ARTICLES

Widespread changes in protein synthesis induced by microRNAs

Matthias Selbach[1], Björn Schwanhäusser[1]*, Nadine Thierfelder[1]*, Zhuo Fang[1], Raya Khanin[2] & Nikolaus Rajewsky[1]

Animal microRNAs (miRNAs) regulate gene expression by inhibiting translation and/or by inducing degradation of target messenger RNAs. It is unknown how much translational control is exerted by miRNAs on a genome-wide scale. We used a new proteomic approach to measure changes in synthesis of several thousand proteins in response to miRNA transfection or endogenous miRNA knockdown. In parallel, we quantified mRNA levels using microarrays. Here we show that a single miRNA can repress the production of hundreds of proteins, but that this repression is typically relatively mild. A number of known features of the miRNA-binding site such as the seed sequence also govern repression of human protein synthesis, and we report additional target sequence characteristics. We demonstrate that, in addition to downregulating mRNA levels, miRNAs also directly repress translation of hundreds of genes. Finally, our data suggest that a miRNA can, by direct or indirect effects, tune protein synthesis from thousands of genes.

MicroRNAs are key *trans*-acting factors that post-transcriptionally regulate metazoan gene expression, and identifying miRNA targets as well as the effect that miRNAs exert on them is a fundamental question for understanding life, health and disease[1–5]. The first identified miRNA targets in *Caenorhabditis elegans* were found to be translationally repressed whereas target mRNA levels were only mildly downregulated. Subsequently, similar cases were reported in mammalian systems[6,7]. Reporter constructs provided experimental evidence that miRNAs can directly repress translation initiation[8–10]. Furthermore, it has been shown that different mechanisms exist by which miRNAs repress protein synthesis or induce mRNA degradation[6,11]. Overexpressing a miRNA in human cell lines causes mostly mild (less than twofold) downregulation of hundreds of mRNAs, of which many are direct targets[12]. Nonetheless, these results do not reveal how much control miRNAs exert on protein synthesis. Because protein synthesis is one of the most important quantities for the phenotype, a fundamental question about gene regulation has therefore remained unanswered.

Identifying miRNA targets has been the subject of a steeply growing number of computational[13–16] and experimental[17–20] approaches. Although certain features of the miRNA-binding site such as seed sites (Watson–Crick consecutive base pairing between mRNAs and the miRNA at position 2–7 counted from its 5′ end) located in the 3′ untranslated regions (3′ UTRs) of mRNAs are important, it is unknown how relevant they are for changes in protein production. Several rules regarding the architecture of miRNA-binding sites have been proposed to explain differences in their efficacy in mRNA degradation versus translational repression[6,21]. However, these rules were based on a few target sites that were studied mostly in reporter assays with non-endogenous proteins. Another study about the effects of miRNA on the proteome was limited by the small number (12) of detected downregulated proteins[22]. Furthermore, different proteins have different turnover times. For example, if a miRNA completely shuts off protein production, steady-state levels of high-turnover proteins will change rapidly whereas stable proteins will be affected later. Therefore, changes in protein concentrations as measured by standard techniques cannot quantify changes in protein synthesis if protein levels are not stationary. In fundamental biological processes such as differentiation, the expression of miRNAs is strongly induced (or switched off) in a relatively small time window[23]. Thus, to assess endogenous regulation of mRNA translation by miRNAs, a technique is needed to measure directly genome-wide changes in protein synthesis shortly after changes in miRNA expression.

pSILAC measures changes in protein production

To overcome these problems, we devised a new variant of SILAC (stable isotope labelling with amino acids in cell culture). In SILAC, proteins are metabolically labelled by cultivating cells in growth medium containing heavy isotope versions of essential amino acids[24,25]. Mass spectrometry can distinguish peptides derived from SILAC-labelled proteins. The ratio of peptide peak intensities reflects differences in corresponding protein abundance. We reasoned that by pulse-labelling with two different heavy stable isotope labels we could measure changes in protein production between two samples. In our pulsed SILAC (pSILAC) method, cells in the two samples are pulse-labelled with two different heavy versions of amino acids. During labelling, all newly synthesized proteins will be 'heavy' or 'medium-heavy' (Fig. 1a). Pre-existing proteins present before labelling remain in the light form and are ignored. Only intensity differences between newly synthesized proteins (medium-heavy and heavy) are considered. Hence, pSILAC quantifies differences in protein production between both samples integrated over the measurement time after the pulse[26]. This is fundamentally different from pulse-labelling with a single label to determine protein turnover or transport[27–29]. We combined pSILAC with state-of-the-art mass-spectrometry-based proteomics[30–32] to measure changes in production of ~5,000 proteins altogether.

We performed transfections to individually overexpress five human miRNAs in HeLa cells. These miRNAs are tissue-specific and virtually absent in HeLa cells (miR-1, miR-155) or expressed in many tissues (miR-16, miR-30a, let-7b) including HeLa cells[33]. At least 90% of all cells could be efficiently transfected (Supplementary Fig. 1), and miRNAs were overexpressed for at least 32 h post-transfection (not shown). Changes in protein production

[1]Max Delbrück Center for Molecular Medicine, Robert-Rössle Str. 10, D-13125 Berlin, Germany. [2]Department of Statistics, 15 University Gardens, University of Glasgow, Glasgow, G12 8QQ, UK.

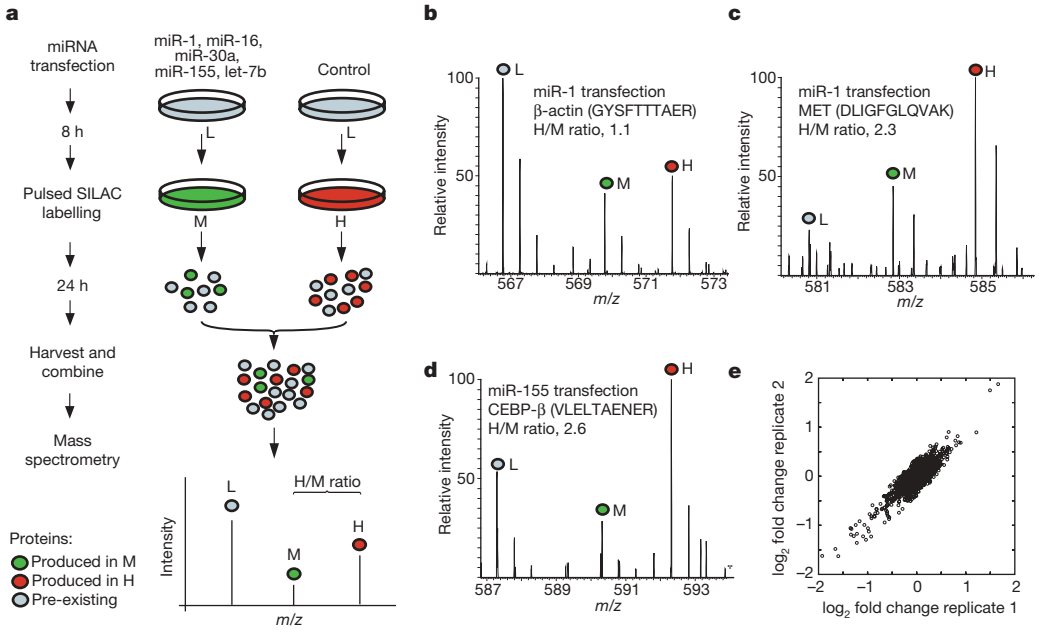
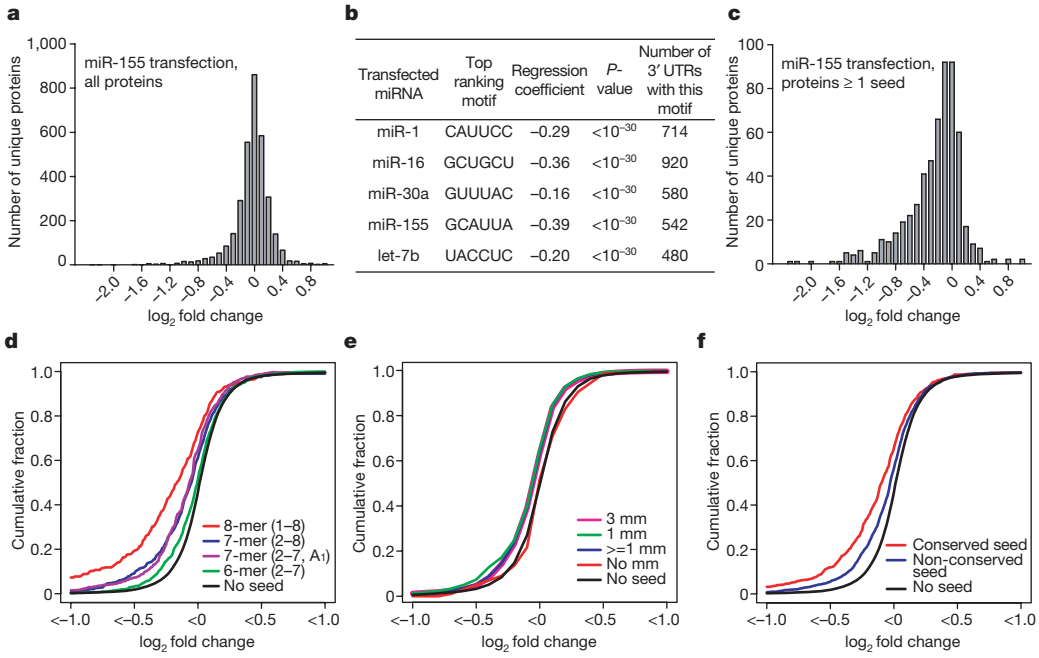


Figure 1 | Global analysis of changes in protein production induced by microRNAs. a, HeLa cells cultivated in normal light (L) medium were either transfected with a miRNA or mock transfected. After 8 h, transfected and control cells were pulse-labelled by transferring them to culture medium containing medium-heavy (M) or heavy (H) isotope-labelled amino acids, respectively (pSILAC). All newly synthesized proteins will appear in the H or M form. Samples were combined after 24 h and analysed by mass spectrometry. Intensity peak ratios between heavy and medium-heavy peptides (H/M ratio) reflect changes in protein production. RNA from the same samples was analysed by microarrays. **b**–**d**, Exemplary peptide mass spectra (sequences are in parentheses). The production of most proteins is unaltered, as shown for a β-actin peptide. In contrast, synthesis of MET and CEBPβ is reduced by miR-1 or miR-155 overexpression. **e**, Reproducibility of pSILAC (biological replicate, see Supplementary Methods).

were measured by pulse-labelling at 8 h post-transfection over a time period of 24 h. Representative mass spectra are shown in Fig. 1b–d. In total, we identified 4,961 proteins in HeLa cells with high confidence (false discovery rate <1%, see Supplementary Methods). Although mass spectrometry is biased to detect more highly expressed genes, this bias was mild and did not affect the detection range (Supplementary Fig. 2). We validated 16 out of 16 selected pSILAC measurements by western blotting (Supplementary Fig. 3). Analysis of biological replicates showed high correlation (Pearson correlation coefficient ~0.9) over the entire dynamic range (Fig. 1e).



Transfected miRNA	Top ranking motif	Regression coefficient	*P*-value	Number of 3′ UTRs with this motif
miR-1	CAUUCC	−0.29	$<10^{-30}$	714
miR-16	GCUGCU	−0.36	$<10^{-30}$	920
miR-30a	GUUUAC	−0.16	$<10^{-30}$	580
miR-155	GCAUUA	−0.39	$<10^{-30}$	542
let-7b	UACCUC	−0.20	$<10^{-30}$	480

Figure 2 | miRNAs downregulate protein synthesis of hundreds of genes. a, Histogram of changes in production of 3,299 proteins in HeLa cells after miR-155 overexpression. **b**, An unbiased search for 3′ UTR motifs that correlate with pSILAC fold changes yielded precisely the miRNA seed sequences. **c**, Proteins with miR-155 seeds tend to be downregulated by miR-155 overexpression. **d**, Cumulative distributions of different seed classes (matches to positions 1–8 (8-mer), 2–8 (7-mers), 2–7 with adenosine in position 1 (2–7, A_1) and 2–7 (6-mer)). **e**, Mismatches (mm) between positions 9 and 11 of the miRNA and target mRNAs with a seed correlate with downregulation. Protein synthesis from mRNAs with perfect complementarity at positions 9–11 (red) and synthesis from mRNAs without seeds (black) is indistinguishable. **f**, Conserved seeds mediate more downregulation than non-conserved seeds. Results are shown for pooled data based on messages with one seed only (**d**–**f**).

mRNA sequence features of repressed proteins

Perhaps surprisingly, pSILAC revealed that miRNA overexpression had, overall, mild effects on the synthesis of most of the 3,000–3,500 proteins quantified in each transfection (shown for miR-155 in Fig. 2a). Because miRNAs are thought to target mRNAs primarily by binding *cis*-regulatory sites in 3′ UTRs, we used a linear-regression-based analysis[34] to identify 3′ UTR sequence motifs that correlated best with changes in protein production. This method performs an unbiased screen for all nucleotide motifs of one to six nucleotides in length. For each miRNA, the most significant motif of all possible 5,460 motifs was precisely the seed of the respective miRNA (Fig. 2b), and correlated with downregulation. The same motif search in 5′ UTRs had no significant results. Searching coding sequences yielded the seed in only two experiments (let-7b, miR-16), and further analyses showed that 3′ UTRs exert the strongest effect (Supplementary Fig. 4). Taking miR-155 as an example, the seed enrichment in downregulated proteins is illustrated by the histogram of fold changes for proteins that contain at least one seed in their mRNA 3′ UTRs (Fig. 2c). Thus, proteins with reduced synthesis are enriched in direct miRNA targets, and a primary motif to mediate this reduction is the 3′ UTR seed. Certain characteristics such as seed-flanking nucleotides have been reported to affect the degree of mRNA degradation by miRNAs[35,36], and we show that these effects are also involved in repressing protein production (Fig. 2d).

When small interfering RNAs (siRNAs) are perfectly complementary to their targets, mRNA cleavage occurs between nucleotides 10 and 11 opposite the siRNA guide strand; in contrast, mismatches in this region strongly reduce cleavage[37–39]. A small-scale study with reporter constructs suggested that siRNA–mRNA pairs with mismatches between nucleotides 9–11 of the siRNA are mainly repressed at the protein level with little effect on the transcript[21]. We found that only seed-containing mRNAs with at least one mismatch were, overall, repressed at the protein level (Fig. 2e). In contrast, protein production from seed-containing mRNAs with perfect base pairing from nucleotides 9 to 11 and mRNAs lacking seeds was indistinguishable. Hence, although mismatches are deleterious to siRNA-mediated cleavage of mRNAs, they correlate with increased repression of protein production by miRNAs. We also found that, on average, repression is more pronounced for conserved than for non-conserved seed sites (Fig. 2f), indicating that our experiments reflect biological relevance and that there are determinants in addition to the seed that mediate efficient downregulation of protein synthesis.

We next quantified how many of the downregulated proteins can be explained by the seed. We recorded how many proteins with at least one 3′ UTR seed site were downregulated by at least *c*-fold as a function of *c* (Fig. 3a). For example, the production of more than 300 proteins with seeds was downregulated by at least 30% ($\log_2$-fold change <-0.5). These proteins amounted to roughly 60–70% of all measured proteins downregulated by at least this much (Fig. 3b). Because the background seed frequency is 10–30% (Fig. 3b, dashed horizontal lines), we can explain up to 60% of the ~300 proteins by the presence of seeds. It remains an open question how many proteins without a seed are direct targets. Nevertheless, pSILAC clearly generates lists of proteins enriched in direct targets. We independently validated the 3′ UTR-dependence of protein production by dual luciferase reporters for eight 3′ UTRs with a seed for either miR-1 or let-7b (see Supplementary Methods). The correlation with the corresponding pSILAC data was high (Fig. 3c).

pSILAC data and target predictions

Having shown that pSILAC data are enriched in direct miRNA targets, we tested how miRNA target predictions correlate with our data. We calculated the fraction of predicted mRNA targets for which protein production was downregulated by at least *c*-fold. The results were consistent for all values of *c* and all miRNAs individually (data not shown). For example, roughly 27% of all 24,238 mRNAs present in the pSILAC data were downregulated more strongly than -0.1 $\log_2$-fold change (Fig. 3d and Supplementary Table 1). A completely random selection would therefore have 27% overlap with pSILAC data. This background accuracy was exceeded by all methods except one based on 5′ UTRs. Simply considering seed sites boosts the accuracy to 44%. This accuracy was only topped by three methods that use evolutionary conservation of seed sites as an additional filter. Almost all other methods, in part based on site-accessibility evaluation, made fewer predictions with less accuracy.

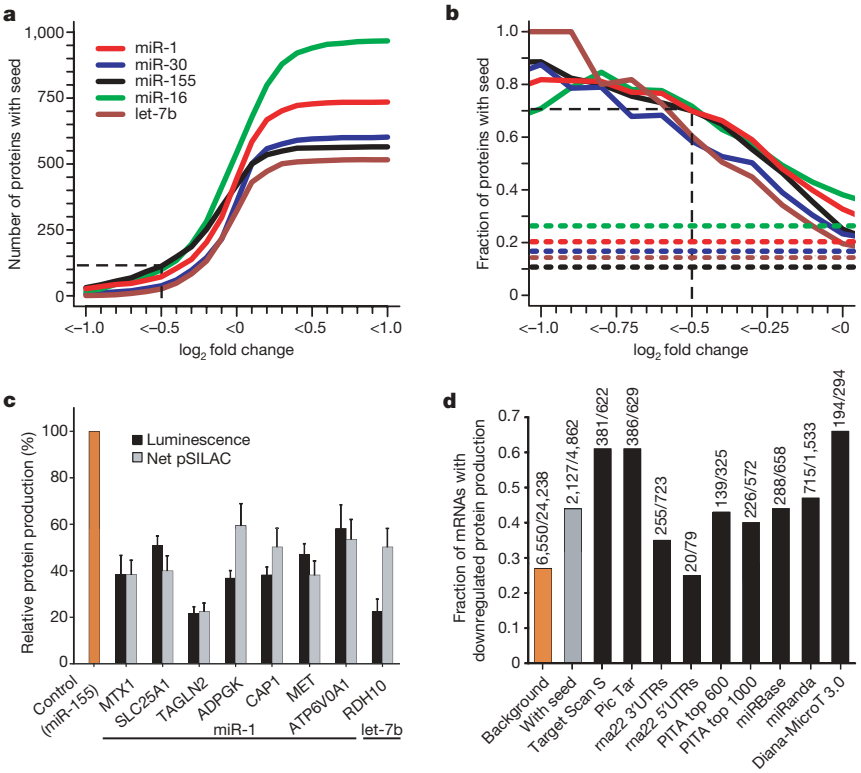


Figure 3 | The miRNA seed explains a large fraction of downregulated protein synthesis. **a**, Cumulative number of proteins with seeds as a function of changes in their production. For a given cutoff, this indicates the number of downregulated seed-containing proteins (shown for -0.5 $\log_2$-fold change). **b**, Fraction of proteins with a seed as a function of repression. Background seed frequency of unchanged proteins (absolute $\log_2$-fold change <0.1) ranges from 10–30% (dashed lines). **c**, Dual luciferase reporter assays for 3′ UTR-mediated regulation by miRNAs (± s.d., $n = 3$). 'Net pSILAC' refers to the difference of pSILAC $\log_2$-fold changes for the miRNA and the control (error bars show 95% confidence interval, see Supplementary Methods). **d**, The fraction of computationally predicted target mRNAs with reduced protein production ($\log_2$-fold change <-0.1) is calculated for all five miRNA data sets pooled.

Translational repression by miRNAs

pSILAC measures changes in the amount of newly synthesized proteins between two samples. This depends on changes of mRNA levels and, in addition, on translational regulation. To discern these two mechanisms, in all pSILAC experiments we measured the mRNA fold changes between the miRNA-transfected sample and the control by Affymetrix microarrays at the beginning of the pulse labelling ($t_1 = 8$ h) and at the end ($t_2 = 32$ h). A total of 69 quantitative polymerase chain reactions with reverse transcription (qRT–PCRs) demonstrated that our microarray data have little compression or other distortion effects in the range where most mRNA fold changes were observed (Supplementary Fig. 5).

For miR-1 as an example, we present the relationship between miRNA-induced fold changes in protein production (pSILAC) and mRNA fold changes (Fig. 4a, b) separately for t_1 and t_2. Very few genes had fold changes of unequal sign and reasonable magnitude ($\geq$1.3-fold). The correlation between mRNA fold changes and pSILAC fold changes became better at t_2. In particular, many genes with downregulated protein production but little mRNA fold changes at t_1 shifted towards greater mRNA fold changes at t_2. Similar overall effects could be seen for the other miRNAs. Nevertheless, the considerable scatter indicates substantial and widespread post-transcriptional regulation of gene expression.

The distribution of fold changes measured by microarray and pSILAC was similar (Fig. 4c, histograms). However, the average number (s) of seeds per gene was higher for more highly downregulated genes. Seed enrichment was not observed for upregulated genes, indicating that the recently reported miRNA-mediated activation of gene expression did not occur under our experimental conditions[40]. For downregulated genes, log-fold changes were linearly correlated with s. Thus, if a target has two seeds, the repressive effect is multiplicative, as has been observed in small-scale studies[15,41]. pSILAC data also support earlier findings[36] that synergistic effects are higher for two nearby seeds (<40 nucleotides) compared to larger spacings (>40 nucleotides; P-value 0.003, one-sided Wilcoxon test). Intriguingly, the slope of s in Fig. 4c is steeper for pSILAC fold changes, suggesting that the multiplicity of a miRNA-binding site in the same 3′ UTR exerts a stronger direct effect on protein production than on mRNA levels. To assess miRNA-mediated changes in translation rates for each gene, we subtracted the $\log_2$ mRNA from the $\log_2$ pSILAC fold changes, and plotted s as a function of these differences (Fig. 4d). The linear decay of s towards the regime of equal fold changes indicates that, in addition to mediating mRNA downregulation[12], the seed also mediates direct repression of translation rates for hundreds of genes.

Endogenous miRNA knockdown

It could be argued that the overexpression of miRNAs can lead to largely non-physiological effects. We therefore used a locked nucleic acid (LNA) approach[42,43] to knockdown let-7b in HeLa cells (Fig. 5a), and measured changes in protein production and mRNA levels as before. Luciferase reporter experiments demonstrated that our knockdown functionally derepressed a known *let-7* target[44] mediated by seed sites (Supplementary Fig. 6). As in the overexpression experiments, an unbiased search for 3′ UTR motifs identified the let-7b seed as the best match. Coding sequences and 5′ UTRs did not yield significant results. Further analyses showed that all effects for seed-mediated targets that we report for the overexpression experiments hold true for the let-7b knockdown after flipping the sign of pSILAC and microarray fold changes, including correlation of target-finding algorithms with pSILAC data (Supplementary Fig. 7). Together, these data suggest that the miRNA overexpression experiments are also physiologically relevant.

let-7b tunes production of thousands of proteins

When we compared the cellular response to let-7b overexpression and knockdown we observed a marked anti-correlation, not only for seed-mediated let-7b targets but also for most of the ~2,700 proteins quantified in both experiments (that is, for both direct and indirect effects; Fig. 5b). For example, when considering all ~130 proteins with a fold change of at least 15% in both the overexpression and knockdown experiments, most were up in one of the experiments but down in the other, irrespective of seeds (Fig. 5c). In contrast, almost all proteins with seeds were down in the overexpression experiment

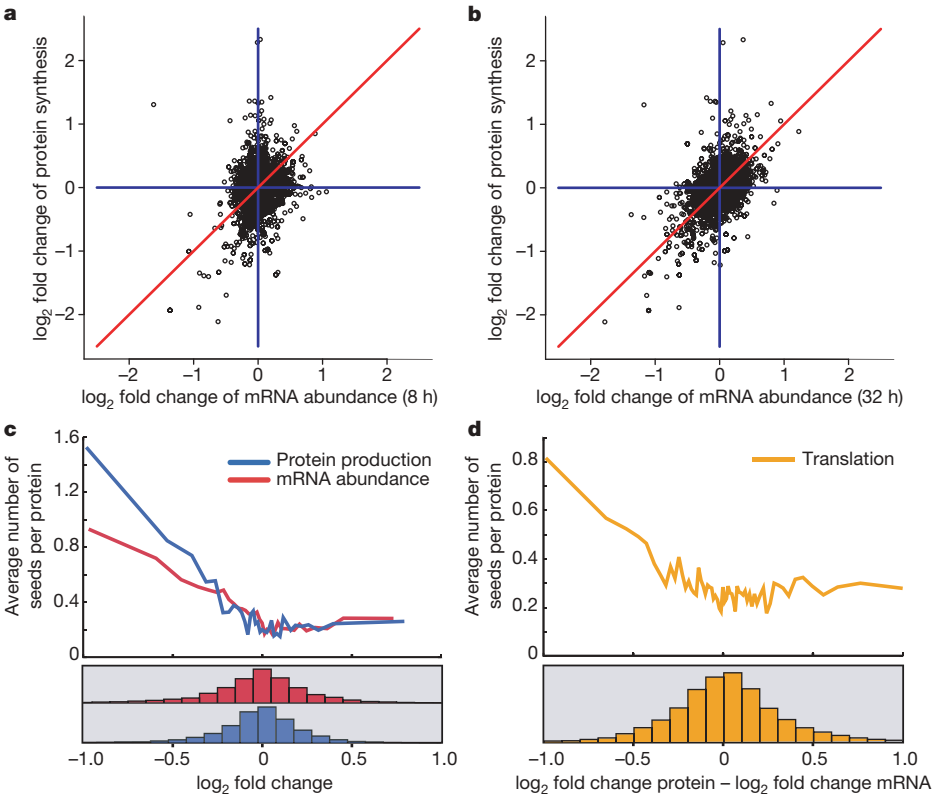


Figure 4 | miRNAs inhibit translation on a genome-wide scale. **a**, Changes in protein production between 8 h and 32 h after miR-1 transfection with mRNA fold changes at 8 h reveal poor overall correlation. **b**, mRNA levels at 32 h correlate remarkably well with changes in protein synthesis. **c**, Overall fold changes of mRNA and protein synthesis are similar (histograms). Reduced protein production and mRNA levels correlate with seed frequency (curves represent proteins ranked by fold change and grouped into bins of 250). **d**, Translational repression by miRNAs is revealed by subtracting mRNA log changes from log changes in protein production. Increased seed frequency, averaged as in **c**, correlates with translational repression. Results are shown for pooled data (**c**, **d**) after discarding genes with mRNA and pSILAC changes of unequal sign.

and up in the knockdown. When averaging the data, we found a linear response of the entire proteome to miRNA misexpression with a slope of −0.3 (Fig. 5b, inset), demonstrating that, on average, let-7b overexpression induced roughly threefold higher $\log_2$-fold changes than let-7b knockdown. Together, these data indicate that upregulation and downregulation of stationary let-7b levels has largely complementary effects on the proteome; that is, let-7b levels can tune protein production from thousands of genes.

Discussion

Here we have measured for the first time changes in cellular protein synthesis in response to miRNA induction or knockdown on a proteome-wide scale. Our results show that a single miRNA can directly downregulate production of hundreds of proteins. In addition to the known effect on global mRNA levels[12], our data strongly indicate that miRNAs translationally repress hundreds of direct target genes. Using an unbiased approach, we identified the seed sequence in the 3′ UTR as a primary motif of miRNA-mediated regulation of protein production. The seed correlated with both mRNA degradation and translational repression (Fig. 4c, d).

Perhaps surprisingly, the repressive effect on individual proteins was relatively small and rarely exceeded fourfold. Because we performed pulsed labelling, this result cannot be explained by persistence of stable proteins. Other investigators observed much higher fold changes (up to 30-fold) in a similar system (double-stranded RNA (dsRNA) transfection in HeLa cells) with artificial reporter constructs[41]. One explanation for this apparent discrepancy is that very few (<0.5%) 3′ UTRs in our data set have more than three seed sites for a given miRNA (and this value is representative for the whole genome) whereas artificial reporter constructs are designed to contain up to six closely spaced miRNA binding sites.

Identifying functionally important miRNA targets is crucial for understanding miRNA functions. By directly measuring changes in protein production, pSILAC data are likely to be more relevant to the phenotypes than microarray data. We also note that a number of targets are almost exclusively repressed at the level of translation and hence missed by microarrays. pSILAC allows assessment of the early effects of miRNAs on translation. This is a considerable advantage over techniques that assay changes in steady-state protein levels and are therefore almost certainly confounded by indirect effects. Although not all changes in peptide peak intensities reflect true differences in protein synthesis, a direct comparison of pSILAC and luciferase measurements yields very similar results over two orders of magnitude[26]. Catalogues of proteotypic peptides will further improve this accuracy and help to achieve full-proteome coverage[45]. pSILAC and microarray data can be queried at http://psilac.mdc-berlin.de.

Although artificially overexpressing miRNAs might cause non-physiological effects, we found that overexpression and knockdown of let-7b inversely modulates protein production, suggesting that such effects do not dominate. Nevertheless, transfecting miRNAs that are not endogenously expressed will probably expose many mRNAs to miRNAs that are never coexpressed in the same cell type. Therefore, it could be argued that a number of target for miR-1 and miR-155 identified by pSILAC are irrelevant *in vivo*. However, transfecting a tissue-specific miRNA into HeLa cells shifts the entire gene expression profile towards that tissue[12]. Furthermore, we show that evolutionarily conserved target sites cause stronger effects than non-conserved sites. Altogether, our data probably contain many physiologically relevant direct targets. These arguments are strengthened by the highly significant correlation of pSILAC data with a number of published miRNA target predictions. Seed-based methods had the highest overlap with pSILAC data. Consistently, many downregulated genes could be explained by seed sites. A number of repressed proteins without seeds are nevertheless probably direct targets of the respective miRNAs. However, although some algorithms include searches for such sites, it seems that they could not identify these non-canonical sites with high success.

Our data indicate that most targets are repressed at both the mRNA and the translational level. As revealed by Fig. 4d, how much both processes contribute to downregulation depends on the individual miRNA–mRNA pair. To test whether targets with strong translational repression share functional properties, we performed gene ontology analysis for proteins with large protein and mRNA fold-change differences ($\log_2$-fold change pSILAC − mRNA < −0.3). Intriguingly, we found over-representation of proteins synthesized at endoplasmic-reticulum-associated ribosomes (gene ontology categories 'intrinsic to membrane' and 'endoplasmic reticulum', corrected *P*-values <0.0001 and <0.005, respectively; Supplementary Table 2). Hence, translational repression seems stronger for mRNAs translated at endoplasmic-reticulum-associated ribosomes compared to free cytosolic ribosomes. Thus, endoplasmic-reticulum-associated ribosomes might be more sensitive to miRNA-mediated translational repression. It is tempting to speculate that mRNAs from free ribosomes but not from endoplasmic-reticulum-associated ribosomes are targeted to processing bodies (P-bodies) for degradation[46]. Because the endoplasmic reticulum is considered to lack proteolytic activity, this finding also suggests that co-translational degradation of nascent peptides is not the predominant mechanism of miRNA-mediated translational repression for this subset of targets[47].

Finally, we showed that overexpression and knockdown of let-7b had largely inverse effects on the protein production of thousands of genes, indicating that altering stationary levels of an endogenously expressed miRNA can tune synthesis levels of a major fraction of the proteome. We noticed that Dicer, which has several *let-7* 3′ UTR seeds, is one of the most strongly upregulated genes in the let-7b knockdown pSILAC (>4-fold) but not in the microarray data

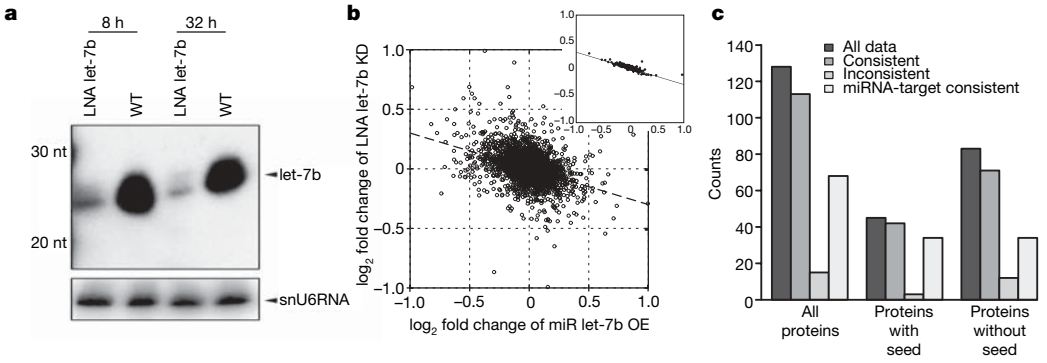


Figure 5 | Endogenous miRNA knockdown. **a**, Northern blotting demonstrates specific and stable let-7b knockdown by means of LNA. nt, nucleotide; WT, wild type. **b**, Scatter plot of changes in protein production in the let-7b overexpression (OE) versus the let-7b knockdown (KD) experiments. The inset shows the same data, averaged over bins of 20 genes. **c**, 'Consistent' refers to proteins with pSILAC fold changes that were upregulated in one experiment but down in the other, and 'inconsistent' refers to all other cases. 'miRNA-target consistent' is the subset of 'consistent' proteins that were downregulated in the overexpression experiment but upregulated in the knockdown.

(<1.3-fold). Therefore, Dicer is likely to be a direct translational target of let-7b. This raises the interesting possibility that let-7b regulates mature miRNA levels, which may in part explain our findings.

METHODS SUMMARY

HeLa cells were transfected with 100 nM synthetic dsRNAs designed to mimic mature endogenous miRNAs using DharmaFECT1 (Dharmacon) at 60–70% confluence, or with LNA-anti-let-7b (BioTez). Mock transfections were performed in the same way but without miRNAs. Eight hours post-transfection, cells were split into new dishes containing medium-heavy and heavy SILAC medium prepared as described[48] and incubated for 24 h until harvest. Corresponding protein and mRNA samples were always derived from the same transfection experiment. For the proteome analysis, miRNA/LNA-transfected cells and corresponding control cells were combined, lysed, and separated by SDS–PAGE. Gel lanes were cut into 15 slices, reduced, alkylated and trypsin-digested. Peptides were extracted and analysed by liquid chromatography–tandem mass spectrometry on a LTQ-Orbitrap hybrid mass spectrometer (Thermo Fisher). All samples were analysed in triplicate resulting in 45 mass spectrometry runs (5 days measurement time) per sample. Raw data files were processed with MaxQuant developed by J. Cox and M. Mann at the Max Planck Institute of Biochemistry (personal communication). False discovery rates were estimated using the target-decoy strategy[49] against an in-house-curated version of the IPI human protein database (version 3.37). In total, we identified 3,097,418 peptides (66,989 unique sequences) with average absolute mass accuracy of 0.65 p.p.m. We identified 4,961 unique proteins with at least two peptides each at a maximum false discovery rate of 1%. In individual experiments we only considered protein quantifications based on at least three peptide quantifications. Microarray analyses were performed with Human Genome U133 Plus 2.0 chips (Affymetrix), normalized by the standard rma()function (http://www.bioconductor.org) and annotated with the current NetAffx annotation file (http://www.affymetrix.com).

Received 8 April; accepted 3 July 2008.
Published online 30 July 2008.

1. Ambros, V. The functions of animal microRNAs. *Nature* **431,** 350–355 (2004).
2. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116,** 281–297 (2004).
3. Bushati, N. & Cohen, S. M. microRNA functions. *Annu. Rev. Cell Dev. Biol.* **23,** 175–205 (2007).
4. He, L. & Hannon, G. J. MicroRNAs: small RNAs with a big role in gene regulation. *Nature Rev. Genet.* **5,** 522–531 (2004).
5. Lai, E. C. miRNAs: whys and wherefores of miRNA-mediated regulation. *Curr. Biol.* **15,** R458–R460 (2005).
6. Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature Rev. Genet.* **9,** 102–114 (2008).
7. Shyu, A. B., Wilkinson, M. F. & van Hoof, A. Messenger RNA regulation: to translate or to degrade. *EMBO J.* **27,** 471–481 (2008).
8. Mathonnet, G. *et al.* MicroRNA inhibition of translation initiation *in vitro* by targeting the cap-binding complex eIF4F. *Science* **317,** 1764–1767 (2007).
9. Thermann, R. & Hentze, M. W. *Drosophila* miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature* **447,** 875–878 (2007).
10. Wakiyama, M., Takimoto, K., Ohara, O. & Yokoyama, S. Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. *Genes Dev.* **21,** 1857–1862 (2007).
11. Liu, J. Control of protein synthesis and mRNA degradation by microRNAs. *Curr. Opin. Cell. Biol.* **20,** 214–221 (2008).
12. Lim, L. P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433,** 769–773 (2005).
13. Bentwich, I. Prediction and validation of microRNAs and their targets. *FEBS Lett.* **579,** 5904–5910 (2005).
14. Hofacker, I. L. How microRNAs choose their targets. *Nature Genet.* **39,** 1191–1192 (2007).
15. Rajewsky, N. microRNA target predictions in animals. *Nature Genet.* **38** (Suppl), S8–S13 (2006).
16. Sethupathy, P., Megraw, M. & Hatzigeorgiou, A. G. A guide through present computational approaches for the identification of mammalian microRNA targets. *Nature Methods* **3,** 881–886 (2006).
17. Beitzinger, M. *et al.* Identification of human microRNA targets from isolated argonaute protein complexes. *RNA Biol.* **4,** 76–84 (2007).
18. Easow, G., Teleman, A. A. & Cohen, S. M. Isolation of microRNA targets by miRNP immunopurification. *RNA* **13,** 1198–1204 (2007).
19. Karginov, F. V. *et al.* A biochemical approach to identifying microRNA targets. *Proc. Natl Acad. Sci. USA* **104,** 19291–19296 (2007).
20. Hendrickson, D. G. *et al.* Systematic identification of mRNAs recruited to argonaute 2 by specific microRNAs and corresponding changes in transcript abundance. *PLoS ONE* **3,** e2126 (2008).
21. Aleman, L. M., Doench, J. & Sharp, P. A. Comparison of siRNA-induced off-target RNA and protein effects. *RNA* **13,** 385–395 (2007).
22. Vinther, J. *et al.* Identification of miRNA targets with stable isotope labeling by amino acids in cell culture. *Nucleic Acids Res.* **34,** e107 (2006).
23. Stefani, G. & Slack, F. J. Small non-coding RNAs in animal development. *Nature Rev. Mol. Cell Biol.* **9,** 219–230 (2008).
24. Mann, M. Functional and quantitative proteomics using SILAC. *Nature Rev. Mol. Cell Biol.* **7,** 952–958 (2006).
25. Ong, S. E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1,** 376–386 (2002).
26. Schwanhaeusser, B., Gossen, M., Dittmar, G. & Selbach, M. Global analysis of cellular protein translation by pulsed SILAC. *Proteomics* (in the press).
27. Milner, E., Barnea, E., Beer, I. & Admon, A. The turnover kinetics of major histocompatibility complex peptides of human cancer cells. *Mol. Cell. Proteomics* **5,** 357–365 (2006).
28. Pratt, J. M. *et al.* Dynamics of protein turnover, a missing dimension in proteomics. *Mol. Cell. Proteomics* **1,** 579–591 (2002).
29. Lam, Y. W., Lamond, A. I., Mann, M. & Andersen, J. S. Analysis of nucleolar protein dynamics reveals the nuclear degradation of ribosomal proteins. *Curr. Biol.* **17,** 749–760 (2007).
30. Cox, J. & Mann, M. Is proteomics the new genomics? *Cell* **130,** 395–398 (2007).
31. Cravatt, B. F., Simon, G. M. & Yates, J. R. III. The biological impact of mass-spectrometry-based proteomics. *Nature* **450,** 991–1000 (2007).
32. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science* **312,** 212–217 (2006).
33. Landgraf, P. *et al.* A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129,** 1401–1414 (2007).
34. Sood, P. *et al.* Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc. Natl Acad. Sci. USA* **103,** 2746–2751 (2006).
35. Nielsen, C. B. *et al.* Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* **13,** 1894–1910 (2007).
36. Grimson, A. *et al.* MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* **27,** 91–105 (2007).
37. Haley, B. & Zamore, P. D. Kinetic analysis of the RNAi enzyme complex. *Nature Struct. Mol. Biol.* **11,** 599–606 (2004).
38. Martinez, J. & Tuschl, T. RISC is a 5′ phosphomonoester-producing RNA endonuclease. *Genes Dev.* **18,** 975–980 (2004).
39. Schwarz, D. S. *et al.* Designing siRNA that distinguish between genes that differ by a single nucleotide. *PLoS Genet.* **2,** e140 (2006).
40. Vasudevan, S., Tong, Y. & Steitz, J. A. Switching from repression to activation: microRNAs can up-regulate translation. *Science* **318,** 1931–1934 (2007).
41. Doench, J. G., Petersen, C. P. & Sharp, P. A. siRNAs can function as miRNAs. *Genes Dev.* **17,** 438–442 (2003).
42. Stenvang, J. *et al.* The utility of LNA in microRNA-based cancer diagnostics and therapeutics. *Semin. Cancer Biol.* **18,** 89–102 (2008).
43. Wahlestedt, C. *et al.* Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc. Natl Acad. Sci. USA* **97,** 5633–5638 (2000).
44. Boyerinas, B. *et al.* Identification of let-7-regulated oncofetal genes. *Cancer Res.* **68,** 2587–2591 (2008).
45. Kuster, B., Schirle, M., Mallick, P. & Aebersold, R. Scoring proteomes with proteotypic peptide probes. *Nature Rev. Mol. Cell Biol.* **6,** 577–583 (2005).
46. Eulalio, A., Behm-Ansmant, I. & Izaurralde, E. P bodies: at the crossroads of post-transcriptional pathways. *Nature Rev. Mol. Cell Biol.* **8,** 9–22 (2007).
47. Nottrott, S., Simard, M. J. & Richter, J. D. Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nature Struct. Mol. Biol.* **13,** 1108–1114 (2006).
48. Ong, S. E. & Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nature Protocols* **1,** 2650–2660 (2007).
49. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods* **4,** 207–214 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Cox and M. Mann for early access to the MaxQuant software package, N. D. Socci for discussions, S. Schmidt, G. Born and N. Huebner for the hybridizations at the MDC microarray facility, C. Sommer for technical assistance, M. Huska and M. Andrade-Navarro for setting up the pSILAC website, P. Sharp for a CXCR4 luciferase construct, M. Peter for the IMP-1 reporters, and the Bundesministerium für Bildung und Forschung for funding mass spectrometry instrumentation. R.K. gratefully acknowledges a DAAD scholarship for research stays at the MDC. pSILAC and microarray data can be queried at http://psilac.mdc-berlin.de.

Author Contributions M.S. and N.R. conceived, designed and supervised the experiments. B.S. and N.T. performed the wet lab experiments. M.S., Z.F., R.K. and N.R. analysed genome-wide data. M.S., R.K. and N.R. interpreted the data. M.S. and N.R. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.S. (matthias.selbach@mdc-berlin.de) and N.R. (rajewsky@mdc-berlin.de).

Selbach Supplementary Methods

Methods

Cell Culture and SILAC media

HeLa cells obtained from LGC Promochem were cultivated at 37°C with 5% CO_2 and split every second or third day. SILAC media were essentially prepared as described previously[48]. Briefly, we used Dulbecco's Modified Eagle's Medium (DMEM) Glutamax lacking arginine and lysine (a custom preparation from Gibco) supplemented with 10% dialyzed fetal bovine serum (dFBS, Gibco). To prepare "heavy" (H) and "medium-heavy" (M) SILAC media we added 84 mg/l $^{13}C_6{}^{15}N_4$ L-arginine plus 40 mg/l $^{13}C_6{}^{15}N_2$ L-lysine or 84 mg/l $^{13}C_6$-L-arginine plus 40 mg/l D_4-L-lysine, respectively. Labeled amino acids were purchased from Sigma Isotec ($^{13}C_6$-L-arginine, $^{13}C_6{}^{15}N_4$ L-arginine and $^{13}C_6{}^{15}N_2$ L-lysine) and Cambridge Isotope Laboratories (D_4-L-lysine). "Light" (L) SILAC medium was prepared by adding the corresponding non-labeled amino acids (Sigma).

Synthetic miRNAs

Synthetic miRNAs designed to mimic mature endogenous miRNAs were purchased from Dharmacon as annealed, 2'-deprotected and desalted duplexes. The miRNA corresponding to miR-1 contained one mismatch in the duplex to facilitate activation of the sense strand[12,51] and was exactly designed as described in the study of Lim *et al.*[12]

RNA duplexes were synthesized as follows (sense 5'-3' / antisense 3'-5'):

miR-1, UGGAAUGUAAAGAAGUAUGUAA / AUAACUUACAUUUCUUCAUACA

miR-16, UAGCAGCACGUAAAUAUUGGCG / AUAUCGUCGUGCAUUUAUAACC

miR-30a, UGUAAACAUCCUCGACUGGAAG / CGACAUUUGUAGGAGCUGACCU

miR-155, UUAAUGCUAAUCGUGAUAGGGGU / ACAAUUACGAUUAGCACUAUCCC

let-7b, UGAGGUAGUAGGUUGUGUGGUU / CCACUCCAUCAUCCAACACACC

LNAs

LNAs (Locked Nucleic Acids) purchased from BioTez (Berlin, Germany) were designed to bind endogenous mature miRNAs with perfect complementarity. The following LNAs (5'-3') were used in this study:

LNA-anti-let-7b, aaccacacaacctactacctca

LNA-anti-miR-21, tcaacatcagtctgataagcta (used as control for LNA-anti-let-7b luciferase experiment)

Generation of luciferase reporter constructs

For pSILAC validation, luciferase reporters carrying 3' UTRs of genes found to be down-regulated by pSILAC upon specific microRNA overexpression were constructed. The 3' UTRs were PCR-amplified from HeLa cDNA (purchased from BioCat, Catalog No.: C1255811) and cloned into *Xho*I and *Not*I sites immediately downstream of the stop codon in the pRL-TK CXCR4 4x vector (a kind gift of Phil Sharp) coding for *Rr*-luc. The artificial CXCR4 4x target site had been removed by digestion beforehand. All constructs were checked by sequencing.

The 3' UTRs of the following genes were cloned:

c-Met (NM_000245)

RDH10 (NM_172037)

CAP1 (NM_006367)

TAGLN2 (NM_003564)

ADPGK (NM_031284)

MTX1 (NM_002455)

SLC25A1 (NM_005984)

ATP6V0A1 (NM_005177)

Primers (5'-3'):

Note that all primers contain flanking restriction sites (5'-3' FW primer = *Xho*I site; 5'-3' RW primer = *Not*I site) for site-directed insertion of the PCR product into the target vector.

CMET_FW 5'-CGGCTCGAGTGCTAGTACTATGTCAAAGCAA-3'

CMET_RW 5'-ATAGTTTAGCGGCCGCTGCATGATTTATCAGAACAACT-3'

RDH10_FW 5'-CGGCTCGAGGAATCTTTTTGTATGGAATATT-3'

RDH10_RW 5'-ATAGTTTAGCGGCCGCCAGTCATTTATAAAACTCCCCA-3'

TAGLN2_FW 5'-CGGCTCGAGTCCCACCCCAGGCCTTGCCC-3'

TAGLN2_RW 5'-ATAGTTTAGCGGCCGCCAAAAATGACAAATTCTTTA-3'

MTX1_FW 5'-CGGCTCGAGTTTGTCCTCACGCTCCCAAG-3'

MTX1_RW 5'-ATAGTTTAGCGGCCGCCAGTGTGAGTGGCTTTATTC-3'

SLC25A1_FW 5'-GCTCTAGAGCCTAGAGAGGCCGCAAGGG-3

SLC25A1_RW 5'-ATAGTTTAGCGGCCGCGCAACAGGATCCGGTTTATT-3'

CAP1_FW 5'-CGGCTCGAGGCGAAGTGCCACTGGGTTCT-3'

CAP1_RW 5'-ATAGTTTAGCGGCCGCCAAGTTTGGTATTAACTTTA-3

ADPGK_FW 5'-CGGCTCGAGGAAGATTCTTAGGGGTAATT-3'

ADPGK_RW 5'-ATAGTTTAGCGGCCGCCCTGAAATGTAAATTGTTTT-3'

ATP6V0A1_FW 5'-CGGCTCGAGGTCCCTGTGAGGGCCGTGTG-3'

ATP6V0A1_RW 5'-ATAGTTTAGCGGCCGCCCGGGGAAGTCAAACATACT-3'

Co-Transfection of synthetic miRNAs and 3' UTR reporter constructs

HeLa cells were co-transfected with synthetic miRNAs and different 3' UTR luciferase reporter constructs using Lipofectamine 2000 (Invitrogen) following the manufacturer's instructions (Lipofectamine™ 2000 siRNA - plasmid co-transfection protocol). The day before transfection, cells were seeded in 24-well plates in antibiotic-free "light" SILAC medium (1×10^5 cells/well). The following day the 80-90% confluent cells were transfected with 180 ng of the respective reporter plasmid and 20 ng pGL3 control

plasmid (Promega), synthetic miRNAs were co-transfected at a final concentration of 100 nM. DNA, RNA and Lipofectamine 2000 were diluted in serum-free DMEM. All transfections were performed in triplicate. Control transfections were performed with *miR-155* as a control since this miRNA did not significantly affect synthesis of the tested proteins. On the next day the medium was changed and cells were harvested 48 h post transfection in 100 µl 1X Passive Lysis Buffer (Promega) according to the manufacturer's instructions (Passive Lysis of Cells Cultured in Multiwell Plates). Cell lysates were cleared by centrifugation in a microcentrifuge for 5 min at 16,000 x g, 4°C.

Dual-Luciferase Assay

Dual-Luciferase Assays (Promega) were performed 48 h post transfection following the manufacturer's protocol (Technical Manual Dual-Luciferase Reporter Assay System) and detected with a MicroLumat Plus LB 96V luminometer (Berthold Technologies). Differing from the protocol, the amounts of cell lysate, LAR II and Stop & Glo Reagent were all divided by two. Each cell lysate was measured three times (3 technical replicates) in a white 96-well plate (nunc). Renilla luciferase activity of the pRL-TK reporter constructs was normalized to the activity of the firefly luciferase of the pGL3 control plasmid (Promega) which served as internal transfection control. The psiCHECK-2 IMP-1 wildtype (wt) and mutated (mt) reporter constructs (a kind gift of Marcus Peter)[44] were transfected similarly with the difference of only using 60 ng plasmid DNA per 24-well. These constructs carry both the readout (Renilla) and the control reporter luciferase gene (Firefly) and were thus only co-transfected with a synthetic miRNA or LNA at a final concentration of 100 nM.

Errors bars were calculated as follows: First, relative errors of the three biological replicates of the respective reporter and its corresponding control were computed. Second, the relative error of the reporter and the control were added up according to the law of error propagation. The resulting relative error was used to calculate absolute errors of the normalized expression values. To estimate the pSILAC error (Fig. 3C) we calculated the standard deviation of all protein quantification for two biological replicates of the *miR-1* transfection experiment (shown in Fig. 1E) after removing 5% outliers. Error bars are shown as +/- two standard deviations.

Transfection and pulsed SILAC labeling

HeLa cells were transfected with synthetic miRNAs (Dharmacon) or LNAs (BioTez) using DharmaFECT1 (Dharmacon) according to the manufacturer's instructions. Cells were plated in 15 cm dishes in antibiotic-free "light" SILAC medium one day before transfection and incubated at 37°C with 5% CO_2 overnight. On the day of transfection, cells were 60-70% confluent. Synthetic miRNAs (LNAs) were used at a final concentration of 100 nM. Control transfections were carried out in parallel under the same conditions using water instead of the miRNA (LNA) (mock-transfection). At 8 h post transfection, cells were trypsinized. 2/5 of the cell suspension was used for RNA isolation (8 h time point). The remaining cells were transferred into two new 10 cm dishes each (3/10 of all cells per plate). One of the two plates containing miRNA (LNA) transfected cells was transferred to "medium-heavy" SILAC medium and one of the mock-transfected cells to "heavy" SILAC medium for pulsed SILAC labeling. The two remaining plates were kept in normal light medium for mRNA analysis. After 24 h, cells of the two SILAC plates were scraped off, combined and spun down (10 min, 600 x g, 4°C) for protein analysis. The corresponding plates were harvested for total RNA isolation (32 h time point).

Sample preparation for mass spectrometry

Combined cell pellets were lysed in 750 µl RIPA buffer consisting of 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% Triton-X100, 1% Sodium deoxycholate and 0.1% SDS for 20 min on ice. The lysates were cleared by centrifugation for 10 min (14,000 rpm at 4°C) and transferred to fresh tubes. 1-D discontinuous SDS-PAGE was performed with the whole-cell lysate using NuPAGE Novex 4 to 12% gradient gels (Invitrogen) under reducing conditions according to the manufacturer's instructions. The gel was cut into 15 slices and each slice was subjected to reduction, alkylation and in-gel digestion with sequence grade modified trypsin (Promega) according to standard protocols[52]. After in-gel digestion peptides were extracted and desalted using StageTips[53] before analysis by mass spectrometry.

Mass spectrometry

Peptide mixtures were analyzed by online LC-MS/MS on a high performance hybrid mass spectrometer (LTQ-Orbitrap, Thermo Fisher). Reversed phase chromatography was performed with the Agilent HPLC 1200 nanoflow system using self-made fritless C_{18} microcolumns (75 µm ID packed with ReproSil-Pur C_{18}-AQ 3-µm resin, Dr. Maisch GmbH, Germany)[54] directly connected to the electrospray ion source of the LTQ-Orbitrap. All gel slices were analyzed three times resulting in 45 LC-MS/MS runs per sample (3 x 15 slices). For each run, 1/3 of the peptide mixture was injected and separated with a 10 to 60% acetonitrile gradient (155 min) in 0.5% acetic acid at a flow rate of 200 nl/min. The LTQ-Orbitrap was operated in the data dependent mode with a full scan in the Orbitrap and five consecutive MS/MS scans in the LTQ. The precursor ion scan/survey MS spectra (m/z 300–1700) were acquired in the Orbitrap part of the instrument (resolution R = 60,000; target value of 1 x 10^6). The five most intense ions were isolated (target value of 5,000; monoisotopic precursor selection enabled) and fragmented in the LTQ part of the instrument by collision induced dissociation (normalized collision energy 35%; wideband activation enabled). Ions with an unassigned charge state and singly charged ions were rejected. Former target ions selected for MS/MS were dynamically excluded for 60 s. Total cycle time for one full scan plus five MS/MS scans was approximately 2 s. Total MS run time was 47 days and 8,821,728 spectra were acquired.

Processing of mass spectrometry data

Identification and quantification of proteins was carried out with version 1.0.7.3 of the MaxQuant software package developed by Jürgen Cox and Matthias Mann (Max Planck Institute of Biochemistry) which is described in more details elsewhere[49]. Briefly, isotope clusters and SILAC triplets were extracted, re-calibrated and quantified in the raw data files with Quant.exe (medium labels: Arg6 and Lys4, heavy labels: Arg10 and Lys8; maximum of three labeled amino acids per peptide; polymer detection enabled; top 6 MS/MS peaks per 100 Da). The generated peak lists (msm files) were submitted to a MASCOT search engine (version 2.2, MatrixScience) and searched against an in house curated concatenated target-decoy database[50] of forward and reversed proteins in the IPI human protein database (version 3.37) supplemented with common contaminants (e.g.

trypsin, BSA). We required full tryptic specificity, a maximum of two missed cleavages and a mass tolerance of 0.5 Da for fragment ions. The initial mass accuracy cut-off on the parent ion was 7 ppm but subsequently narrowed down by filtering based on hits to reversed peptides in the target-decoy database (see below). Oxidation of methionine and acetylation of the protein N-terminus were used as variable modifications, carbamidomethylation of cysteine as a fixed modification. Filtering of putative MASCOT peptide identifications, assembly of proteins and re-quantification was performed with Identify.exe (part of MaxQuant). We required a minimum peptide length of 6 amino acids and a minimum of two peptides per protein group (with at least one of the two being unique in the database). False discovery rates were estimated based on matches to reversed sequences in the concatenated target-decoy database. We required a maximum false discovery rate of 1% at both the peptide and the protein level. With these thresholds we identified 3,097,418 peptides (66,989 unique peptide sequences) from 6,432,045 submitted MS/MS spectra. Average absolute mass accuracy of identified peptides was 0.65 ppm. Peptides were assigned to protein groups (that is a cluster of a base protein plus additional proteins matching to a subset of the same peptides). Protein groups containing matches to proteins from the reversed database or contaminants were discarded. Overall we identified 4,962 protein groups as the HeLa cell proteome. To quantify changes in protein production we calculated the median of all H/M peptide ratios using only unique peptides and non-unique peptides assigned to the protein group with the highest number of peptides ("Occam's razor" peptides). For subsequent data analysis we only considered protein quantifications based on at least three independent H/M ratio measurements leading to 3,000-3,500 quantified protein groups in individual samples. Reproducibility (Fig. 1E) was checked by performing two completely independent *miR-1* transfection experiments on different days. Both samples were processed and analyzed by mass spec (each on 15 slices measured in triplicates). 2,287 proteins were identified and quantifiable in both samples according to our quantification criteria after removal of 5% outliers.

RNA isolation

Total RNA was extracted using Trizol Reagent (Invitrogen) following the manufacturer's protocol with slight modifications: Cells were not lysed directly in the culture dish as

cells had to be split into new dishes. Instead, for the first cell harvest time point, cells were washed once with 1X D-PBS (Gibco) and trypsinized with 3 ml of 0.05% Trypsin-0.53 mM EDTA * 4 Na (Gibco) per 15 cm dish. The reaction was stopped by adding 7 ml DMEM Glutamax lacking arginine and lysine supplemented with 10% dFBS (Gibco). 4 ml of the 10 ml cell supsension, i.e. 2/5 of the cells, were collected in a falcon tube. Cells were pelleted by centrifugation for 5 min at 300 x g, 4°C, in a Heraeus Multifuge 3 S (Heraeus). The supernatant was aspired, the pellet washed with 1X D-PBS and centrifuged as described. The supernatant was carefully removed and the cells were lysed by adding 1 ml Trizol Reagent. To ensure homogenization, the cell lysate was passed through a 20G needle 8-10 times, total RNA was isolated as described in the protocol. RNA pellets were resuspended in 20 µl RNase-free sterile water, RNA quantity was assessed spectrophotometrically using the NanoDrop ND-1000 UV-VIS Spectrophotometer (Thermo Fisher). For the second cell harvest time point, cells growing in 10 cm dishes were treated the same way but not split into new dishes.

Northern Blotting

Total RNA was isolated from LNA-anti-let-7b and mock-transfected HeLa cells with Trizol Reagent (Invitrogen) as described above. Briefly, 15 µg total RNA per lane and a radioactive labeled RNA marker (Decade marker, Ambion) were resolved on a 15% denaturing polyacrylamide gel and transferred onto a Hybond-N+ membrane (Amersham, GE Life Sciences) at 700 mA for 1 h in a Trans-Blot Semi-Dry Transfer Cell (BioRad). The blot was cross-linked using a Stratalinker (Stratagene) and prehybridized for 1 h at 50°C in hybridization buffer (5X SSC, 20 mM Na_2HPO_4 pH 7, 1% SDS, 1X Denhardt's solution, 10 mg/µl sonicated salmon sperm DNA) being followed by an overnight incubation at 50°C in hybridization buffer containing the 5'-^{32}P-labeled probe. On the next day the blot was washed twice with 5X SSC and 1% SDS at 50°C for 10 min each followed by a third 10 min wash with 1X SSC and 1% SDS. Exposition was performed on a Kodak BioMax MR film at -80°C for six days. Before reprobing the blot with the labeled snU6RNA probe to check for equal loading, the blot was stripped by incubating 3X in a 1% SDS-solution for 10 min at 80°C. Hybridization of the snU6RNA probe was done the same way using 65°C as hybridization temperature and exposing only one day on film. Probes were generated by end-labeling 30 pmol of DNA oligonucleotide (Sigma)

complementary to let-7b or snU6RNA with Optikinase (Usb) and 3 µl of [γ-^{32}P] ATP (3000 Ci/mmol, PerkinElmer) as described by the manufacturer. Labeled probes were purified with MicroSpin G-25 columns (GE Healthcare) and added to the hybridization solution. The 5'-^{32}P-radiolabeled oligodeoxynucleotide probes (5'-3') were:

oligo let-7b, AACCACACAACCTACTACCTCA

oligo snU6RNA, TATGGAACGCTTCACGAATTTGCGTGTCAT

Microarrays

Microarrays were prepared according to the "GeneChip Expression Analysis Technical Manual" from Affymetrix using the GeneChip One-Cycle Target Labeling and Control Reagents kit (Affymetrix), following the manufacturer's protocol. Starting material for each array were 2 µg of total RNA. Fragmented and biotinylated cRNA was hybridized to Human Genome U133 Plus 2.0 Arrays using the GeneChip Hybridization Oven 640. Washing and staining of the probe arrays was performed in the GeneChip Fluidics Station 450. After completion of the wash protocols, arrays were scanned in the GeneChip Scanner 3000 7G. The applied GeneChip operating software was version 1.4.

Microarray data analysis

The output of microarrays was normalized by the standard rma() function from the Bioconductor R-library (www.bioconductor.org)[55]. To annotate Affymetrix probe sets to Refseq identifiers, the current NetAffx Annotation file was downloaded from the Affymetrix website (http://www.affymetrix.com). For the transcript (RefSeq) with multiple probes, the average logarithm expression values for all corresponding probes were taken. Fold-changes were defined as differences between the intensities of misexpressions and controls (log2 ratios). The mock-transfected control corresponding to the same miRNA transfection experiment was used where applicable (three control samples for *miR-1* and 2 samples for *miR-30a*). These were also used to calculate the correlation of microarrays and qRT-PCR (for *miR-1*). For the other miRNA transfection experiments (*miR-155*, *miR-16*, and *let-7b*) we used the median of three controls or two controls taken at the 8 h and 32 h time point, respectively.

Quantitative Reverse Transcription PCR

To validate the results of the microarray analysis, the expression of 23 genes upon *miR-1* over-expression was reanalyzed via 2-step quantitative reverse-transcription polymerase chain reaction (qRT-PCR). We used three different samples of total RNA derived from three *miR-1* over-expression experiments (two of them harvested 8 h post transfection, one harvested 32 h). The exact same samples were used for both microarray analysis and qRT-PCR. Single-stranded cDNA was generated by reverse transcription of total RNA in a 20 µl reaction volume using SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. Each 20 µl reaction contained 1.5 µg of total RNA, 250 ng of random primers (Invitrogen), 1 µl of 10 mM dNTP Mix (10 mM each dATP, dGTP, dCTP and dTTP), 4 µl 5X First Strand Buffer, 2 µl 0.1 M DTT, 1 µl RNasin (Promega), 1 µl of SuperScript III Reverse Transcriptase and RNase-free water. Prior to the qPCR, cDNA was diluted 1:60. For each gene and sample, qPCR was run in triplicates using combinations of primer pairs and TaqMan probes targeting mRNA sequences of the genes listed in table 1. Primer Pairs and TaqMan probes were designed using the Primer Express Software (Applied Biosystems) and were purchased from BioTez. For 21 of the 23 genes we were able to design primers spanning two exons, thus eliminating the possiblitity of genomic DNA amplification. Per 384 well, 5 µl of diluted cDNA and 10 µl 2X TaqMan Universal PCR Master Mix (Applied Biosystems) were used in a 20 µl reaction. Primers were added to a final concentration of 500 nM, the TaqMan probe to a concentration of 200 nM. The increase in reporter signal was captured in real time with the ABI Prism 7900HT Sequence Detection System (Applied Biosystems), the program starting with 2 min 50°C, being followed by a 10 min-denaturation at 95°C, 40 cycles of 95°C for 15 s each, concluding with 1 min at 60°C. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), shown to remain constant upon *miR-1* over-expression by microarray analysis, was chosen to serve as internal control for the normalization of all qPCR products. Relative quantification results were calculated according to the ddCt method. Primers and probes for qRT-PCR were as follows (Gene name, RefSeq ID, forward primer, reverse primer, TaqMan probe):

• ADAR, NM_001111, CTGGCCGCCATCATTATGA, CCTTTTAGGCTGAGAGAATCTCCTTT, FAM-CGTCAGCTTGGGAACAGGGAATCG-TAMRA

• ADPGK, NM_031284, TCACGACATTGCCCAGGTT, AATTAAAGCTGCATTTCCTCCTACA, FAM-TCAGAGTTCCCAGGAGCCCAGCACT-TAMRA

• ATP6VOA1, NM_005177, CTGGCGACTACGTGCACAAGT, CGGAACCCTTCACAGATTTTCT, FAM-TCATTTTCTTCCAAGGCGATCAG-TAMRA

• AXL, NM_001699, TCTGCATGAAGGAATTTGACCAT, TCTCGTTCAGAACCCTGGAAA, FAM-CCAACGTCATGAGGCTCATCGGTG-TAMRA

• BRI3BP, NM_080626, GCCGCTTCTTCTGGATCGT, GCTCGCCCTCGTACTTGTG, FAM-TCCTGTTTTCCATGTCCTGCGTGTACA-TAMRA

• CAP1, NM_006367, ACTTGGCCTGGTATTCGATGA, CTTTACCCATTACCTGAACTTTGACA, FAM-TGGTGGGCATTGTGGAGATAATCAACAGTAAG-TAMRA

• CDCP1, NM_022842, TGGTTCCACCCCAGAAATGT, CTCGATGATGCACAGACGTTTTAT, FAM-CGGCTTCAGCATTGCAAACCGC-TAMRA

• DHX15, NM_001358, GATGGTGTGGTGTTTGTGATTGAT, CACCAAAAGGGACTCAACTCTGAT, FAM-CTGGATTTGCGAAACAGAAGGTCTACAATCCTC-TAMRA

• EGFR, NM_005228, GCGTCTCTTGCCGGAATGT, GGCTCACCCTCCAGAAGGTT, FAM-AGCCGAGGCAGGGAATGCGTG-TAMRA

• GAPDH, NM_002046, CTCTGCCCCCTCTGCTGAT, TGATGATCTTGAGGCTGTTGTCA, FAM-TTCGTCATGGGTGTGAACCATGAGAAGT-TAMRA

• G6PD, NM_000402, CAAGAAGCCGGGCATGTT, GTAGGCGTCAGGGAGCTTCAC, FAM-TGGACCTGACCTACGGCAACAGATACAAGA-TAMRA

• LRRC8A, NM_019594, TACCCCAACTCCACCATTCTG, CAAACCAGTGCAGTCGGTTCT, FAM-CCTGGACCGGCACCAGTACAACTACG-TAMRA

• MTX1, NM_002455, GAGAGGTCATCTCAGTTCCACACA, TGCCGAGCTGACAGATCATAA, FAM-ATCACCCACCTTCGAAAAGAGAAGTACAATGC-TAMRA

• NOTCH2, NM_024408, TGCTGTTGTCATCATTCTGTTTATTATTC, CTCACGACGCTTGTGATTGC, FAM-ATGGCAAAACGAAAGCGTAAGCATGG-TAMRA

• OAT, NM_000274, TTATGCCGGGATTCGACATC, GCCACATTTGGATCCTGAAGA, FAM-ATCTGCCCGCACTGGAGCGTG-TAMRA

• PTPLB, NM_198402, CTGGCCACGGCGTACCT, GGACCAGACCAACCGCTATAAC, FAM-ATGTGGTGATGACAGCCGGGTGG-TAMRA

• SFRS9, NM_003769, GTGCCCTTCGCCTTCGT, ATAACCATTTCTTCCATAAATAGCATCCT, FAM-CGCTTCGAGGACCCCCGAGATG-TAMRA

• SFXN1, NM_022754, CGTAGCAACAGCTCTAGGACTCAAT, ACGGCAGCAAAGGGAACA, FAM-CCAAGCATGTCTCACCACTGATAGGACGTT-TAMRA

• SLC25A1, NM_005984, CCAGGCCATCCGCTTCTT, AGAGGGTTCATGGGCTTGTTG, FAM-CCTGCGCAACTGGTACCGAGGG-TAMRA

• SNX6, NM_021249, AAGATGAAACAGGAACTGGAAGCT, GCCACACGACACAGGAACAC, FAM-TGGCAATATTCAAGAAGACAGTTGCGATGC-TAMRA

• TAGLN2, NM_003564, ACTGTGGACCTCTGGGAAGGA, CCCATCATCTCGGGCTACTG, FAM-ACGCTGATGAATCTGGGTGGGCTG-TAMRA

• THBS1, NM_003246, CATCCGCAAAGTGACTGAAGAG, CTGTACTGAACTCCGTTGTGATAGC, FAM-TGAGCTGAGGCGGCCTCCCCTA-TAMRA

• TPM3, NM_152263, GAGATCGGTAGCCAAGCTAGAAAA, CTAATGGCCTTGTACTTCAGTTTCTG, FAM-ACAATTGATGACCTGGAAGATGAGCTCTATGC-TAMRA

• TWF1, NM_002822, ATATTCATTGCATGGTCTCCAGATC, AAATTCCTTCTTCAGAGTTGCTCTTG, FAM-TCTCATGTTCGTCAAAAAATGTTGTATGCAGC-TAMRA

Western blotting

In order to validate selected miRNA targets identified by the pulsed SILAC approach, HeLa cells were transfected with either one of the synthetic miRNAs or the LNA targeting *let-7b* as described above (see '*Transfection and pulsed SILAC labeling'*). 32 h post transfection cells were harvested, proteins were isolated and separated by SDS-PAGE. Gels were blotted onto PVDF membranes using the iBlot dry blotting system (Invitrogen) according to the manufacturer's instructions. Unspecific binding sites were blocked with Roti-Block (Carl Roth, Germany) blocking reagent for 1 h at room temperature (RT). Primary antibodies against CEBPβ (sc-150), Kras (sc-30), Annexin2 (sc-48397), Tropomyosin (sc-28543), Twinfilin-1 (sc-51241), FGF-2 (sc-74412), Integrin α2 (sc-53353), α-Adducin (sc-25731), EGFR (sc-03), PICALM (sc-6433), SNX6 (sc-50373), IMP-1(sc-21026) and cMet (sc-161) were purchased from Santa Cruz Biotechnologies or Sigma-Aldrich (β-Actin, A5441). All primary antibodies were applied at a 1:1,000 dilution in TTBS (140 mM NaCl, 25 mM Tris-HCl, 0.1% Tween 20, pH 7.4) over night at 4°C. Blots were washed 3x in TTBS and incubated either with an anti-mouse, an anti-rabbit or an anti-goat secondary antibody (all from Amersham) conjugated to horseradish peroxidase diluted 1:3,000 in TTBS for 1 h. After three more washing

steps in TTBS the bound secondary antibodies were detected with the Western Blot Chemiluminescence Reagent Plus for ECL immunostaining (PerkinElmer).

Analysis of transfection efficiency

The BLOCK-iT fluorescent oligo (Invitrogen) is a fluorescein-labeled, non-targeted dsRNA oligomer allowing for visual monitoring of transfection efficiency. Cells were transfected with BLOCK-iT fluorescent oligos as described in '*Transfection and pulsed SILAC labeling*'. 8 h post transfection cells were washed with 1X D-PBS (Gibco) and fixed with 4% paraformaldehyde (PFA) in D-PBS. Transfection efficiency was assessed by comparing the amount of transfected, fluorescing cells with non-transfected cells using epifluorescence microscopy (Leica DM-R).

Mapping protein identifiers to transcripts

Protein identifiers were first mapped to NCBI Entrez Gene gene numbers by the IPI cross-reference file (version 3.37) from EBI database (http://www.ebi.ac.uk). The gene2refseq file (downloaded on 28th of Sep, 2007) from the NCBI database (ftp://ftp.ncbi.nlm.nih.gov) was used to map NCBI Entrez Gene gene numbers to Refseq identifiers. For each protein group, the first protein identifier in a group was assigned to corresponding Refseq identifiers. If the first protein did not have any corresponding Refseq identifiers the mapping was done for the second protein in the protein group. Context features such as 8mers, M8-7mers, A1-7mer, 6mers and mismatches at positions 9 to 11 (Fig. 2) were assigned to proteins (with their corresponding protein fold-changes) using the mapping to transcripts. If a protein group had several mapped RefSeq identifiers, with corresponding context features and mRNA fold-changes, each was counted separately.

Human 3' UTR, 5' UTR and CDS sequences based on human reference sequence (NCBI Build 36.1) were extracted from UCSC Genome Browser (http://genome.ucsc.edu). In order to assign 3' UTR characteristics (number of seeds, number of conserved seeds) to unique protein identifiers we used the maximum number of seeds from all transcript identifiers mapped to the protein. This mapping slightly overestimates the count of seed numbers. Mapping wherein the seed number is taken as the rounded median of the seed numbers of corresponding transcripts was also tested and

the results were unchanged. Conserved seed number for a microRNA was assigned to a protein group in a similar way.

Identifying 3' UTR motifs correlated with changes in protein production

We used a linear regression model[34] to identify significant motifs in 3' UTRs which correlate best with the global changes in protein synthesis. Protein identifiers were mapped to mRNA identifiers as described. If a protein is mapped to multiple mRNAs, we randomly assigned one of the mRNAs to it.

Seed conservation

We used mammalian orthologous 3' UTR alignments constructed previously[56]. Conserved seeds were defined as being present at the same position in an alignment and identical in human, chimp, rat, mouse, and dog. In cases where one of the chimp, rat, or mouse sequences was missing, the 3' UTR alignment was still kept but discarded if more than one species was not represented.

Target prediction comparisons

The predictions of PicTar[56,57] were downloaded from http://pictar.bio.nyu.edu/. TargetScanS[58] predictions were downloaded from UCSC Table Browser (http://genome.ucsc.edu). Since predictions of TargetScanS are based on miRNA families, all predictions of the corresponding family were used for each miRNA. The predictions of PITA[59] were downloaded from http://genie.weizmann.ac.il/pubs/mir07/mir07_data.html. As PITA predicts thousands of targets for each miRNA we selected the top 600 or top 1,000 predictions based on their score to ensure that the number of predictions is comparable with the other algorithms. rna22 predictions[60] based on 3' UTRs and 5' UTRs were obtained from http://cbcsrv.watson.ibm.com/rna22_download.html. miRbase predictions (Version 5) [61] were downloaded from http://microrna.sanger.ac.uk. Diana-MicroT (Version 3.0, http://microrna.gr/) [M. Maragkakis, P. Alexiou, A.G. Hatzigeorgiou, unpublished] and miRanda predictions[62,63] were obtained directly from the authors. The medium-confidence predictions of Diana-MicroT and all predictions of miRanda were taken. Since proteomic data was mapped to Refseq identifiers, for those predictions which are not based on Refseq identifiers, we mapped them to Refseq via Biomart (Version 0.6,

http://www.biomart.org). For TargetScanS, which requires perfect miRNA seeds from position 2 to position 7, the mapped Refseq identifiers that contained imperfect seeds were removed. Predictions from different algorithms for each of the five microRNAs in the dataset were mapped to the proteomic data using the mapping of protein identifiers to transcripts as described above. If a protein could be mapped to multiple mRNAs, we used all of them with the same protein log2 fold changes. The lists of transcripts with a corresponding seed were constructed in a similar way. The fraction of down-regulated transcripts for a given cut-off (-0.1 log2 fold change) was computed for each prediction algorithm and reported in the histogram for pooled proteomic data. Some of the predicted targets are not expressed in Hela cells or escaped detection for technical reasons. Therefore, the correlation of target prediction with our experimental data is only valid for the subset of targets we could identify.

[51] Khvorova, A., Reynolds, A., and Jayasena, S. D., Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209-216 (2003).

[52] Shevchenko, A. et al., In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 1, 2856-2860 (2006).

[53] Rappsilber, J., Ishihama, Y., and Mann, M., Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* 75, 663-670 (2003).

[54] Ishihama, Y., Rappsilber, J., Andersen, J. S., and Mann, M., Microcolumns with self-assembled particle frits for proteomics. *J Chromatogr A* 979, 233-239 (2002).

[55] Gentleman, R. C. et al., Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5, R80 (2004).

[56] Krek, A. et al., Combinatorial microRNA target predictions. *Nat Genet* 37, 495-500 (2005).

[57] Lall, S. et al., A genome-wide map of conserved microRNA targets in C. elegans. *Curr Biol* 16, 460-471 (2006).

[58] Lewis, B. P., Burge, C. B., and Bartel, D. P., Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120, 15-20 (2005).

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

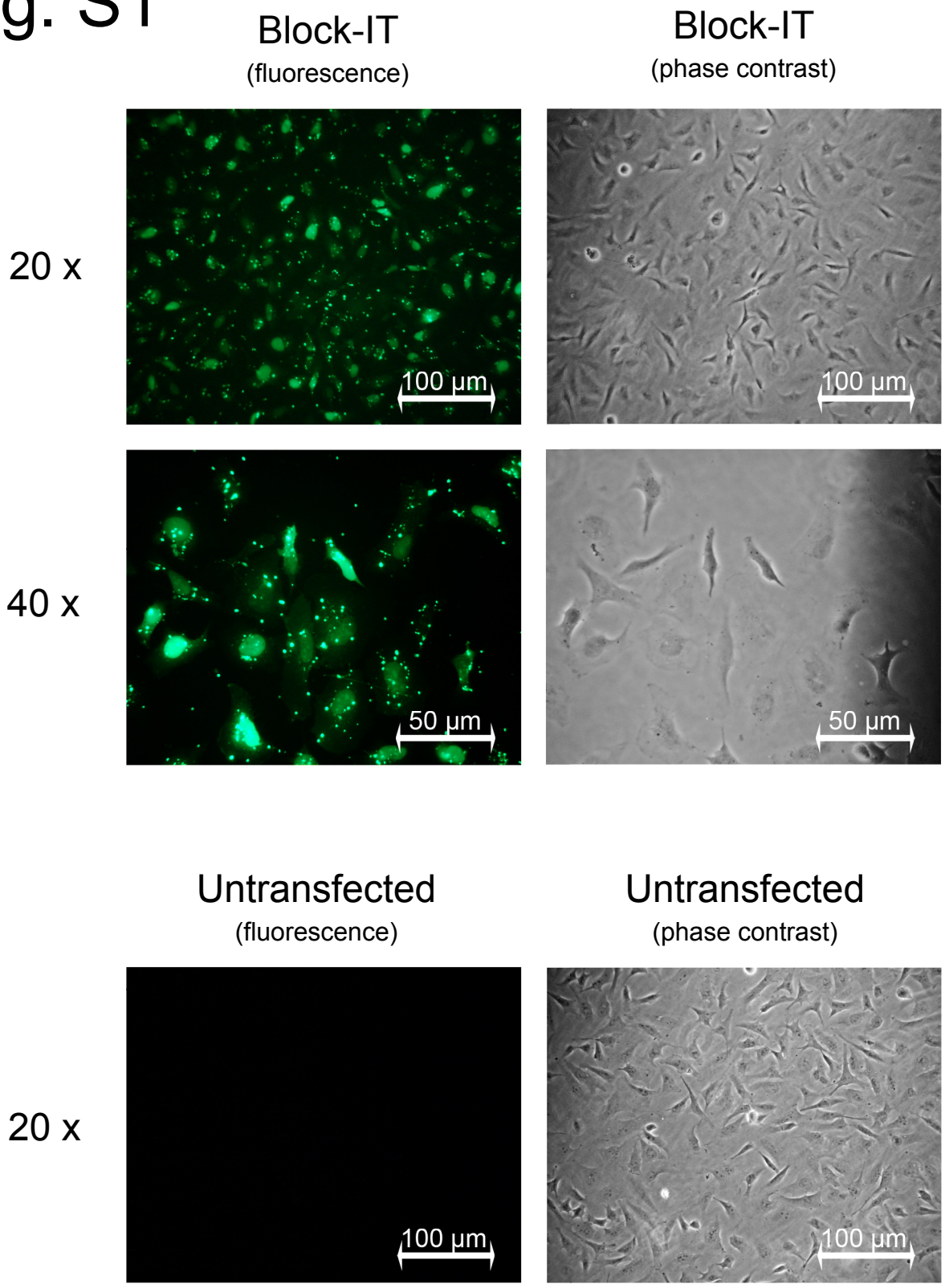
[60] Miranda, K. C. et al., A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* 126, 1203-1217 (2006).

[61] Griffiths-Jones, S. et al., miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 34, D140-144 (2006).

[62] Betel, D. et al., The microRNA.org resource: targets and expression. *Nucleic Acids Res* 36, D149-153 (2008).

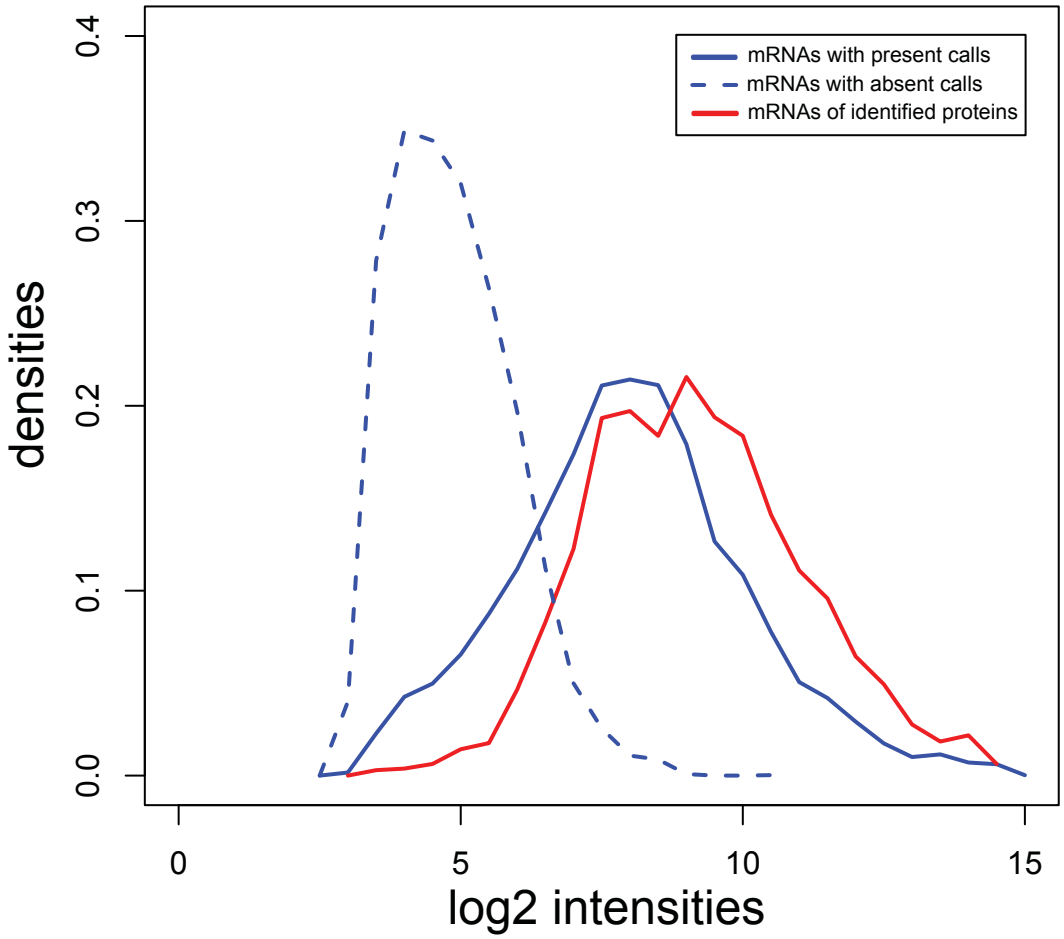
[63] John, B. et al., Human MicroRNA targets. *PLoS Biol* 2, e363 (2004).

Fig. S1



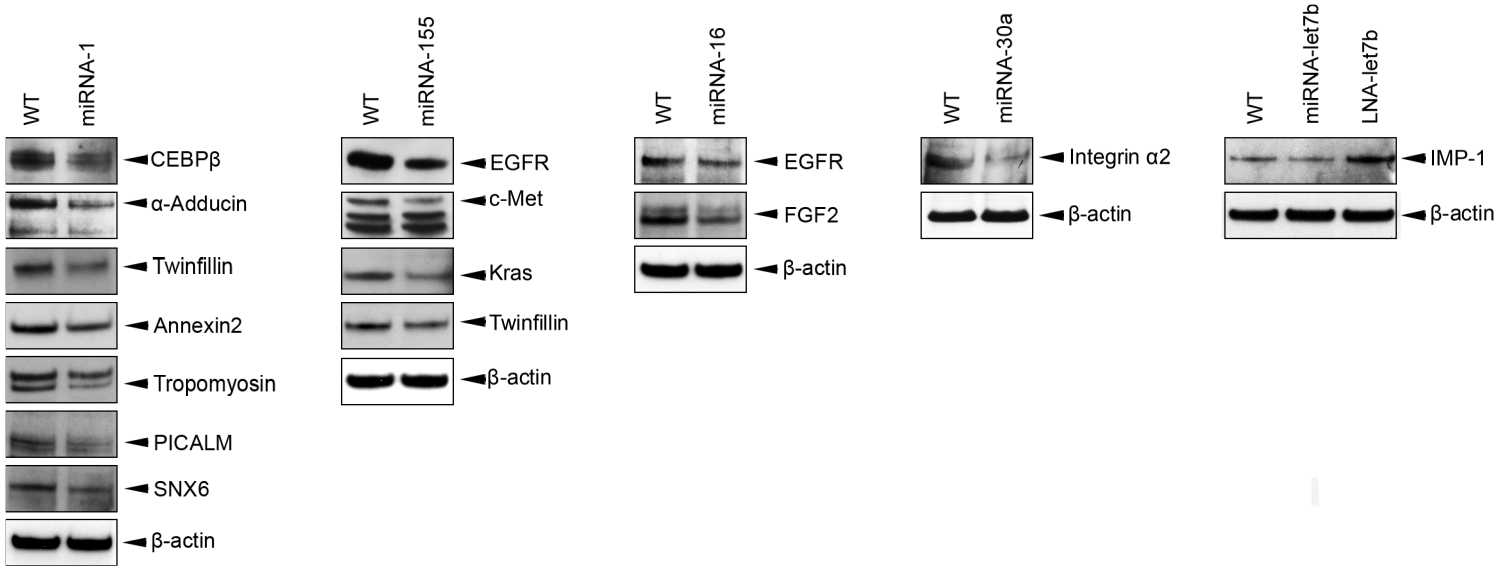
Analysis of transfection efficiency. HeLa cells were transfected with fluorophore-conjugated dsRNA (Block-IT, Invitrogen) according to our miRNA transfection protocol. Phase contrast images are presented as a reference.

Fig. S2



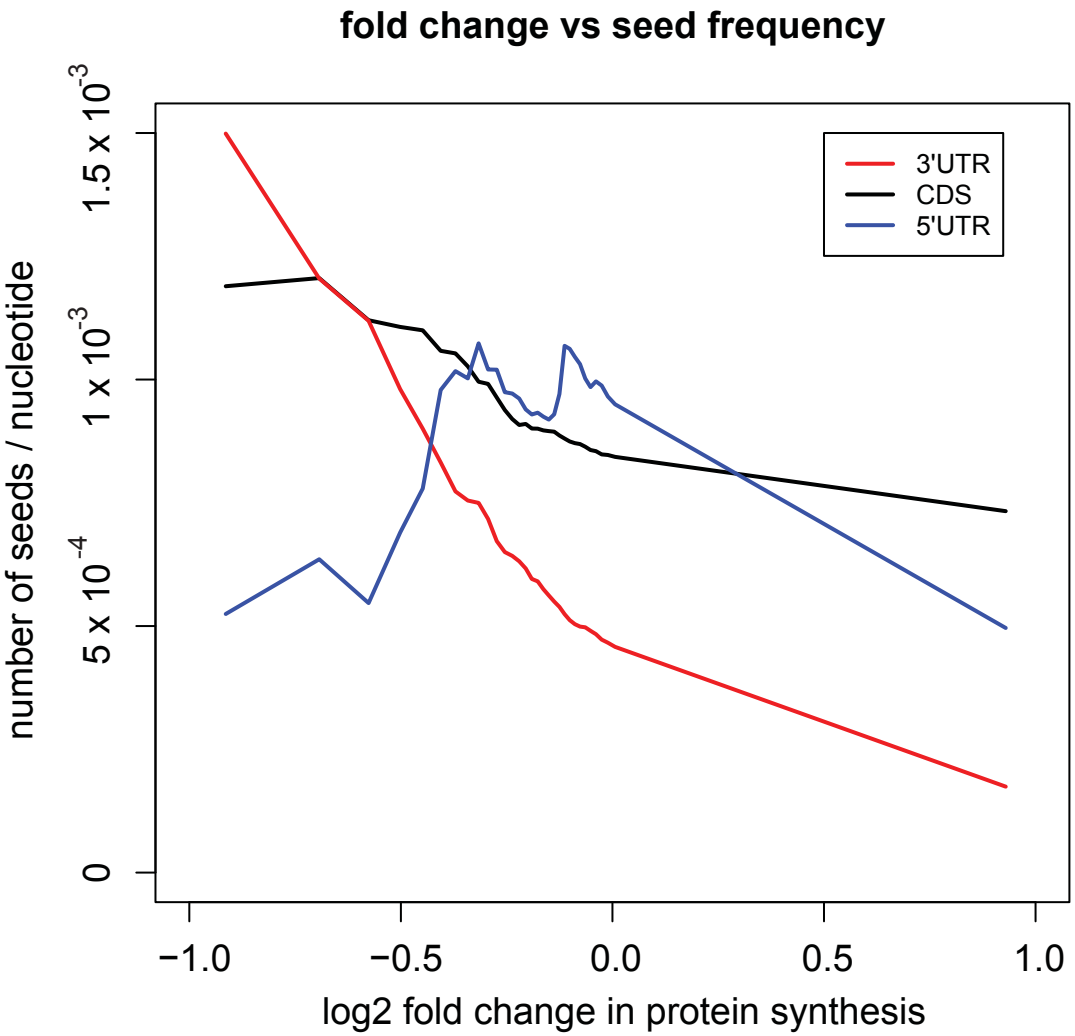
Dynamic range of detection. Signal intensities for all 29,729 mRNAs measured in a microarray experiment (blue) and the subset of 4778 mRNAs for which we could quantify corresponding proteins (red) are shown. Solid and dashed blue lines designate mRNAs with present or absent calls, respectively. Normalizing with the bioconductor gcrma() function and using a different probeset annotation (Dai M, et al., Nucleic Acids Res. 2005;33(20):e175) produced similar results.

Fig. S3



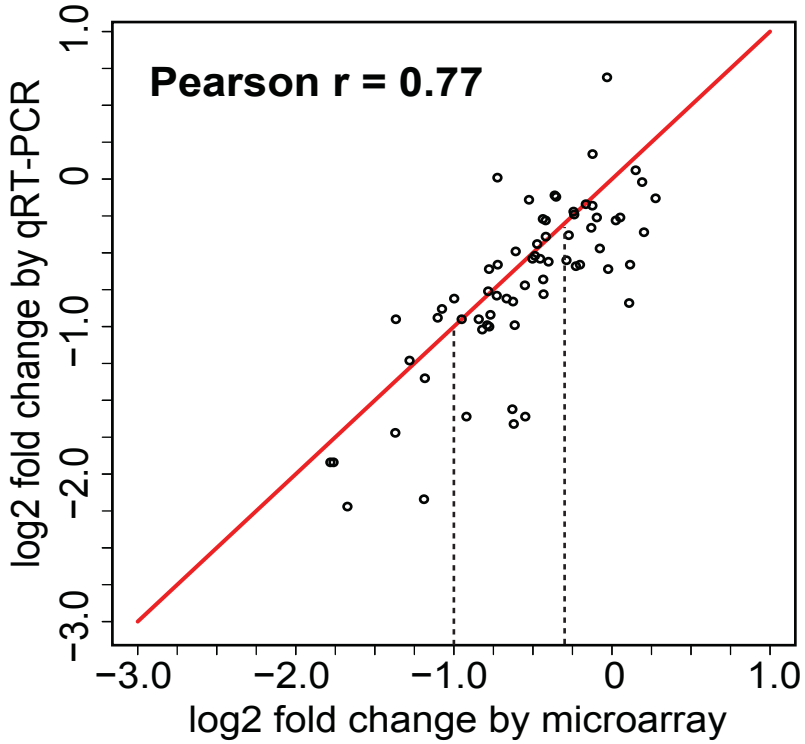
Verification of targets identified by pSILAC with western blots. HeLa cells were transfected with the miRNA, LNA or control treated and analysed by western blotting with antibodies against c-Met, EGFR, α-Adducin, CEBPβ, Twinfilin, Annexin2, Tropomyosin, PICALM, SNX6, Kras, FGF2, Integrin-α2 and IMP-1. β-Actin served as a loading control.

Fig. S4



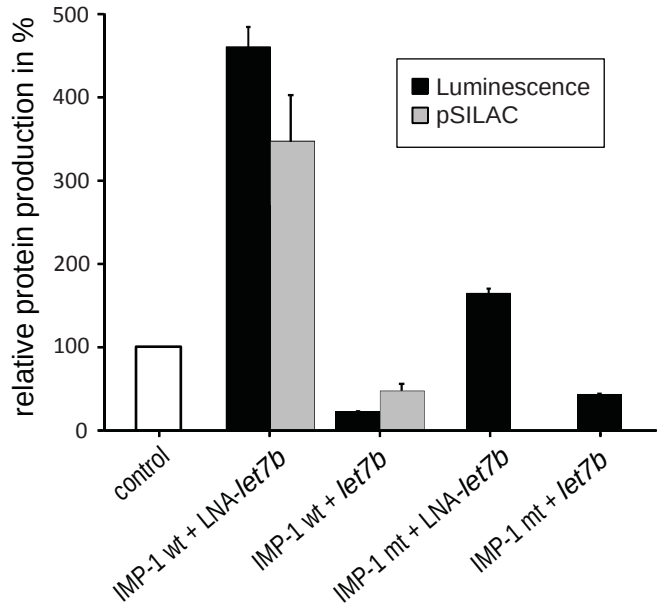
Seed frequencies in coding sequences (CDSs), 5' untranslated regions (UTRs) and 3' UTRs of mRNAs and changes in production of the corresponding proteins upon miR-16 overexpression. Seed frequencies for each mRNA were calculated by dividing the number of seeds by the number of nucleotides. Seed frequencies per unique protein were averaged over all mRNAs mapping to a protein. Proteins were sorted by their fold changes and seed frequencies were calculated for bins of 100 proteins. Seeds in 3' UTRs are strongly correlated with reduced protein production. This correlation is weaker but still detectable for CDSs. Seeds in 5' UTRs have no apparent correlation with overall changes in protein synthesis.

Fig. S5



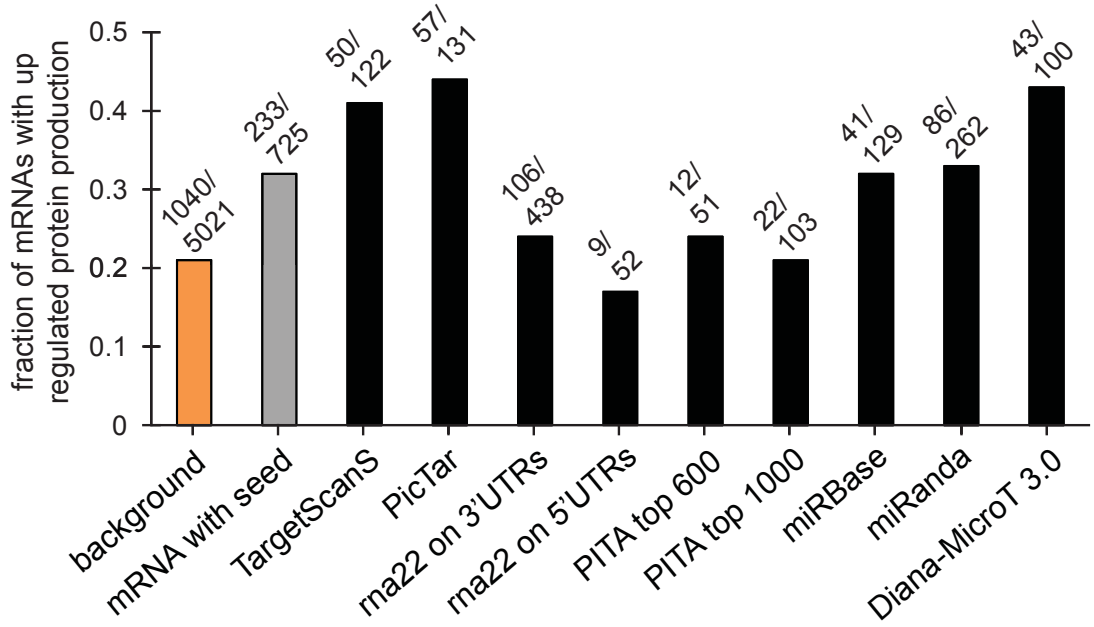
Accuracy of microarray measurements. A total of 69 quantitative RT-PCR (qRT-PCR) measurements were compared to Affymetrix microarray data obtained from the same samples. The correlation was very good in the range where the vast majority of genome-wide miRNA induced mRNA fold changes were observed (log2 fold change between -0.3 and -1.0, dashed lines)

Fig. S6



Dual-luciferase reporter assays for 3' UTR mediated regulation of insulin-like growth factor 2 mRNA-binding protein 1 (IMP-1) by *let7b*. HeLa cells were co-transfected with IMP-1 reporter constructs and either *let7b,* LNA-*let7b* or respective control oligos (LNA-*miR-21* or *miR-155*). Changes in luciferase expression are presented as changes relative to controls. Knocking down endogenous *let7b* enhances luciferase expression while *let7b* over expression represses luciferase activity of the wild-type construct (IMP-1 wt). Mutating the seed (IMP-1 mt) reduces *let7b* mediated regulation. Measurements from the pSILAC experiments are shown for comparison.

Fig. S7



Comparison of let7b knock down data with computational target predictions.
The fraction of predicted let7b target mRNAs with increased protein production (log2 fold change > 0.1) is shown. Since we found that some up-regulated TargetScanS predictions in the LNA let7b experiment were lost due to the mapping, we used an additional mapping procedure for TargetScanS. We first took the union of RefSeq identifiers mapped by biomart (version 0.6) and the cross-reference table from the "known genes" dataset (assembly May 2004) of the UCSC genome browser database (http://genome.ucsc.edu). Then for each RefSeq identifier that was present in PicTar predictions but not in the mapped predictions of TargetScanS, we checked it via the UCSC genome browser and included it in case it was a predicted TargetScanS target. The slightly higher accuracy of PicTar compared to TargetScanS comes from a number of targets with conserved imperfect seeds with compensatory base pairings. TargetScanS does not predict such targets.

Supplementary Table 1: Correlation of target predictions with changes in protein production for five miRNAs (*miR-1, miR-16, miR-30a, miR-155, let7b*)

prediction algorithm	number of predicted targets mapped to Refseq	number of targets measured by pSILAC	number of down-regulated targets (log2FC < -0.1)	fraction of down-regulated targets (log2FC < -0.1)	Reference
TargetScanS	2842	622	381	61%	[1]
PicTar	3289	629	386	61%	[2]
rna22 on 3'UTRS	4112	723	255	35%	[3]
rna22 on 5'UTRS	607	79	20	25%	[3]
PITA top 600	3000	325	139	43%	[4]
PITA top 1000	5000	572	226	40%	[4]
miRbase	3347	658	288	44%	[5]
miRanda	8605	1533	715	47%	[6]
Diana-MicroT 3.0	1678	294	194	66%	Hatzigeorgiou, A., unpublished

1 Lewis, B. P., Burge, C. B., and Bartel, D. P., Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120** (1), 15 (2005).

2 Krek, A. et al., Combinatorial microRNA target predictions. *Nat Genet* **37** (5), 495 (2005).

3 Miranda, K. C. et al., A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* **126** (6), 1203 (2006).

4 Kertesz, M. et al., The role of site accessibility in microRNA target recognition. *Nat Genet* **39** (10), 1278 (2007).

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

6 Betel, D. et al., The microRNA.org resource: targets and expression. *Nucleic Acids Res* **36** (Database issue), D149 (2008).