Cell Reports
Resource

An Optimized microRNA Backbone for Effective Single-Copy RNAi

Christof Fellmann,[1,*] Thomas Hoffmann,[2] Vaishali Sridhar,[1] Barbara Hopfgartner,[2] Matthias Muhar,[2] Mareike Roth,[2] Dan Yu Lai,[1] Inês A.M. Barbosa,[2] Jung Shick Kwon,[1] Yuanzhe Guan,[1] Nishi Sinha,[1] and Johannes Zuber[2,*]
[1]Mirimus Inc., 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA
[2]Research Institute of Molecular Pathology (IMP), Dr. Bohr-Gasse 7, 1030 Vienna, Austria
*Correspondence: fellmann@mirimus.com (C.F.), zuber@imp.ac.at (J.Z.)
http://dx.doi.org/10.1016/j.celrep.2013.11.020

SUMMARY

Short hairpin RNA (shRNA) technology enables stable and regulated gene repression. For establishing experimentally versatile RNAi tools and minimizing toxicities, synthetic shRNAs can be embedded into endogenous microRNA contexts. However, due to our incomplete understanding of microRNA biogenesis, such "shRNAmirs" often fail to trigger potent knockdown, especially when expressed from a single genomic copy. Following recent advances in design of synthetic shRNAmir stems, here we take a systematic approach to optimize the experimental miR-30 backbone. Among several favorable features, we identify a conserved element 3′ of the basal stem as critically required for optimal shRNAmir processing and implement it in an optimized backbone termed "miR-E", which strongly increases mature shRNA levels and knockdown efficacy. Existing miR-30 reagents can be easily converted to miR-E, and its combination with up-to-date design rules establishes a validated and accessible platform for generating effective single-copy shRNA libraries that will facilitate the functional annotation of the genome.

INTRODUCTION

RNAi is an evolutionarily conserved mechanism of posttranscriptional gene regulation that uses small RNAs produced through multicomplex machinery to guide suppression of complementary transcripts. The most well-characterized endogenous RNAi triggers, microRNAs (miRNAs), are expressed as hairpin-like structures in primary transcripts (pri-miRNAs), recognized and released as precursor miRNAs (pre-miRNAs) by the Drosha/DGCR8 microprocessor complex, and further processed in the cytoplasm by the Dicer/TRBP complex into mature small RNA duplexes, of which one strand serves as guide to downregulate complementary transcripts (for review, see Bartel, 2004; Carthew and Sontheimer, 2009; Filipowicz et al., 2008). Simultaneous to the discovery and ongoing exploration of miRNA pathways, synthetic RNAi triggers have been developed to experimentally program the RNAi machinery for loss-of-function genetics. Beyond transfection of small interfering RNAs (siRNAs), providing an efficient approach for transient gene knockdown, short hairpin RNAs (shRNAs) can be expressed from DNA vectors that enable stable and regulated loss-of-function studies in a gene-by-gene or pool-based format (for review, see Hannon and Rossi, 2004; Mohr and Perrimon, 2012).

Among numerous existing shRNA reagents (for review, see Hu and Luo, 2012; Pan et al., 2012), two basic expression systems need to be distinguished. Simple stem-loop shRNAs transcribed from Pol-III promoters enter the processing machinery at the level of Dicer and can be effective RNAi triggers. However, their enforced expression can saturate endogenous miRNA pathways and result in severe toxicities (Grimm et al., 2006). Moreover, recent studies reveal that Dicer is imprecise in processing commonly used stem-loop designs, which increases the likelihood of aberrant guide- and passenger-strand mediated off-target effects (Gu et al., 2012). An alternative approach is based on embedment of synthetic shRNA stems into the context of endogenous miRNAs. The resulting "shRNAmir" structures serve as natural substrates in miRNA biogenesis pathways and can trigger potent knockdown, as has been demonstrated for a number of miRNA backbones, including miR-30, miR-155, and miR17-92 (Chung et al., 2006; Liu et al., 2008; Zeng et al., 2002).

Although studies comparing knockdown efficacies of simple stem-loop and miRNA-embedded shRNAs have reported conflicting results (Boden et al., 2004; Boudreau et al., 2008; Silva et al., 2005), shRNAmir-based systems offer several technical advantages: (1) Like endogenous miRNAs, shRNAmirs can be expressed from Pol-II promoters, which has enabled tissue-specific RNAi studies (Hinterberger et al., 2010) as well as the development of robust tetracycline (Tet)-regulated RNAi systems (Dickins et al., 2005; Stegmeier et al., 2005); (2) in Pol-II transcripts, shRNAmirs can be placed in the 3′ UTR of a reporter to directly monitor shRNA expression (Premsrirut et al., 2011; Stegmeier et al., 2005; Zuber et al., 2011a); (3) like miRNAs, multiple shRNAmirs can be expressed as a polycistron, providing a setup for combinatorial RNAi studies (Chung et al., 2006; Zhu et al., 2007); (4) shRNAmirs are less prone to cause toxicities by interfering with endogenous miRNA pathways (Boudreau

et al., 2009; Castanotto et al., 2007; McBride et al., 2008; Premsrirut et al., 2011); and (5) the natural loop configuration of shRNAmirs such as miR-30 complies with a recently discovered ''loop-counting rule'' (Gu et al., 2012) that ensures precise Dicer cleavage and reduces off-target effects.

The key limitation of existing shRNAmir reagents is our inability to predict shRNAs that trigger potent target knockdown (>80% reduction in protein expression) when expressed from a single genomic integration. Although such single-copy efficiency is a critical prerequisite for various key applications such as pool-based RNAi screening and transgenic RNAi, shRNA reagents are typically untested at this stringency level. In fact, until recently, shRNAmir design strategies were mainly based on siRNA data sets, and many previous studies evaluating shRNAmir constructs were solely based on plasmid DNA transfection (Chung et al., 2006; Silva et al., 2005; Zeng et al., 2002), which fails to reflect the relevant single-copy setting (Fellmann et al., 2011). At the same time, our still limited understanding of structural requirements for effective miRNA biogenesis hampers the rational optimization of shRNAmir reagents.

In an effort to functionally identify highly effective miR-30-based shRNAs, and systematically explore shRNA processing requirements, we have recently established an RNAi Sensor assay able to measure knockdown potency of thousands of shRNAs at single-copy conditions in a pooled format (Fellmann et al., 2011). Our study revealed that potent single-copy shRNAs are rare and share distinct sequence features that can be associated with specific miRNA biogenesis steps. Notably, the most potent shRNAs were also characterized by a marked bias for incorporating the guide strand, suggesting that the use of these shRNAs will not only boost on-target knockdown levels, but also reduce passenger-mediated off-target effects. Beyond these advances in identifying single-copy effective shRNAs, results from this and a parallel study (Premsrirut et al., 2011) indicated that single-copy expression of even the most potent shRNAs results in mature small RNA levels well below the most abundant endogenous miRNAs, raising the possibility that miR-30-based RNAi reagents could be further improved by optimizing the miRNA scaffold. To explore this option, here we systematically evaluate how conserved sequence features in the miR-30 backbone impact the potency of shRNAmirs under strict single-copy conditions. Among several improved design variants, the most potent (termed ''miR-E'') incorporates a conserved sequence element 3′ of the basal stem and can easily be implemented in existing miR-30 reagents to globally enhance mature small RNA production and knockdown potency.

RESULTS

Although many key players in miRNA biogenesis are known, our understanding of substrate RNA sequence features required for effective pri- and pre-miRNA processing is still limited (Chiang et al., 2010), hampering the rational optimization of shRNAmir backbones. Because conserved sequence elements in pri-miRNAs can impact miRNA biogenesis (Feng et al., 2011; Han et al., 2006), we examined whether any such elements have been altered in the experimental miR-30 backbone. Compared to the natural human MIR30A, the common miR-30 backbone differs in three potentially relevant features (Figures 1A and S1A): (1) The synthetic miR-30 stem has no bulge and harbors the intended guide on the opposite strand; (2) for designing the 3′ overhangs of synthetic shRNAs, two conserved base pairs flanking the loop were changed from CU/GG to UA/UA; and (3) the introduction of XhoI/EcoRI restriction sites for shRNA cloning led to a mutation in a highly conserved region 3′ of the basal stem.

To investigate the impact of these alterations on shRNAmir knockdown efficiency, we designed 11 miR-30 variants that revert each feature individually or in combination back to the endogenous MIR30A configuration (Figure S1B). As experimental shRNA for these studies, we selected a previously characterized intermediate Pten shRNA (Pten.1524, 62% protein knockdown at single copy) (Fellmann et al., 2011), whose target site lies directly adjacent to that of a highly potent Pten shRNA (Pten.1523, 96% knockdown), which makes it unlikely that the limited potency of Pten.1524 is due to target site inaccessibility (Ameres et al., 2007). To test the knockdown potential of different backbone variants harboring Pten.1524, we employed a newly developed fluorescence-based reporter assay that enables accurate quantification of single-copy protein knockdown levels of up to 20 experimental and several control shRNAs side by side (Figures S1C–S1F).

Compared to potent control shRNAs, Pten.1524 in the common miR-30 configuration produced intermediate knockdown levels (Figure 1B). Although several miR-30 variants involving bulges and reversed stem configurations had no or a negative impact, four designs significantly enhanced reporter knockdown ($p < 0.001$). Strikingly, two of the tested alternative backbones increased the knockdown potency of Pten.1524 to levels in the range of our most potent control shRNAs. One of them, miR-R1B1, was designed to closely resemble the stem-loop structure of the endogenous MIR30A, whereas the other configuration, termed ''miR-E'', restores conserved sequence features 3′ of the basal stem by repositioning the EcoRI cloning site to a nonconserved region further 3′ and modifying the 5′ miR-30 scaffold opposing EcoRI in a way that minimizes changes to the overall secondary structure (Figures S1B and S1G).

Although both miR-R1B1 and miR-E boosted knockdown to a similar degree, we decided to focus our subsequent analyses on miR-E, because existing miR-30 reagents and methodology can be easily adapted to this design, whereas the implementation of miR-R1B1 would require de novo shRNA cloning and revision of many established protocols. Moreover, in a reversed passenger-guide configuration, the critical 5′ end of the guide strand would not be determined by Dicer, which is known to cleave miR-30 at the intended site with 98% accuracy (Gu et al., 2012), but by the microprocessor complex, which appeared to be less precise in our previous analyses (Fellmann et al., 2011).

To confirm the strong effects of miR-E, we transduced NIH 3T3 fibroblasts under single-copy conditions with our potent and intermediate Pten shRNAs in conventional miR-30, miR-E, and an ineffective reversed stem configuration and analyzed Pten knockdown using western blotting (Figure 1C). Again, the miR-E design dramatically improved the potency of Pten.1524, and even for Pten.1523 enhanced target knockdown. To test whether the miR-E design could be further improved, we

Cell PRESS
Open ACCESS

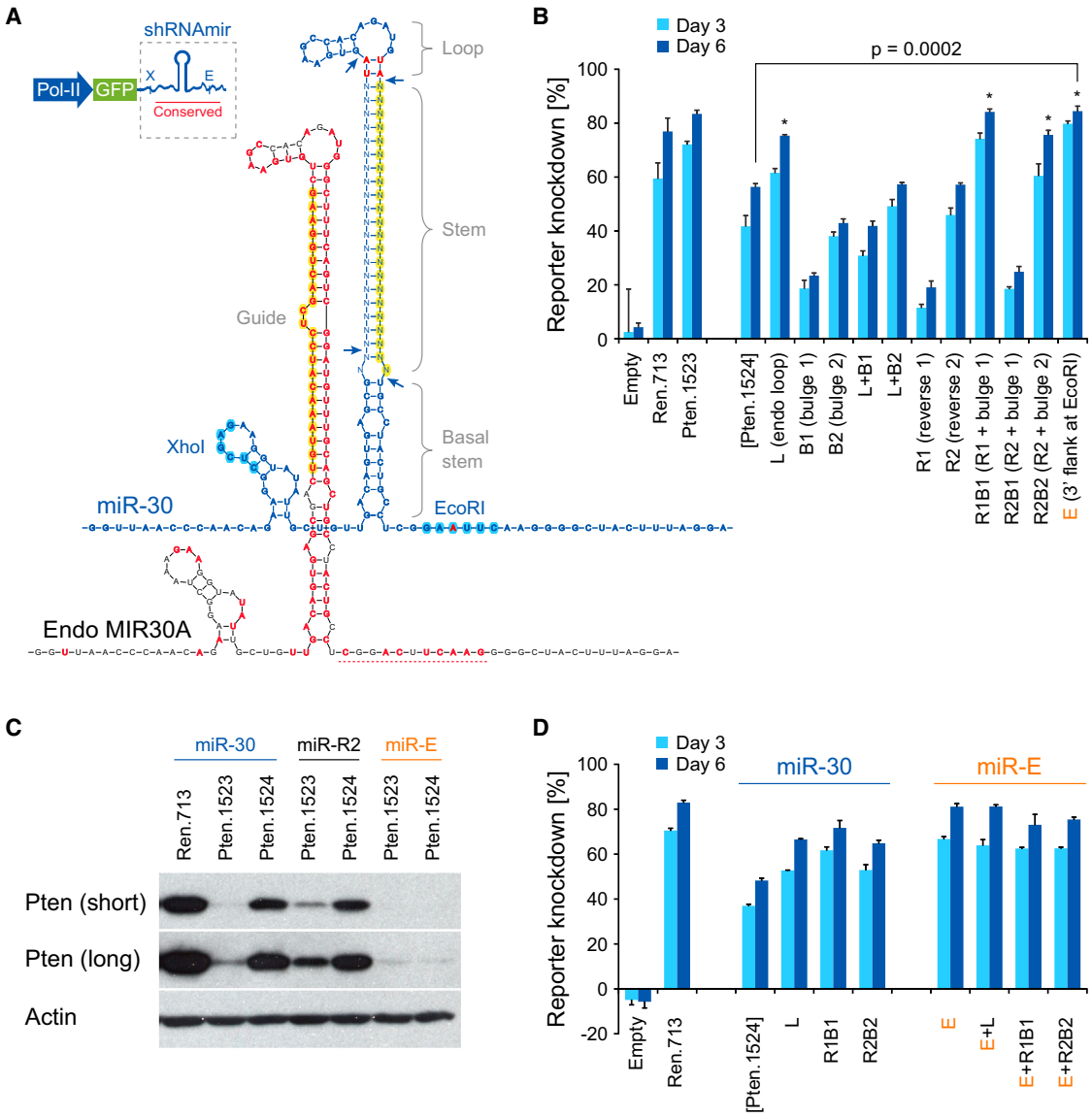

Figure 1. Evaluation of Structural Variants of the Common miR-30 shRNAmir Backbone

(A) Comparison of sequence and predicted structure of the endogenous human MIR30A (black) and the experimental miR-30 backbone (blue). In MIR30A, conserved nucleotides (see Figure S1A for details) are printed in dark red, the guide strand (miR-30a-5p) is highlighted in yellow, and a conserved region 3′ of the basal stem is underlined. In the experimental miR-30, variable target dependent nucleotides are shown as "N", the guide strand is highlighted in yellow, restriction sites used for shRNA cloning are highlighted in blue, and all conserved nucleotides that are altered compared to MIR30A are printed in red. Arrows indicate canonical Drosha and Dicer cleavage sites.

(B) Reporter-based evaluation of shRNAmir backbone variants. Reporter cells expressing dTomato tagged with target sites of the probed shRNAs were transduced at single copy with LMN vectors expressing the indicated shRNAs. All tested backbone variants (Figure S1B) contain Pten.1524. dTomato fluorescence intensity of shRNA-expressing cells was quantified at the indicated time points (see Figures S1C–S1F for details). Values represent means of biological triplicates; error bars represent the SEM. Asterisks, backbone variants that show a highly significant ($p < 0.001$) increase in knockdown potency at day 6 compared to miR-30 Pten.1524.

(C) Immunoblotting for Pten in NIH 3T3 fibroblasts expressing the indicated shRNAs from miR-30, miR-R2, or miR-E under single-copy conditions (see Figure S1G for details). Short and long exposures are shown; miR30-based Ren.713 served as negative control.

(D) Reporter-based evaluation of backbone features in combination. Stem-loop features resulting in significantly improved knockdown (L, R1B1 and R2B2 in B) were tested side by side in the conventional miR-30 backbone or with the miR-E feature. Values are means of biological triplicates; error bars represent the SEM.

See also Table S1 and Figure S1.

constructed miR-E variants harboring the other three modifications found to increase miR-30 knockdown in our initial analysis (Figures 1B and S1B). Although all three modifications again enhanced knockdown of miR-30-based Pten.1524, none of them was able to further improve the potent effects of the miR-E design (Figure 1D). Together, these results illustrate that the effectiveness of miRNA-based shRNAs can be strongly enhanced by optimizing the shRNAmir backbone and reveal one particularly powerful miR-30 design variant that can easily be implemented in established reagents.

Following these initial studies, we explored whether the miR-E design would enhance the knockdown potency of other shRNAs and in other contexts. To this end, we converted a panel of previously tested Mcl1, Bcl2, and Pten shRNAs of different strengths to the new format and evaluated their potency in miR-30 and miR-E side by side (Figures 2A–2C). In almost all cases, the miR-E design resulted in a marked increase in single-copy knockdown and (similar to Pten.1524) turned various intermediate shRNAs into potent RNAi triggers. Beyond studies in fibroblasts, we also evaluated our panel of Pten shRNAs in other murine cell types, including immortalized hepatocytes (BNL CL.2), endothelial cells (SVEC4-10), and kidney cells (TCMK-1) (Figure 2C). The knockdown improvements associated with miR-E were comparable in all four cell lines, demonstrating that these effects are not restricted to specific cell contexts. To explore whether these effects extend to non-mammalian vertebrates, we also tested our two Pten shRNAs using a reporter assay in chicken embryonic fibroblasts and observed similar improvements with miR-E (Figure S2A).

Although these data illustrate that miR-E can result in dramatic improvements of intermediate shRNAs, it remained less clear how substantial these effects would be in case of already validated potent miR-30 shRNAs. Because in these cases protein knockdown quantitation may underestimate the biological importance of improvements in shRNAmir design, we explored this question using functional assays. Specifically, we generated miR-E versions of seven validated shRNAs targeting genes known to be required for survival of MOLM-13 acute myeloid leukemia (AML) cells (Zuber et al., 2011b, 2011c) and tested their effects in competitive proliferation assays by monitoring the relative fraction of cells expressing the shRNA-linked fluorescence reporter (Figures 2D, S2B, and S2C). Strikingly, use of the miR-E backbone led to more rapid depletion of shRNA-expressing AML cells in all seven cases, demonstrating that even established potent miR-30 shRNAs can be enhanced through conversion to the miR-E design. At the same time, an established potent control shRNA targeting Renilla Luciferase (Ren.713) (Zuber et al., 2011a) had no effects in miR-30 or miR-E (Figures 2E and S2C), indicating that use of the miR-E backbone does not result in general toxicities.

To explore whether miR-E would generally improve the single-copy potency of shRNAs cloned de novo based on up-to-date design rules (Dow et al., 2012; Fellmann et al., 2011), we tested 11 predictions targeting Dnmt3a in miR-30 and miR-E using our reporter assay. Although in miR-30 these new designs produced variable knockdown effects and only two highly potent shRNAs in the range of Ren.713, the miR-E format improved knockdown potency and triggered >80% single-copy knockdown in all 11 cases (Figures 2F and S2D–S2F). Similar results have been observed for more than 50 shRNAs targeting other genes (data not shown), whereas we have yet to encounter a single case where implementing the miR-E design has a negative impact. Of note, several of the enhanced shRNAs (i.e., Mcl1.1334, Mcl1.1792, Mcl1.2018, MYC.1834) have previously been identified through systematic testing of every possible target site (Fellmann et al., 2011), illustrating that miR-E can even boost knockdown of the most potent miR-30 shRNAs. Together, these results validate miR-E as a superior backbone for single-copy miRNA-based RNAi.

Although miR-E differs from the previous miR-30 design in several nucleotides (Figure S1G), we hypothesized that restoring the conserved region 3′ of the basal stem (Figure 1A) provided the decisive feature for its improved function. To explore this region in more detail, we constructed miR-E-Pten.1524 variants carrying single point mutations in each of eight consecutive nucleotides in this region and tested their single-copy knockdown efficiency using two independent assays (Figures 3A and 3B). Point mutations in three conserved nucleotides severely impaired the knockdown potency of miR-E to or even below levels observed using the miR-30 design. Mutations in three additional positions resulted in a potency reduction that was only detected by western blot analysis, whereas the only two positions completely unaffected by mutation were nonconserved. Together, these results reveal that the potent effects of miR-E crucially depend on the presence of a conserved -ACNNC- motif 3′ of the basal stem, whereas the other three conserved nucleotides (adding up to an -ACNUCAA- motif) may also contribute to its optimal function. Interestingly, endogenous miRNAs show a propensity for ACNNC/GCNNC/UCNNC motifs in the 3′ flank (Figure 3C), and the presence of these motifs correlates with high mature miRNA levels (Figure S3B), suggesting that a 3′ (non-C)CNNC motif is a common miRNA feature associated with optimal miRNA biogenesis.

Besides the position of this feature in the 3′ backbone flank, two observations strongly suggested that miR-E results in increased knockdown levels due to enhanced pri-miRNA processing. First, compared to miR-30, single-copy expression of transcripts encoding GFP and miR-E shRNAs resulted in significantly lower GFP fluorescence levels (Figure 3D), indicating that a larger fraction of these transcripts are recognized and processed as pri-miRNAs and thus no longer available for reporter protein expression (Figure 3E). Second, presumably for similar reasons, packaging of miR-E containing retroviruses yields significantly lower virus titers, leading to reduced infection efficiencies (Figure S3C). To test this hypothesis, we performed cotransfection assays using siRNAs targeting Microprocessor components and a dual-color reporter vector expressing GFP-coupled miR-30 or miR-E shRNAs alongside a secondary mCherry transcript to control for transfection efficiency. Again, use of miR-E resulted in strongly reduced GFP expression levels compared to miR-30, whereas in both cases siRNA-mediated suppression of DROSHA or DGCR8 resulted in a significant increase in GFP expression to similar levels (Figure S3D). Together, these data indicate that the superior knockdown potential of miR-E-based shRNAs is due to enhanced pri-miRNA

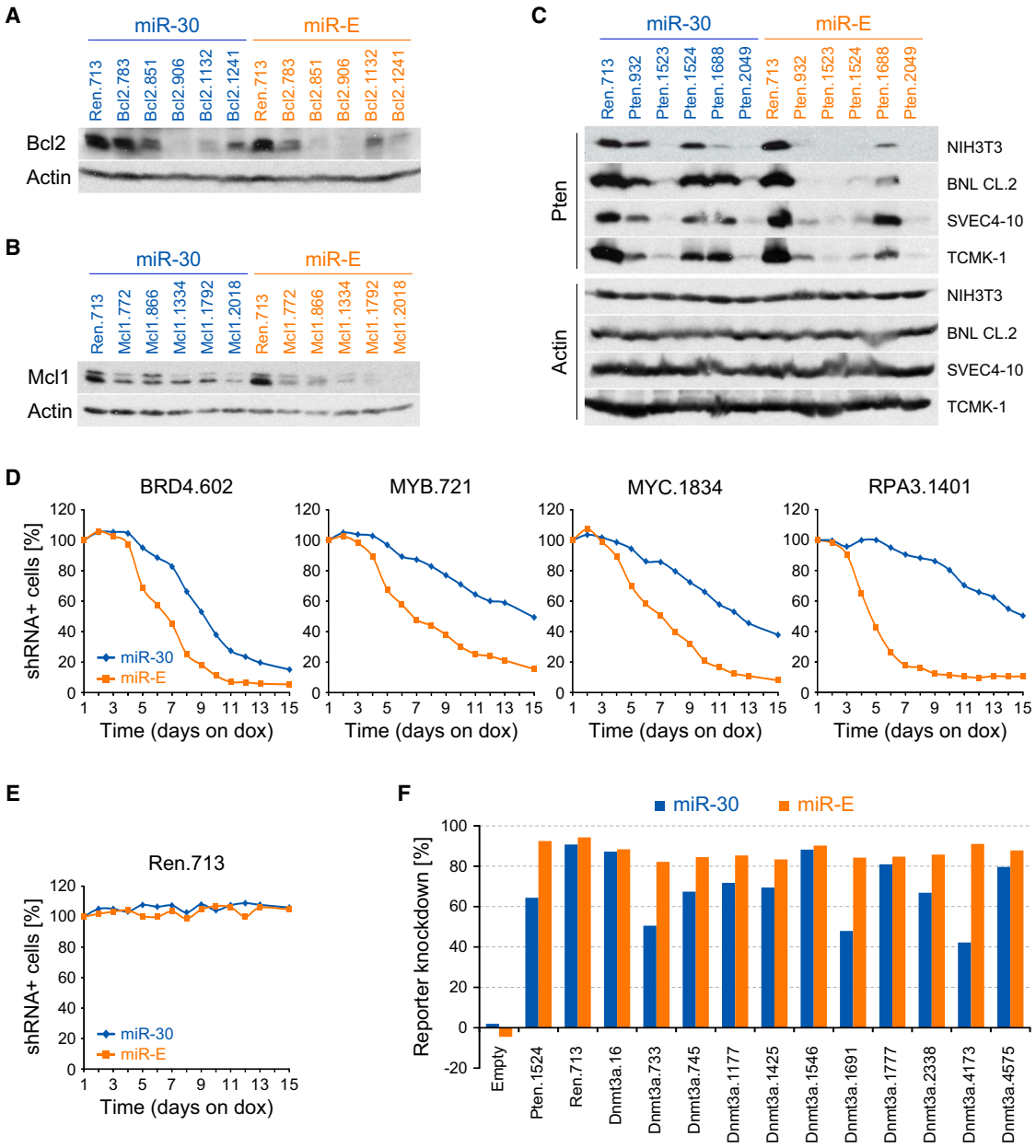

Figure 2. Validation of the miR-E Backbone

(A and B) (A) Bcl2 and (B) Mcl1 western blotting in NIH 3T3 cells transduced with LMP expressing the indicated shRNAs from either the miR-30 or miR-E backbone at single copy (see also Table S1).

(C) Comparison of Pten knockdown in fibroblasts (NIH 3T3), immortalized hepatocytes (BNL CL.2), endothelial cells (SVEC4-10), and kidney cells (TCMK-1) transduced at single copy with LMP expressing the indicated shRNAs from either the miR-30 or miR-E backbone (see also Figure S2A).

(D) Competitive proliferation assays evaluating established shRNAs targeting genes known to be essential in MOLM-13 leukemia cells. Tet-ON competent MOLM-13 cells were infected with a vector conditionally expressing the indicated shRNAs from the miR-30 or miR-E backbone. Infected cells were mixed with uninfected cells, and the percentage of shRNA expressing cells monitored upon shRNA induction by doxycycline (dox) (see also Figure S2B).

(E) Same competition assay as in (D), run with a neutral control shRNA (Ren.713). Note, no cytotoxic nor cytostatic effects were observed for neither miR-30 nor miR-E (see also Figure S2C).

(F) Reporter-based quantification of knockdown efficiency of various Dnmt3a shRNAs expressed from either the miR-30 or miR-E backbone under single-copy conditions. Mouse embryonic fibroblasts expressing a dTomato reporter tagged with target sites of the probed shRNAs were transduced at single copy with the indicated shRNAs. An empty vector, Ren.713 and Pten.1524 shRNAs served as controls (see Figures S2D–S2F for details).

See also Table S1 and Figure S2.

recognition and processing, resulting in a substantial increase in pre-miRNA and mature small RNA levels.

To test the impact of these effects under relevant single-copy conditions, we transduced NIH 3T3 fibroblasts with single LMP retroviruses expressing Ren.713 or one of three tested Pten shRNAs in the miR-30 or miR-E backbone and subsequently quantified pri-miRNA levels using quantitative RT-PCR (qRT-PCR), as well as mature miRNA abundance using deep sequencing of small RNA libraries. For all four shRNAmirs, use of the miR-E design resulted in a strong reduction of pri-miRNA levels (Figure 3F) and a massive (7- to 32-fold) increase in mature small RNA levels (Figure 3G). Although the most striking effects were observed for Pten.1524, even the potent miR-30 shRNAs Ren.713 and Pten.1523 yielded ten times more mature small RNAs when expressed from the miR-E backbone. At the same time, the absolute level of synthetic small RNAs produced from miR-E constructs remained well below the most strongly expressed endogenous miRNAs, and endogenous miRNA abundances were not affected by the use of miR-E backbones when compared to miR-30 (Pearson r = 0.99). In conjunction with previous analyses of miR-30 shRNAs (Premsrirut et al., 2011), these results demonstrate that single-copy expression of miR-E-based shRNAs has no major effects on endogenous miRNA biogenesis and thus provides a validated approach for minimizing toxicities associated with shRNA expression in other RNAi systems (Castanotto et al., 2007; Grimm et al., 2006).

We also took advantage of our deep-sequencing data set to analyze the accuracy of Drosha and Dicer processing (Figure 3H). Drosha cleaved at its canonical site in 25%–70%, whereas a large fraction of experimental pre-shRNAs resulted from cleavage 1–2 nt closer to the loop. Interestingly, while miR-30 and miR-E versions of individual shRNAs resulted in very similar cleavage patterns, specific shRNAs were associated with strong biases, suggesting that the synthetic stem sequence affects the accuracy of the Microprocessor. In stark contrast, Dicer cleaved experimental pre-miRNAs at the canonical site in 82%–100% of cases. These dramatic differences between Drosha and Dicer confirmed our rationale for keeping the experimental guide on the 3′ stem side. Only in this configuration the 5′ end of the guide, which is critical for shRNA seed-sequence determination, strand selection, and overall potency (Fellmann et al., 2011; Frank et al., 2010), will be determined by Dicer, whereas the less precise Drosha cleavage will produce variable 3′ guide ends.

DISCUSSION

The miR-30 system is a well-established and commonly used platform for shRNAmir-based RNAi that has proved its effectiveness and versatility in countless cell types and experimental settings, ranging from high-throughput screens to transgenic mice. Here, we have tackled the key limitation of current shRNAmir reagents—the frequent ineffectiveness of computationally predicted shRNAmirs to trigger potent knockdown when expressed under low-or single-copy conditions. Our study identifies and validates an improved miR-30 backbone termed miR-E, which generally boosts knockdown potency through strongly enhanced pri-miRNA processing that leads to 10- to 30-fold higher mature small RNA levels.

As a key feature underlying the improved function of miR-E, our study identifies an evolutionarily conserved element involving an ACNNC motif 3′ of the basal stem and demonstrates its decisive role for optimal performance of MIR30A-based shRNAmir reagents. Although our analysis was focused on experimental applications, a recent study has used an unbiased high-throughput assay to explore features associated with effective processing of endogenous pri-miRNAs (Auyeung et al., 2013). As one out of three sequence features, this study identified a similar motif (CNNC) at the same position and implicates it in recruitment of a splicing factor (SRSF3, also known as SRp20) that appears to be required for optimal pri-miRNA recognition and processing of endogenous miRNAs. Beyond establishing the relevance of this feature for the design of experimental shRNAmirs, our study shows that the nucleotide 5′ of CNNC can have a strong impact on pri-miRNA processing, and in endogenous miRNAs is globally biased against C. The other two features identified by Auyeung et al. (2013) include a UG at the 5′ basal stem basis and a bias for UGUG at the 5′ loop end, which had a much weaker impact on processing efficiency. Although the basal UG motif is present in all miR-30 variants, the apical UGUG has been changed to AGUG in miR-30, and restoring the original loop indeed moderately improved knockdown potency of miR-30-based shRNAs (L in Figure 1B). Though we do not rule out that alternative loops may provide opportunities to further improve the miR-E design, their implementation in existing reagents and protocols will be less straightforward and requires more extensive testing.

The simple yet effective nature of miR-E lends itself for direct implementation in any laboratory performing vector-based RNAi studies. Importantly, because the design features of miR-E do not affect the core stem-loop structure, existing miR-30 shRNAs can be easily converted through PCR subcloning into the miR-E backbone (Figure 4A), which has already been implemented in a variety of constitutive and Tet-regulated shRNA expression vectors (Figures 4B, 4C, S4C, and S4D). For de novo shRNA cloning, combining the superior processing features of miR-E with optimized Sensor-based shRNA designs (Table S3) boosts the chances of identifying potent single-copy shRNAs. Based on our experience with the miR-E backbone, >50% of top Sensor-based predictions for a given gene trigger >80% protein knockdown at single copy. In addition to their effectiveness, Sensor-based shRNAs are characterized by an extreme bias for loading the intended guide strand (Fellmann et al., 2011) and therefore provide a means to minimize passenger-mediated off-target effects. Hence, the combination of Sensor-based predictions and miR-E opens a promising avenue for generating focused and genome-wide shRNA libraries that will truly cover each gene with multiple effective shRNAs, reduce the likelihood of general and passenger-mediated off-target effects, and constitute a validated and versatile tool for high-throughput functional genetics in the postgenomic era.

EXPERIMENTAL PROCEDURES

Vectors, Backbone, and shRNA Cloning

The miR-E backbone and all variants were constructed according to sequences provided in Figure S1B using custom oligonucleotides (IDT) and

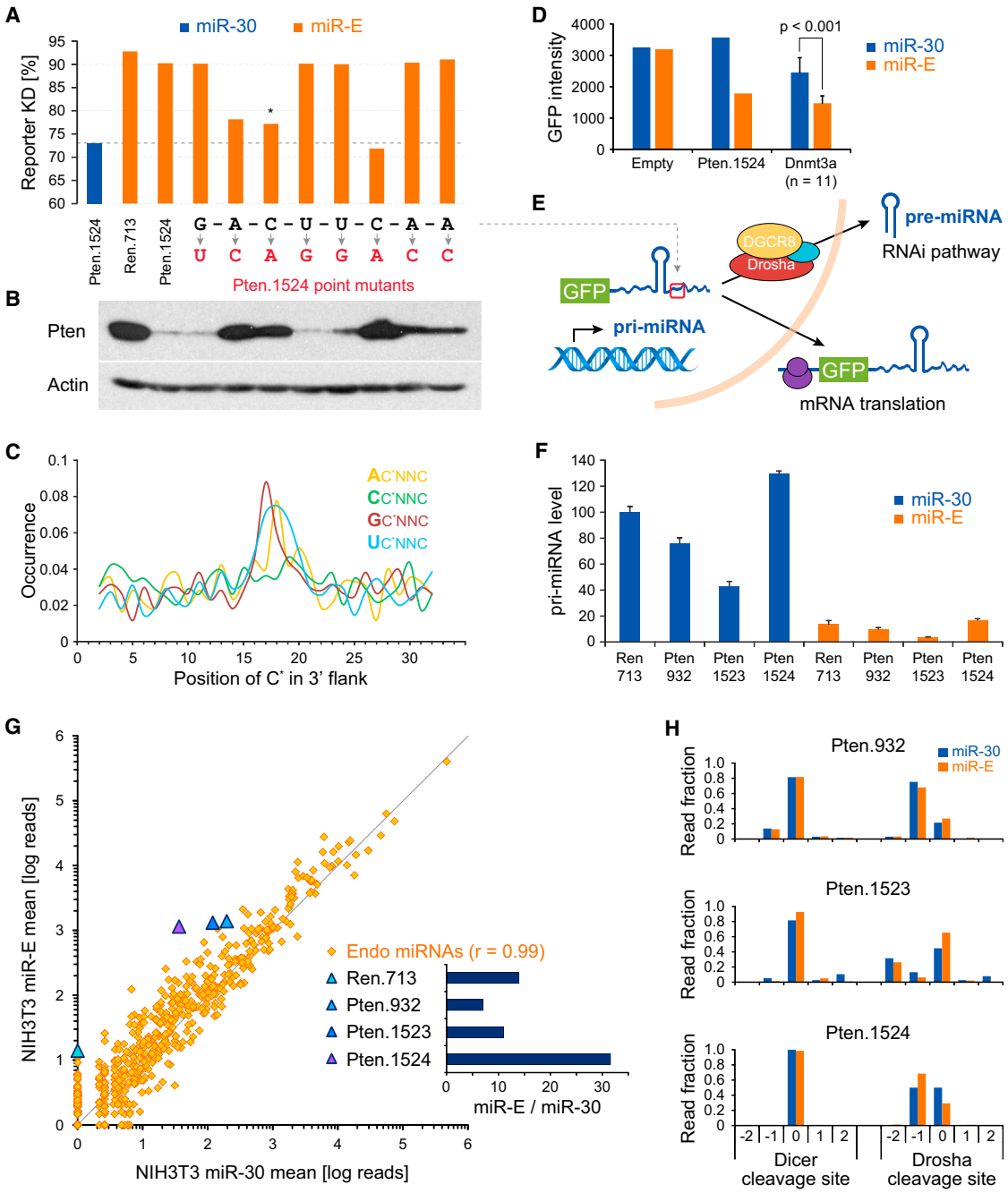

Figure 3. miR-E Enhances Knockdown Potency through Improved shRNA Biogenesis

(A and B) (A) Reporter-assay and (B) NIH 3T3 immunoblotting based evaluation of point mutations in a conserved region 3′ of the basal shRNA stem (sequence shown in black; see Figure 1A for details). Cells were transduced with the indicated control shRNAs or Pten.1524 in backbones containing the shown point mutations (red) under single-copy conditions. Asterisk, this mutant corresponds to the conventional miR-30 in this region.

(C) Sequence motif analysis in 3′ flanks of endogenous human microRNAs. Shown are the relative frequencies of occurrence for the given sequence motifs (NCNNC) at the indicated positions 3′ of the Drosha cleavage site. See Figure S3A for an analysis of 5′ flanks; see also Figure S3B and Table S2.

(D) Quantification of GFP fluorescence intensities of cells transduced at single copy with LMN vectors expressing GFP-coupled miR-30 or miR-E shRNAs (see also Figure S3C). Error bars represent SD.

(E) Schematic showing the dual use of fluorophore-miRNA polycistronic transcripts (pri-miRNA) for both protein synthesis (GFP, mRNA translation) and miRNA biogenesis (miR-30/miR-E, RNAi pathway). A red box and arrow highlight the region mutated in (A) and (B).

(F) Quantification of pri-miRNA transgene levels in NIH 3T3s expressing the indicated miR-30- or miR-E-based shRNAs from a single-copy genomic integration. The data show normalized means with error bars indicating the SD of triplicates (see also Figure S3D).

(legend continued on next page)

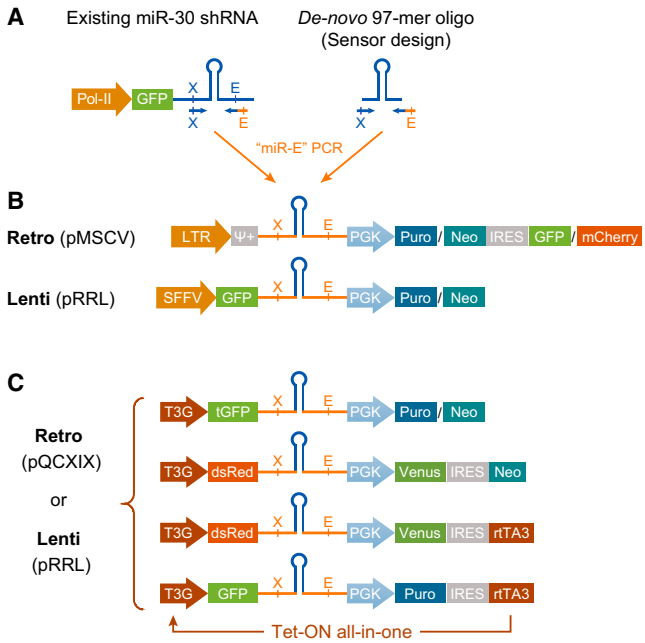

Figure 4. Implementation of the miR-E Design

(A) Schematic of PCR-based cloning of existing miR-30 shRNAs, or de novo synthesized shRNA oligonucleotides, into miR-E recipient vectors to implement the improved miR-E backbone design.

(B) Validated constitutive retro- and lentiviral miR-E expression vectors, cloned and tested in various versions featuring different drug selection (Puro/Neo) and fluorescent (GFP/Cherry) markers (see Figures S3A–S3D for details).

(C) Validated Tet-regulated miR-E expression vectors (all available in a retro- and lentiviral backbone) expressing fluorescence-coupled miR-E shRNAs from an optimized Tet-responsive element promoter (T3G), which strongly reduces leaky shRNA expression. Vector variants harboring rtTA3 feature a positive feedback loop to boost rtTA3 expression upon dox treatment (Zuber et al., 2011a) and enable single-vector ("all-in-one") Tet-ON shRNAmir expression studies.

See also Table S3 and Figure S4.

standard cloning techniques. Retroviral vectors (Figure S4C) were constructed based on existing miR-30 shRNA expression vectors in pMSCV or pQCXIX backbones (Dickins et al., 2005; Zuber et al., 2011a, 2011c); lentiviral vectors (Figure S4D) were constructed in the pRRL backbone (Dull et al., 1998) based on derivatives of pRRL.PPT.SF.GFPpre (Schambach et al., 2006). Existing miR-30 shRNAs were converted to miR-E by simple PCR amplification using the primers miRE-Xho-short-fw (5′-AGAAGGCTCGAGAAGGTATATTGC-3′) and miRE-EcoPlasmid-rev (5′-GCTCGAATTCTAGCCCCTTGAAGTCCGAGG-3′), 10 ng plasmid template, and the PfuUltra HF kit (Agilent Technologies), and amplification products were XhoI/EcoRI cloned into miR-E recipient vectors. For de novo generation of miR-E shRNAs, 97-mer oligonucleotides (IDT Ultramers) coding for the respective shRNAs (Table S1) were PCR amplified using the primers miRE-Xho-fw (5′- TGAACTCGAGAAGGTATATTGCTGTTGACAGTGAGCG-3′) and miRE-EcoOligo-rev (5′-TCTCGAATTCTAGCCCCTTGAAGTCCGAGGCAGTAGGC-3′), 0.05 ng oligonucleotide template, and the PfuUltra HF kit (Agilent), and cloned into miRE recipient vectors. More detailed protocols are provided in the Supplemental Information, and all vectors are available upon request.

Sensor Rules and shRNA Design

For de novo prediction of shRNAs, siRNA guide predictions targeting common regions of all known transcript variants (NCBI Gene, http://www.ncbi.nlm.nih.gov/gene/) were obtained using the DSIR algorithm (Vert et al., 2006). DSIR-predicted guide strands were subsequently filtered using "Sensor rules" to enrich for predictions harboring sequence features associated with effective shRNAmir processing and potent knockdown (Fellmann et al., 2011). A complete list of up-to-date Sensor-based shRNA predictions for all human and mouse coding genes are provided in Table S3.

Cell Culture, Retroviral Transduction, and Reporter Assay for shRNA Knockdown Quantification

Cell lines were retrovirally transduced under single-copy conditions as previously described (Fellmann et al., 2011). Competitive proliferation assays were performed as previously described in Tet-ON competent MOLM-13 cells (Zuber et al., 2011c). To quantify protein knockdown potency under single-copy conditions, we established a scalable two-vector/two-color reporter assay involving a reporter vector (TtNPT) constitutively expressing a hybrid transcript encoding dTomato (or Ametrine) and harboring up to 20 shRNA target sites (cloned using IDT gBlocks) in its 3′ UTR (Figure S1C). Stable reporter cells generated using retroviral transduction and selection were subsequently infected at single copy with shRNAmir expression vectors coexpressing GFP (e.g., LMP, LEPG, LENG; Figure S1D). dTomato (or Ametrine) expression was quantified in shRNA expressing (GFP^+) cells using flow cytometry after 3–6 days and compared to in-sample GFP^- control cells, as well as to parallel samples harboring control shRNAs of known potency (Figures S1C–S1F). More detailed protocols and additional experimental procedures (describing qRT-PCR, small RNA sequencing, and immunoblotting analyses) are provided in the Supplemental Information.

Analysis of Sequence Motif Conservation and microRNA Representation

Human pre-miRNA and mature miRNA data were obtained from miRbase (Kozomara and Griffiths-Jones, 2011), release 20. All pre-miRNAs were extended by 100 nt on each side by alignment to the genome. Drosha cleavage sites were determined by matching the mature miRNAs to the corresponding pre-miRNAs (842 miRNAs with annotated sense and antisense strand). MicroRNAs where the start of the sense and antisense were >60 nt apart (22) were eliminated, leaving 820 human miRNAs for analysis. Mature small RNA sequencing data of HEK293T were generated previously (Fellmann et al., 2011) and reanalyzed by alignment to the human mature miRNA collection. For miRNA representation analyses, only the guide strands (miR strand, defined by the higher read count) were used; 515 such miRNA guide strands were identified that had reads in HEK293T cells as well as an annotated sense and antisense strand.

Statistical Analysis

Linear dependence of variables was evaluated using the Pearson product-moment correlation coefficient r. Measurement predictability is given by error bars indicating the standard error of the mean (SEM) of biologically independent triplicates. Unless otherwise indicated, p values were calculated using an unpaired t test. Significance of differential miRNA representation between classes with or without a certain sequence motif was assessed using the

(G) Deep-sequencing analysis of mature small RNAs of endogenous mouse microRNAs and synthetic shRNAs. Compared are read numbers of mature small RNAs in NIH 3T3s transduced at single copy with LMP expressing one of four shRNAs from either miR-E (y axis) or miR-30 (x axis). Values for endogenous miRNAs are means of four replicates. r, Pearson correlation coefficient. Endo miRNAs, endogenous microRNAs. The inset shows the fold change in mature small RNA levels when expressing the indicated shRNAs from miR-E compared to miR-30.

(H) Drosha and Dicer cleavage accuracy for three Pten shRNAs expressed from either miR-30 or miR-E. Shown are the guide strand read fractions at each position. Cleavage is indicated with respect to the canonical cleavage site (see Figure 1A for details).

See also Table S2 and Figure S3.

Wilcoxon-Mann-Whitney U test, with Šidák correction to adjust for inflation of the α level due to large sample sets.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.11.020.

ACKNOWLEDGMENTS

We thank G.J. Hannon, S.W. Lowe, S.L. Ameres, and members of the Zuber lab for fruitful discussions and helpful suggestions, M. Weißenböck and B.Y. Yoon for excellent technical assistance, and P.K. Premsrirut for general support. We thank A. Schwarzer, A. Schambach, and C. Baum for providing pRRL.PPT.SF.GFPpre lentiviral constructs. We gratefully acknowledge X. Zhou for assistance with small RNA library preparation and M. Hammell for advice on statistical analysis. We also thank the IMP/IMBA Biooptics and Molecular Biology service facilities for excellent technical support. C.F. is a founder and employee and J.Z. a member of the scientific advisory board of Mirimus, a company that develops RNAi-based reagents and transgenic mice and that has applied for intellectual property protection of some of the methods described here. The Research Institute of Molecular Pathology (IMP) is generously funded by Boehringer Ingelheim.

Received: July 10, 2013
Revised: October 9, 2013
Accepted: November 11, 2013
Published: December 12, 2013

REFERENCES

Ameres, S.L., Martinez, J., and Schroeder, R. (2007). Molecular basis for target RNA recognition and cleavage by human RISC. Cell *130*, 101–112.

Auyeung, V.C., Ulitsky, I., McGeary, S.E., and Bartel, D.P. (2013). Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. Cell *152*, 844–858.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281–297.

Boden, D., Pusch, O., Silbermann, R., Lee, F., Tucker, L., and Ramratnam, B. (2004). Enhanced gene silencing of HIV-1 specific siRNA using microRNA designed hairpins. Nucleic Acids Res. *32*, 1154–1158.

Boudreau, R.L., Monteys, A.M.A.S., and Davidson, B.L. (2008). Minimizing variables among hairpin-based RNAi vectors reveals the potency of shRNAs. RNA *14*, 1834–1844.

Boudreau, R.L., Martins, I., and Davidson, B.L. (2009). Artificial microRNAs as siRNA shuttles: improved safety as compared to shRNAs in vitro and in vivo. Mol. Ther. *17*, 169–175.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642–655.

Castanotto, D., Sakurai, K., Lingeman, R., Li, H., Shively, L., Aagaard, L., Soifer, H., Gatignol, A., Riggs, A., and Rossi, J.J. (2007). Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. Nucleic Acids Res. *35*, 5154–5164.

Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E., et al. (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev. *24*, 992–1009.

Chung, K.-H., Hart, C.C., Al-Bassam, S., Avery, A., Taylor, J., Patel, P.D., Vojtek, A.B., and Turner, D.L. (2006). Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155. Nucleic Acids Res. *34*, e53.

Dickins, R.A., Hemann, M.T., Zilfou, J.T., Simpson, D.R., Ibarra, I., Hannon, G.J., and Lowe, S.W. (2005). Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. Nat. Genet. *37*, 1289–1295.

Dow, L.E., Premsrirut, P.K., Zuber, J., Fellmann, C., McJunkin, K., Miething, C., Park, Y., Dickins, R.A., Hannon, G.J., and Lowe, S.W. (2012). A pipeline for the generation of shRNA transgenic mice. Nat. Protoc. *7*, 374–393.

Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D., and Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. J. Virol. *72*, 8463–8471.

Fellmann, C., Zuber, J., McJunkin, K., Chang, K., Malone, C.D., Dickins, R.A., Xu, Q., Hengartner, M.O., Elledge, S.J., Hannon, G.J., and Lowe, S.W. (2011). Functional identification of optimized RNAi triggers using a massively parallel sensor assay. Mol. Cell *41*, 733–746.

Feng, Y., Zhang, X., Song, Q., Li, T., and Zeng, Y. (2011). Drosha processing controls the specificity and efficiency of global microRNA expression. Biochim. Biophys. Acta *1809*, 700–707.

Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat. Rev. Genet. *9*, 102–114.

Frank, F., Sonenberg, N., and Nagar, B. (2010). Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. Nature *465*, 818–822.

Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F., and Kay, M.A. (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. Nature *441*, 537–541.

Gu, S., Jin, L., Zhang, Y., Huang, Y., Zhang, F., Valdmanis, P.N., and Kay, M.A. (2012). The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo. Cell *151*, 900–911.

Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., Sohn, S.Y., Cho, Y., Zhang, B.-T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell *125*, 887–901.

Hannon, G.J., and Rossi, J.J. (2004). Unlocking the potential of the human genome with RNA interference. Nature *431*, 371–378.

Hinterberger, M., Aichinger, M., Prazeres da Costa, O., Voehringer, D., Hoffmann, R., and Klein, L. (2010). Autonomous role of medullary thymic epithelial cells in central CD4(+) T cell tolerance. Nat. Immunol. *11*, 512–519.

Hu, G., and Luo, J. (2012). A primer on using pooled shRNA libraries for functional genomic screens. Acta Biochim. Biophys. Sin. (Shanghai) *44*, 103–112.

Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. *39* (Database issue), D152–D157.

Liu, Y.P., Haasnoot, J., ter Brake, O., Berkhout, B., and Konstantinova, P. (2008). Inhibition of HIV-1 by multiple siRNAs expressed from a single microRNA polycistron. Nucleic Acids Res. *36*, 2811–2824.

McBride, J.L., Boudreau, R.L., Harper, S.Q., Staber, P.D., Monteys, A.M., Martins, I., Gilmore, B.L., Burstein, H., Peluso, R.W., Polisky, B., et al. (2008). Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi. Proc. Natl. Acad. Sci. USA *105*, 5868–5873.

Mohr, S.E., and Perrimon, N. (2012). RNAi screening: new approaches, understandings, and organisms. Wiley Interdiscip. Rev. RNA *3*, 145–158.

Pan, Q., van der Laan, L.J.W., Janssen, H.L., and Peppelenbosch, M.P. (2012). A dynamic perspective of RNAi library development. Trends Biotechnol. *30*, 206–215.

Premsrirut, P.K., Dow, L.E., Kim, S.Y., Camiolo, M., Malone, C.D., Miething, C., Scuoppo, C., Zuber, J., Dickins, R.A., Kogan, S.C., et al. (2011). A rapid and scalable system for studying gene function in mice using conditional RNA interference. Cell *145*, 145–158.

Schambach, A., Galla, M., Modlich, U., Will, E., Chandra, S., Reeves, L., Colbert, M., Williams, D.A., von Kalle, C., and Baum, C. (2006). Lentiviral vectors pseudotyped with murine ecotropic envelope: increased biosafety and convenience in preclinical research. Exp. Hematol. *34*, 588–592.

Silva, J.M., Li, M.Z., Chang, K., Ge, W., Golding, M.C., Rickles, R.J., Siolas, D., Hu, G., Paddison, P.J., Schlabach, M.R., et al. (2005). Second-generation

shRNA libraries covering the mouse and human genomes. Nat. Genet. *37*, 1281–1288.

Stegmeier, F., Hu, G., Rickles, R.J., Hannon, G.J., and Elledge, S.J. (2005). A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. Proc. Natl. Acad. Sci. USA *102*, 13212–13217.

Vert, J.P., Foveau, N., Lajaunie, C., and Vandenbrouck, Y. (2006). An accurate and interpretable model for siRNA efficacy prediction. BMC Bioinformatics *7*, 520.

Zeng, Y., Wagner, E.J., and Cullen, B.R. (2002). Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol. Cell *9*, 1327–1333.

Zhu, X., Santat, L.A., Chang, M.S., Liu, J., Zavzavadjian, J.R., Wall, E.A., Kivork, C., Simon, M.I., and Fraser, I.D. (2007). A versatile approach to multiple gene RNA interference using microRNA-based short hairpin RNAs. BMC Mol. Biol. *8*, 98.

Zuber, J., McJunkin, K., Fellmann, C., Dow, L.E., Taylor, M.J., Hannon, G.J., and Lowe, S.W. (2011a). Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. Nat. Biotechnol. *29*, 79–83.

Zuber, J., Shi, J., Wang, E., Rappaport, A.R., Herrmann, H., Sison, E.A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M., et al. (2011b). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature *478*, 524–528.

Zuber, J., Rappaport, A.R., Luo, W., Wang, E., Chen, C., Vaseva, A.V., Shi, J., Weissmueller, S., Fellmann, C., Taylor, M.J., et al. (2011c). An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance. Genes Dev. *25*, 1628–1640.