends. Microbiol.*, **10**, 332–339.

Mocarski,E.S., Jr. (2004) Immune escape and exploitation strategies of cytomegaloviruses: impact on and imitation of the major histocompatibility system. *Cell Microbiol.*, **6**, 707–717.

Muckstein,U. *et al.* (2006) Thermodynamics of RNA-RNA binding. *Bioinformatics*, **22**, 1177–1182.

Nachmani, D. *et al.* (2010) The human cytomegalovirus microRNA miR-UL112 acts synergistically with a cellular microRNA to escape immune elimination. *Nat. Immunol.*, **11**, 806–813.

Papadopoulos,G.L. *et al.* (2009) The database of experimentally supported targets: a functional update of TarBase. *Nucleic Acids Res.*, **37**, D155–D158.

Rehmsmeier,M. *et al*. (2004) Fast and effective prediction of microRNA target duplexes. *RNA*, **10**,1507–1517.

Reinhart,B.J. *et al.* (2000) The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*, **403**, 901–906.

Schafer,A. *et al.* (2007) Cloning and analysis of microRNAs encoded by the primate gamma-herpesvirus rhesus monkey rhadinovirus. *Virology*, **364**, 21–27.

Selbach,M. *et al.* (2008) Widespread changes in protein synthesis induced by microRNAs. *Nature*, **455**, 58–63.

Shin,C. *et al.* (2010) Expanding the microRNA targeting code: functional sites with centered pairing. *Mol. Cell*, **38**, 789–802.

Slack,F.J. *et al.* (2000) The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol. Cell*, **5**, 659–669.

Smith,C. *et al*. (2010) Freiburg RNA tools: a web server integrating INTARNA, EXPARNA and LOCARNA. *Nucleic Acids Res.*, **38,** W373–W377.

Stern-Ginossar,N. *et al.* (2007) Host immune system gene targeting by a viral miRNA. *Science*, **317**, 376–381.

Stern-Ginossar,N. *et al.* (2008) Human microRNAs regulate stress-induced immune responses mediated by the receptor NKG2D. *Nat. Immunol.*, **9**, 1065–1073.

Tay,Y. *et al.* (2008) MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature*, **455**, 1124–1128.

Wienholds,E. and Plasterk,R.H. (2005) MicroRNA function in animal development. *FEBS Lett*., **579,** 5911–5911.