Widespread context dependency of microRNA-mediated regulation

Florian Erhard, Jürgen Haas, Diana Lieber, et al.

Genome Res. 2014 24: 906-919 originally published online March 25, 2014
Access the most recent version at doi:10.1101/gr.166702.113

Supplemental Material

http://genome.cshlp.org/content/suppl/2014/04/09/gr.166702.113.DC1.html

Related Content

A genomic portrait of the genetic architecture and regulatory impact of microRNA expression in response to infection
Katherine J. Siddle, Matthieu Deschamps, Ludovic Tailleux, et al.
Genome Res. May 2014 24: 850-859

References

This article cites 51 articles, 19 of which can be accessed free at:
http://genome.cshlp.org/content/24/6/906.full.html#ref-list-1

Articles cited in:
http://genome.cshlp.org/content/24/6/906.full.html#related-urls

Creative Commons License

This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to Genome Research go to:
http://genome.cshlp.org/subscriptions

© 2014 Erhard et al.; Published by Cold Spring Harbor Laboratory Press
Widespread context dependency of microRNA-mediated regulation

Florian Erhard, Jürgen Haas, Diana Lieber, Georg Malterer, Lukasz Jaskiewicz, Mihaela Zavolan, Lars Dölken, and Ralf Zimmer

Gene expression is regulated in a context-dependent, cell-type-specific manner. Condition-specific transcription is dependent on the presence of transcription factors (TFs) that can activate or inhibit its target genes (global context). Additional factors, such as chromatin structure, histone, or DNA modifications, also influence the activity of individual target genes (individual context). The role of the global and individual context for post-transcriptional regulation has not systematically been investigated on a large scale and is poorly understood. Here we show that global and individual context dependency is a pervasive feature of microRNA-mediated regulation. Our comprehensive and highly consistent data set from several high-throughput technologies (PAR-CLIP, RIP-chip, 4sU-tagging, and SILAC) provides strong evidence that context-dependent microRNA target sites (CDTS) are as frequent and functionally relevant as constitutive target sites (CTS). Furthermore, we found the global context to be insufficient to explain the CDTS, and that flanking sequence motifs provide individual context that is an equally important factor. Our results demonstrate that, similar to TF-mediated regulation, global and individual context dependency are prevalent in microRNA-mediated gene regulation, implying a much more complex post-transcriptional regulatory network than is currently known. The necessary tools to unravel post-transcriptional regulations and mechanisms need to be much more involved, and much more data will be needed for particular cell types and cellular conditions in order to understand microRNA-mediated regulation and the context-dependent post-transcriptional regulatory network.

[Supplemental material is available for this article.]
bind at a certain binding site. This is an example of individual context, since an RBP is able to affect individual target sites of a microRNA. In contrast, dynamic microRNA expression determines a global context, i.e., if a microRNA is differentially expressed, all of its targets are affected. This important distinction is often neglected; and often, only the global context is considered (Carroll et al. 2013).

Several examples of individual context-specific microRNA-mediated regulation can also be found in the literature (for a review, see Pasquini et al. 2012). Bhattacharyya et al. (2006) identified the RNA binding protein, ELAVL1, as a derepressor for miR-122 regulation of the SLC7A1 mRNA. In normal hepatocarcinoma cells, SLC7A1 is repressed by a miR-122 target site in its 3′-UTR. Under different stress conditions, ELAVL1 is released from the nucleus into the cytoplasm, which abolishes SLC7A1 repression. The exact mechanism, however, remains unclear. Intriguingly, ELAVL1 has also been implicated in activating a target site of the microRNA let-7 in the 3′-UTR of MYC (Kim et al. 2009), which indicates that ELAVL1 can both induce and prevent microRNA-mediated regulation. In addition to ELAVL1, DND1 (Kedde et al. 2007) and PUM1 (Kedde et al. 2010) have also been identified to influence microRNA regulation. There may be other RNA binding proteins that interfere with or facilitate microRNA/target interactions.

These examples illustrate that the presence of a functional target site is not sufficient for regulation. It may be active under certain conditions but nonfunctional in a different context. Presently, our knowledge about context-dependent microRNA-mediated regulation is only based on a few examples, and the underlying molecular mechanisms are poorly understood.

By immunoprecipitation of microRNA/target complexes using monoclonal antibodies to RISC complex components followed by high-throughput sequencing of the protein-protected microRNA target sites, the complete targetome of cellular and viral microRNAs has become accessible. More than 10,000 putative microRNA binding sites, so-called clusters, are obtained in a single HTS-CLIP (high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation) or PAR-CLIP (photo-activated gbonucleotide-enhanced crosslinking immunoprecipitation) experiment. Although the annotation of the responsible microRNA to an identified cluster still leaves room for improvement, >75% of microRNA target interactions can be correctly annotated, thereby allowing in-depth analyses of microRNA regulatory networks (Gottwein et al. 2011; Haecker et al. 2012; Riley et al. 2012; Skalsky et al. 2012; Erhard et al. 2013a).

To study context-specific microRNA-mediated regulation, we generated AGO2-PAR-CLIP data from two human B-cell lines. In addition, we reanalyzed two recently published sets of AGO2-PAR-CLIP data from two different human B-cell lines (Gottwein et al. 2011). These four cell lines represent different stages of B-cell development and are either infected by Kaposi’s sarcoma-associated herpes virus (KSHV), coinfected by both KSHV and Epstein-Barr Virus (EBV), or not infected. Thus, each cell line provides a distinct context for microRNA-mediated regulation. All data sets were reanalyzed using a new algorithm called PArma (Erhard et al. 2013a). PArma considers the topology of the microRNA/target interaction and the position of UV light-induced cross-links in more detail than state-of-the-art methods, and it provides quality control scores for both—the identification of microRNA target site clusters and the annotation of the interacting microRNA to these sites. For two of these four cell lines, we generated three additional data sets, including RIP-chip, 4sU-tagging-derived RNA half-lives, and large-scale SILAC-based proteomics. This allowed us to comprehensively analyze the effect of context-dependent microRNA/target interactions on the recruitment of the target mRNAs to AGO2 complexes, on target RNA stability, and on target protein levels. By considering viral as well as host microRNAs, we investigated both microRNA/target interactions that coevolved within a species as well as interactions of an exogenous microRNA with endogenous target sites. The results provide compelling evidence that context dependency of microRNA-mediated regulation is not restricted to a few examples but is a widespread and general feature of post-transcriptional regulation mediated by both cellular and viral microRNAs.

**Results**

**Differential analysis of PAR-CLIP data**

To comprehensively study regulation of cellular gene expression by both cellular and Kaposi’s sarcoma-associated herpes virus (KSHV)-encoded microRNAs, we applied AGO2-PAR-CLIP to two human B-cell lines—the body cavity-based lymphoma cell line BCBL1, which is latently infected with KSHV, and the Burkitt lymphoma cell line DG75, which is KSHV negative. Applying PArma (Erhard et al. 2013a) with stringent criteria (see Supplemental Methods), we identified 15,577 clusters, 12,333 of which mapped to known transcripts (Ensembl v60).

In order to assess the quality of the PAR-CLIP data sets, we first computed the positional distribution of all target sites in mRNAs (Fig. 1A). Target sites of viral microRNAs shared the well-described features of cellular microRNA target sites: They preferentially bind to the 3′ untranslated region (3′-UTR) and rarely to the 5′-UTR of transcripts (Grimson et al. 2007; Hafner et al. 2010). Within the 3′-UTR, target sites tend to accumulate at the very beginning, i.e., immediately after the stop codon, and at the transcript end, i.e., immediately upstream of the poly-A tail (Grimson et al. 2007). We furthermore checked the accuracy of the microRNA assignment to target sites by confirming that virtually no reads mapped to KSHV microRNA target sites in the KSHV negative cell line DG75 (a feature that is not used by PArma to assign microRNAs) (see Fig. 1B). The few instances with random reads in DG75 may nevertheless be bona fide KSHV microRNA target sites. As we observed random reads spread across a multitude of transcripts at low frequency, these reads presumably result from infrequent unspecific immunoprecipitates or insufficient removal of background total RNA rather than microRNA-specific signatures. This is further supported by a significantly lower frequency of T to C conversions and lower consistency across replicates for these reads (Fig. 1B; Supplemental Fig. S1).

We further validated our PAR-CLIP data set using published data for the same cell lines: (1) PAR-CLIP targets are highly consistent with RIP-chip data (Fig. 1C,D; Dölken et al. 2010); (2) KSHV microRNA targets are selectively enriched in BCBL1 and not DG75 in the RIP-chip experiments (cf. Fig. 1C,D); and (3) PAR-CLIP target sites lead to a measurable reduction of target mRNA half-lives (Fig. 1E; Supplemental Table S2).

To be able to perform a more in-depth analysis on KSHV microRNA targets in human B-cells, we also included recently published PAR-CLIP data from two additional B-cell lines, namely BC1 and BC3 (Gottwein et al. 2011). We reanalyzed all data sets using PArma, which yielded 21,628 clusters, 16,425 of which mapped to known transcripts (see Supplemental Table S1; see Supplemental Table S6 for individual read mapping statistics).
Figure 1. (Legend on next page)
Intriguingly, the overlaps of target sites of both ubiquitously expressed cellular and KSHV microRNAs were surprisingly small (Fig. 2A,B). Such extreme differences of called target sites may be due to experimental bias or context dependency, i.e., a major fraction of microRNA target sites is only active in some of the cell lines considered.

There are some traces of batch effects (Fig. 2B), but experimental bias does not explain all differences: Replicate experiments for all four cell lines yielded highly reproducible results (Fig. 2B,C; Supplemental Fig. S1), and differences were also observable for inner-laboratory comparisons, i.e., there are target sites of constitutively expressed microRNAs that are reproducibly missing in DG75 and reproducibly observable in BCBL1 and vice versa, and these cell lines have been measured simultaneously, using the exact same protocol in the same laboratory (the same is true for BC1 and BC3) (see Supplemental Discussion for a detailed analysis of technical bias and further discussion). To further analyze potential context dependency of microRNA/target site interactions, we used a per-microRNA differential analysis of all PAR-CLIP experiments (Supplemental Figs. S4, S8–S11).

Intriguingly, when we considered all target sites of a single microRNA, there was no clear correlation of target sites across cell lines (Fig. 2C; Supplemental Fig. S2). Instead, distinct clusters of target sites emerged, for instance, several kshv-miR-K12-4-3p target sites that appear to be active in BCBL1 only and not in BC1 or BC3. This suggests that context-dependent microRNA-mediated regulation may be substantially more important than generally expected. An obvious explanation for this would be that all these target sites are not expressed in BC1 and BC3 or are expressed at very low levels such that PAR-CLIP is not able to identify them. This would be one kind of global context. Low levels or absence of mRNA is detectable by expression measurements that are indicated on the left side of the heatmaps in Figure 2C and is, in fact, not a general feature of these target sites, arguing in favor of an individual context that determines the activity of many of these sites. Additionally, there are mutually exclusive target sites (MES): Target sites of the same microRNA that are missing in BCBL1 and active in BC1 or BC3 and target sites missing in BC1 and BC3 but active in BCBL1, i.e., the target sites in one condition are not a subset of the target sites in the other condition. Thus, the presence of MES cannot be explained by a higher expression or activity of the respective microRNA or mRNA, i.e., by global context.

Context-dependent target sites of KSHV microRNAs

Taken together, our differential analysis of PAR-CLIP data suggests that microRNA-mediated regulation is substantially and generally dependent on the cellular context. To experimentally test this hypothesis, we used three sets of additional high-throughput methods to investigate the consequences of context-dependent microRNA-mediated regulation. First, using RIP-chip, we tested whether context-dependent microRNA/target interactions, as found in the PAR-CLIP data, had a measurable impact on the recruitment of the target mRNA to RISC in their specific context only. Second, using microarray-based transcriptomics, including metabolic labeling of RNA, and SILAC-based proteomics experiments, we tested whether such context-dependent microRNA/target interactions also have a measurable impact on mRNA half-lives and on mRNA as well as on protein levels of their targets in their specific context only. All these experiments were performed by comparing DG75 to BCBL1. We selected all KSHV microRNAs that showed a KSHV-specific activity pattern, i.e., the set of target sites was depleted of reads in DG75 and included reproducible target sites of all three KSHV positive cell lines (Supplemental Figs. S3, S8, S9). Furthermore, all selected KSHV microRNAs are expressed in all KSHV-positive cell lines according to the microRNA reads from the PAR-CLIP experiment (see Supplemental Table S4) and showed a clear pattern of MES as introduced above.

Context-dependent microRNA targets are associated with RISC in a context-dependent manner

First, we looked at the recruitment of the mRNA targets of these KSHV microRNAs to AGO2-complexes. We recently used RIP-chip to identify KSHV and EBV microRNA targets in human B-cells (Dölk et al. 2010). Since then, we performed two additional RIP-chip replicates of the KSHV-positive cell line, BCBL1, to perform a more solid statistical analysis (Erhard et al. 2013b). Data were normalized using principal component analysis as described (Erhard et al. 2013b) and differential enrichment values were computed for BCBL1 and DG75 as the second principal component (PC2), indicating whether an mRNA is more strongly associated with RISC in BCBL1 in comparison to DG75. All PAR-CLIP target sites were mapped to genes; and genes with any KSHV target site in BCBL1, with a constitutive target site in all KSHV-positive cell lines, and with exclusive sites in BCBL1 or BC1/BC3 were compared to all other genes with any PAR-CLIP target site as background (Fig. 3).

The differential RIP-chip enrichment was significantly shifted toward higher values for genes with BCBL1 exclusive sites in comparison to the background (P < 0.0007, Kolmogorov-Smirnov test), indicating that BCBL1-exclusive target sites indeed lead to a stronger association of the target mRNA in BCBL1 to RISC. This was also true for constitutive KSHV target sites (P < 0.0009) as well as for all KSHV target sites active in BCBL1 (P < 3.10^{-8}). However, BC1/BC3-specific target sites, which were not active in BCBL1,
Figure 2. Comparison of PAR-CLIP data sets. (A) The number of target sites observed only in individual cell lines (outermost labeled circles), in two cell lines (circles on the edges between cell lines), and in all three cell lines (center circle), for KSHV microRNA target sites. Relatively few target sites appear to be active in multiple cell lines (see Supplemental Fig. S2A for overlaps of cellular microRNA targets). (B) Summary of all pairwise overlaps for clusters of cellular and viral microRNAs in all data sets. The Jaccard index ($J$) is the number of clusters in the intersection divided by the total number of clusters in any of the two experiments. Jaccard indices of $\sim{70\%}$ for all replicate measurements indicate high reproducibility, whereas comparisons across cell lines show relatively low overlap ($J < 40\%$) (see also Supplemental Fig. S2). (C) The PAR-CLIP read heatmap for target sites of the KSHV microRNA miR-K12-4-3p (see Fig. 1B for more information about PAR-CLIP read heatmaps). Between KSHV-positive cell lines, there is no correlation, but there are distinct clusters of target sites. No obvious dependency between clusters and mRNA expression level is observable.
were indistinguishable from the background (Fig. 3A; Supplemental Table 1). In particular, genes with active KSHV microRNA target sites in BCBL1 showed a twofold enrichment of genes that are significantly (PC2 score > 2) more associated with RISC in BCBL1 than in DG75 over background genes. In contrast, genes with KSHV microRNA target sites that are exclusively active in BC1 or BC3 and not in BCBL1 are indistinguishable from background genes (Fig. 3B).

This provides strong evidence that a major fraction of the KSHV microRNA target sites identified by PAR-CLIP exclusively in BC1/BC3 do not mediate a strong recruitment of their target mRNA to RISC in BCBL1, i.e., they are indeed context-dependent target sites. Context-dependent microRNA/target interactions, as defined by differential analysis of PAR-CLIP data, can thus be confirmed using an independent RIP-chip experiment.

**Target mRNA stability is affected in a context-dependent manner**

Next, we analyzed the context-dependent effects of the KSHV microRNAs on target RNA stability. Since microRNAs can induce destabilization of the mRNA transcripts (Bartel 2009), microRNA/target interactions that are active in BCBL1 should decrease the target mRNA half-life in BCBL1 as compared to DG75. Target sites inactive in BCBL1 (and only active in BC1/BC3) in contrast should not decrease mRNA half-life.

Previously, we applied metabolic labeling of newly transcribed RNA followed by microarray analysis to separate newly synthesized and preexisting RNA (Dölken et al. 2008). We computed RNA half-lives based on the ratios of newly synthesized to total RNA for both DG75 and BCBL1 (Dölken et al. 2010) and considered the differences in target mRNA half-lives in between BCBL1 and DG75.

Intriguingly, the mRNA half-life of KSHV microRNA targets in BCBL1 was decreased by ~20 min ($P < 3.10^{-2}$) on average, whereas for KSHV microRNA targets not active in BCBL1, no significant decrease was observed (Fig. 3C; Supplemental Table 1). Furthermore, the half-life difference values of BCBL1-exclusive target genes were significantly smaller than half-life difference values of BC1- or BC3-exclusive target genes ($P < 0.008$, Wilcoxon rank sum test) (Fig. 3D). Thus, context-dependent microRNA/target interactions impact on mRNA stability in a context-dependent manner.

Interestingly, constitutive KSHV microRNA target sites showed an even stronger decrease in the mRNA half-life than for context-dependent target sites (>35 min on average, $P < 10^{-6}$). A possible explanation is that constitutive microRNA/target interactions are less susceptible to the cellular context, resulting in more substantial target suppression. Therefore, constitutive interactions likely represent the most important targets for the virus.

**Protein levels are differentially regulated for context-dependent microRNA targets**

We now asked whether context-dependent microRNA targets are also reflected in steady-state mRNA or protein levels in two different contexts. It is important to note that protein levels in a cell depend on multiple factors, including protein half-lives and microRNA-independent post-transcriptional regulation, most of which are well described to have a substantially greater impact on protein levels than generally exerted by microRNAs. Therefore, targets of the viral microRNAs may not necessarily show differential expression between DG75 and BCBL1 on protein or mRNA levels (Supplemental Fig. S3; Dölken et al. 2010). Especially viral microRNAs are likely to counteract the cellular response to infection (Cullen 2011; Kincaid and Sullivan 2012), which is reflected by the fact that KSHV microRNAs target multiple induced genes (Dölken et al. 2010).

Indeed, when mRNA or protein levels were considered individually, no significant shift in expression fold changes was observed for any set of microRNA targets (Fig. 3E; Supplemental Fig. S3). Thus, in spite of the fact that mRNA half-lives are significantly decreased by KSHV microRNAs, there is no observable effect on steady-state levels of either mRNAs or proteins. However, if protein fold changes are normalized to mRNA fold changes, a small but statistically significant difference can be observed between BCBL1-specific targets and BC1/BC3-specific targets ($P < 0.01$, Wilcoxon rank sum test) (Fig. 3F). Since this normalization effectively removes all effects of mRNA levels and half-lives, this indicates that KSHV microRNAs not only have an impact on mRNA half-life in a context-dependent manner but also on how many proteins are produced per mRNA molecule. Constitutive targets of KSHV microRNAs did not show this pattern, presumably because of their strong impact on mRNA half-lives (Fig. 3).

Taken together, RIP-chip data, RNA half-life data, as well as mRNA and protein expression data, provide good evidence that a substantial amount of KSHV microRNA target sites as found by differential analysis of PAR-CLIP data is indeed context-dependent, resulting in a differential association with RISC and a context-dependent impact on target gene expression.

**Context-dependent target sites of cellular microRNAs**

We next selected context-dependent microRNA/target interactions that are either active in BCBL1 or DG75 but not in both. Thus, we first selected all microRNAs that are not differentially expressed between BCBL1 and DG75 (less than twofold) and are reliably detected in the PAR-CLIP experiments (at least 100 reads in all four data sets). Furthermore, all microRNAs had to have at least 20 target sites as identified by a 7-mer seed by PARma. All identified microRNAs showed a clear pattern of context dependency in their target sites (Supplemental Figs. S10, S11). Using the same criteria as in the analysis of KSHV microRNAs, context-dependent target sites were defined (Supplemental Fig. S4; see Supplemental Table S5).

Again, context-dependent microRNA/target interactions as defined by the differential PAR-CLIP analysis resulted in highly significant differential association with RISC (Fig. 4A; Supplemental Table 1). Specifically, context-dependent targets are more than twofold enriched in significantly differentially RISC-associated mRNAs (PC2 score > 2) for both cellular contexts. Furthermore, target mRNA half-lives are again significantly lowered by the context-dependent activity of the microRNA/target interactions ($P < 0.00002$, Wilcoxon rank sum test) (Fig. 4B). Thus, as in the analysis of KSHV microRNAs, context-dependent target sites of cellular microRNAs also lead to differential RISC-association and have functional impact on target mRNA half-lives in a context-dependent manner.

The analysis of steady-state expression levels revealed a clear pattern of context-dependent targets: Both sets of context-dependent targets are clearly shifted in comparison to the background with respect to both mRNA and protein fold changes (Fig. 4C,D). Specifically, genes tend to have higher expression in the context where the microRNA/target interactions are active, indicating that a significant number of targets are context-dependent because
Figure 3. (Legend on next page)
they are not expressed or are expressed at low levels in one condition, i.e., affected by the global context.

Importantly, this is not solely due to a completely abrogated expression in the nonactive context since proteins are detected for almost half of all context-dependent targets in both cell lines; and in more than two-thirds of the cases, the fold change is smaller than twofold (Fig. 4E). Thus, it is not the absence or presence of target mRNAs that lead to context dependency of target sites. Rather, this indicates a complex dependency of the target site activity on the exact target mRNA expression levels. However, there may be a subpopulation within both sets of context-dependent targets, where a missing activity of a target site may be explained by the complete absence of the target mRNA (Fig. 4F).

mRNA levels and flanking sequence motifs explain context-dependent microRNA/target interactions

Thus, we analyzed to which extent mRNA expression levels contribute to the cellular context and whether there are other factors that are necessary to explain the widespread context dependency of target sites. First, we tested whether the target mRNA level is the only contributor that constitutes the cellular context for microRNA-mediated gene regulation.

Read counts are not only subject to biological variance but also to a substantial amount of sampling noise since many clusters only have a few dozen reads. To compare PAR-CLIP read count fold changes with mRNA fold changes in a more robust manner, it is therefore important to estimate the extent of this sampling noise. We used a population-based estimate of variance using a conditional gamma distribution (see Supplemental Methods). This approach is similar to recent methods to estimate the significance of differential expression in RNA-seq data (Anders and Huber 2010; Robinson et al. 2010). Importantly, our noise model nicely reflects variation observed in replicate experiments (Supplemental Fig. S6).

If this noise model is applied to the comparison of mRNA fold change-corrected PAR-CLIP target sites, >50% of all context-dependent target sites, i.e., at least 14% of all target sites of the selected set of cellular microRNAs, cannot be explained as judged by the F-value distribution (Fig. 5). This means that in these cases, the PAR-CLIP read count fold change is significantly higher than expected from the corresponding mRNA fold change, and this difference also cannot be explained by sampling noise inherent to low-count data such as PAR-CLIP. Thus, target site activities are not simply linearly dependent on mRNA levels.

Furthermore, as illustrated in Figure 5B, there are several instances in which the target gene is not differentially expressed (i.e., data points around zero on the mRNA log2 fold change axis) but where the target sites show a >16-fold elevated activity. In these cases, mRNA expression alone clearly cannot explain target site activity. Thus, other factors contribute to context-specific microRNA function.

RNA binding proteins (RBPs) likely constitute such additional contributors. Thus, we performed a motif search in regions flanking context-dependent target sites (seed site ±80 bp). For motif discovery, we used MERCI (Vens et al. 2011), which is based on efficiently enumerating all discriminative k-mers of two sets of sequences. Specifically, we searched for k-mers that do not occur in the negative set and occur at least n times in the positive set, and we only considered target sites from mRNAs that are not differentially expressed. n was chosen according to the total number of sequences in the positive set. MERCI identified 20–30 k-mers when we compared target sites of cellular microRNAs exclusively present in BCBL1 to those exclusively present in DG75 and target sites of viral microRNAs exclusively present in BCBL1 to those in BC1/BC3 or vice versa (Fig. 6A; Table 1; Supplemental Table S3). These discriminative k-mers occur in 75%–90% of all context-dependent target sites that cannot be explained by the mRNA level, and as few as five motifs already can explain 30%–40% of all sites. In contrast, discriminative k-mers found by chance in randomized sequences only occur in a considerably lower number of sequences (Supplemental Fig. S7). Thus, these motifs are likely candidates of binding sites for RBPs contributing to context-dependent recognition of target sites by microRNAs.

In summary, from all context-dependent target sites identified by PAR-CLIP and validated by RIP-chip experiments, 4S U tagging-based mRNA half-lives, and mRNA and protein expression measurements, >90% can either be explained by differential mRNA levels or by the presence of a putative RBP binding motif (Fig. 6B).

Context-dependent target sites are less conserved than constitutive sites

Finally, we asked whether context-dependent target sites have distinct evolutionary conservation patterns as compared to constitutive target sites. Following the approach of Friedman et al. (2009), for each target site we computed the branch length along the phylogenetic tree of 46 vertebrates by summing all branches in which the seed of a cluster is fully conserved in the genome-wide multiple alignment of 46 vertebrate species. The branch length thus incorporates both the evolutionary age as well as the loss of a target site in specific lineages. Specifically, a target site that emerged in the last common ancestor of primates and rodents and has not been lost in any primate or rodent lineage, has a branch length of 2.342 (Fig. 7, shaded areas).

Figure 3. PAR-CLIP targets in RIP-chip experiments, mRNA half-life measurements, and expression measurements. (A) Differential RIP-chip enrichment scores (PC2 scores; positive values indicate higher enrichment in BCBL1 than in DG75). Generally, KSHV microRNA targets active in BCBL1 are significantly shifted toward higher values as compared to all other genes with any PAR-CLIP target site, in contrast to KSHV target sites exclusively active in BC1 or BC3 and not in BCBL1. Illustrates this further: The enrichment of genes with any KSHV site, with a constitutive or a BCBL1 exclusive site over genes with BC1/BC3 exclusive sites among all genes with PC2 score > 2 is about twofold in all cases. (C) Distributions of half-life differences between BCBL1 and DG75 for all genes with PAR-CLIP target sites. Thus, positive values indicate a longer mRNA half-life in BCBL1 than in DG75. Genes with KSHV microRNA targets active in BCBL1 tend to have shorter half-lives in BCBL1 than in DG75. This is highly significant for all BCBL1 target genes as well as the constitutive targets but not for BCBL1-specific targets, even if their half-life is on average ~ 20 min shorter in BCBL1 than in DG75. However, KSHV microRNA targets that are inactive in BCBL1 do not show any shift in their half-lives. (D) The difference between targets active exclusively in BCBL1 is statistically significantly different from targets active exclusively in BC1 or BC3, when their ranks among all PAR-CLIP targets are considered. (E) Genes are scattered according to their mRNA log2 fold changes between BCBL1 and DG75 on the x-axis and to their protein log2 fold changes on the y-axis. In both dimensions, none of the KSHV target sets is significantly down-regulated on either the mRNA or protein level (Supplemental Fig. S5). However, target sites active in BCBL1 appear to be shifted toward the bottom right. These sites correspond to genes whose protein level fold change between BCBL1 and DG75 is lower than expected from the mRNA level. (F) The ranks of protein fold changes normalized to their mRNA levels for all gene sets considered. Normalized protein fold changes are significantly lower for genes with BCBL1-specific target sites than for genes with target sites inactive in BCBL1 (P < 0.01, Wilcoxon rank sum test) (see also Supplemental Fig. S3).
Intriguingly, constitutive target sites of conserved cellular microRNAs are significantly more strongly conserved than context-dependent sites ($P < 0.003$, two-sided Kolmogorov-Smirnov test). For instance, although $>80\%$ of constitutive sites are conserved beyond the last common ancestor of primates and rodents, only $65\%$ of context-dependent sites are conserved beyond this clade; and $20\%$ of context-dependent sites even show a signature of recent evolution within the primate lineage. Importantly, this does not reflect the overall conservation level of the respective 3'-UTRs, but is specific to the seed sites (Supplemental Fig. S5).

Target sites of viral microRNAs, independent of whether they are context-dependent or constitutive, show patterns of much weaker conservation. This can be expected because there are no conserved viral microRNAs, and pathogenicity of KSHV may rather induce positive selection of its microRNA target sites on host mRNAs.

**Discussion**

In this study, we analyzed PAR-CLIP data from four human B-cell lines, three of which are infected with Kaposi’s sarcoma-associated herpes virus (KSHV), using an improved computational approach to identify target sites of both cellular and viral microRNAs (PARma) (Erhard et al. 2013a). The overlap in target sites between the four cell lines was surprisingly low (~40\%), indicating a large set of context-dependent microRNA/target interactions. Three additional sets of high-throughput data (RIP-chip, 4sU-tagging-derived RNA half-lives, and SILAC proteomics data) supported this observation: Context-dependent microRNA targets are associated with RISC in a context-dependent manner and have a measurable functional impact on their targets in a context-dependent manner. This was observed for the targets of both cellular and viral microRNAs. The latter offered an important control as they were exclusively expressed in the cells expressing the viral microRNAs. Thus, we propose a new layer of complexity in microRNA targeting: Depending on the cellular context, specific microRNA/target interactions may be active or not, even if both microRNA and target mRNA are expressed. Furthermore, we could show that the evolutionary conservation differs between context-dependent and constitutive target sites, indicating that selective pressure may be different for context-dependent and constitutive target sites or that they have different evolutionary ages.

Cellular context may be formed directly by the quantities of microRNAs and mRNAs. Dependent on the exact copy numbers of microRNAs and mRNAs in each cell, intricate regulatory mechanisms may emerge leading to highly complex patterns of regulation (Mukherji et al. 2011). Furthermore, due to the many-to-many nature of regulatory蛋白 interactions and targets, microRNAs and mRNAs are embedded in a highly complex regulatory network (Hobert 2008). Our analyses indicate that the quantities of microRNAs and target mRNAs are direct contributors to the cellular context. However, based on our results, >50\% of all observed context-dependent microRNA/target interactions cannot be explained by microRNA or mRNA levels and, therefore, are likely dependent on indirect factors.

Differential isoform usage might constitute relevant contexts. It may lead to target sites that are active in one condition and inactive in another, even if the corresponding microRNA is constitutively expressed and the target appears constitutively expressed on the gene level. Especially differential usage of polyadenylation sites has been shown to be a major contributor to the cellular context during development and proliferation (Sandberg et al. 2008; Boutet et al. 2012). However, in our data sets, alternative polyadenylation sites cannot be distinguished, and more involved experiments would be necessary to identify context-dependent target sites due to possible alternative polyadenylation.

The presence of RNA binding proteins (RBPs) may prevent microRNA binding to nearby sites (Bhattacharyya et al. 2006) or also induce binding (Kim et al. 2009). In a recent study, the whole RNA binding proteome of a cell line was examined by PAR-CLIP coupled to high resolution mass spectrometry (Baltz et al. 2012). This study revealed two important aspects of RBPs: First, in a single cell type, about 800 different RBPs can be identified. This unexpectedly high number of RBPs allows for highly complex combinatorics of competitive or activating RBP–microRNA interactions. Second, crosslinking events were observed for almost 30\% of all uridines in 3'-UTRs, suggesting that mRNAs are broadly covered by RBPs. Indeed, we could identify a handful of sequence motifs that are able to explain a large fraction of context-dependent target sites, indicating that RBPs may play important roles in shaping the cellular context for microRNA-mediated regulation.

Thus, there is an intriguing analogy of the transcriptional and post-transcriptional layer of regulation: DNA, which is the material for transcriptional regulation, is covered by histones, transcription factors, and other DNA binding proteins; and the composition and dynamics of these proteins contribute to the cellular context (The ENCODE Project Consortium 2012). This cellular context determines to which extent a certain transcription factor can bind to a specific target site and exerts its regulatory role. Context-dependent regulatory networks may differ dramatically across different cell types or conditions (Neph et al. 2012). Similarly, mRNAs, which are the units for post-transcriptional regulation, are covered by RBPs, and we argue that their composition and dynamics contribute to a cellular context for microRNA-mediated regulation. Additionally, factors other than these covering proteins...
may further shape the cellular context for both transcriptional and post-transcriptional regulation: For transcriptional regulation, distinct modifications of chromatin or the DNA may also determine context. Furthermore, chromosomal conformations may place distal binding sites of transcription factors to promoters of different genes in three-dimensional space and may therefore also be important.

mRNAs may even provide more opportunities for context-dependent regulation: Although DNA usually is restricted to a single cellular compartment, the nucleus, the life cycle of mRNAs may span multiple compartments and subcompartments. This cellular localization may itself be regulated, and depending on the localization, mRNAs may be translated or not. For instance, sequestering of mRNAs to P-bodies by microRNAs leads to a reduced

Figure 5. Comparison of mRNA fold changes to PAR-CLIP read count fold changes. (A) Scatterplot comparing mRNA fold changes to PAR-CLIP read count fold changes of all target sites of the cellular microRNAs analyzed. For the PAR-CLIP data, a pseudocount of 1 was used. Green dots represent target sites that can be explained by the mRNA fold change while respecting sampling noise of the read counts, whereas orange and red dots correspond to significant outliers (\(P < 0.05\) and \(P < 0.01\), respectively). The \(P\)-value distribution in C of all these target sites suggests that at least 14.9% (363 instances with \(P < 0.01\) of overall 2436 target sites after subtraction of baseline indicated by the horizontal line) of all differential target site activities cannot be explained by the mRNA fold change and sampling noise. B and D illustrate this for the context-dependent microRNA/target interactions only. Here, >50% of all sites cannot be explained by mRNA levels (see also Supplemental Fig. S5).
translation and mRNA decay (Pasquinelli 2012). Furthermore, the single-stranded mRNA gives rise to complex secondary and tertiary structures, and it has been shown that the accessibility of target sites determines whether or not microRNAs can bind to the mRNA (Kertesz et al. 2007). Interestingly, the conformation of RNAs is highly flexible and may be reshaped in a context-dependent way: Kedde et al. (2010) have shown that the activation of the RNA binding protein, PUM1, induces a local change in a hairpin structure of the 3′-UTR of the CDKN1B mRNA. Upon PUM1 activation, an inaccessible binding site of miR-221/miR-222 is opened for binding, leading to an efficient repression of CDKN1B.

The differential analysis of a collection of high-quality large-scale experiments for microRNA target site discovery indicates that context-dependent microRNA targeting is not restricted to a few examples, but is a widespread phenomenon and a general feature of microRNA-mediated regulation. Under natural conditions, several factors may contribute to the cellular context that may determine which target sites of a microRNA are active and which are not. Thus, true target sites discovered in an experiment performed in a particular cell line may not be active in a different context. In addition, experiments that impose non-natural conditions on cells may identify target sites that are active only in such a non-natural context and at the same time may miss naturally occurring target sites.

Thus, additional experiments to unravel context-dependent microRNA targets are of great importance for both the identification of microRNA targets in particular contexts as well as the investigation of key contributors that determine cellular context.

Methods

Cell lines
DG75-eGFP (named DG75 throughout this paper) and BCBL1 were cultured in RPMI medium supplemented with 10% fetal calf serum and pen/strep.

PAR-CLIP and sequencing
PAR-CLIP on DG75 and BCBL1 was performed by the Zavolan laboratory as described (Kishore et al. 2011; Jaskiewicz et al. 2012), and the data are available at GEO (accession number: GSE43909; see Supplemental Table S6 for sequencing statistics). The PAR-CLIP sequencing data for BC1 and BC3 from Gottwein et al. (2011) have been downloaded from GEO (accession number: GSE32113). We applied PARMa to the whole collection of all PAR-CLIP data sets as described (Erhard et al. 2013a).

SILAC-based proteomics
SILAC and LC-MS/MS were performed as described in the Mann laboratory at MPI.
Figure 7. Conservation of target sites. Distributions of branch lengths of target sites are illustrated (see main text for a definition of branch lengths). Shaded regions indicate the maximal branch lengths of target sites conserved in primates, in primates and rodents, in mammals, and in vertebrates. All cellular microRNAs considered here are conserved in vertebrates. Constitutive target sites of these microRNAs are significantly more conserved (P < 0.0004, two-sided Kolmogorov-Smirnov test) than context-dependent target sites. Moreover, neither context-dependent nor constitutive target sites of viral microRNAs show evidence for evolutionary conservation (see also Supplemental Fig. S5).

Acknowledgments

for Biochemistry in Munich. The raw files from the mass spectrometer have been analyzed using MaxQuant (version 1.2.2.5) (Cox and Mann 2008), using standard parameters against all human proteins from Ensembl (v60).

RIP-chip analysis

For the RISC-IPs, $5 \times 10^8$ cells were taken for each replicate and processed as previously described (Dölken et al. 2010) using 6 μg purified monoclonal hAGO2 antibody (a-hAGO2; 11A9) or monoclonal BrdU-antibody (Abcam; used as control).

RNA half-life measurements by 4SU-tagging

The RNA half-life data for DG75 and BCBL1 have been published previously (Dölken et al. 2010). In brief, newly transcribed RNA was labeled for 1 h by adding 100 μM 4SU to the cell culture medium. Total RNA was prepared using TRIzol, and newly transcribed RNA was purified as described (Dölken et al. 2008). Three replicates of newly transcribed, total, and preexisting RNA were measured.

References


Author contributions: F.E. carried out the analyses, interpreted data, and wrote the manuscript. D.L. and G.M. carried out experiments. J.H. supervised the experiments. L.D. designed the experiments, contributed ideas, and helped to draft the manuscript. All authors read and approved the final manuscript.

and M.Z. carried out the PAR-CLIP experiments. R.Z. supervised the project and helped to draft the manuscript. All authors read and approved the final manuscript.


Received September 13, 2013; accepted in revised form March 24, 2014.