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).

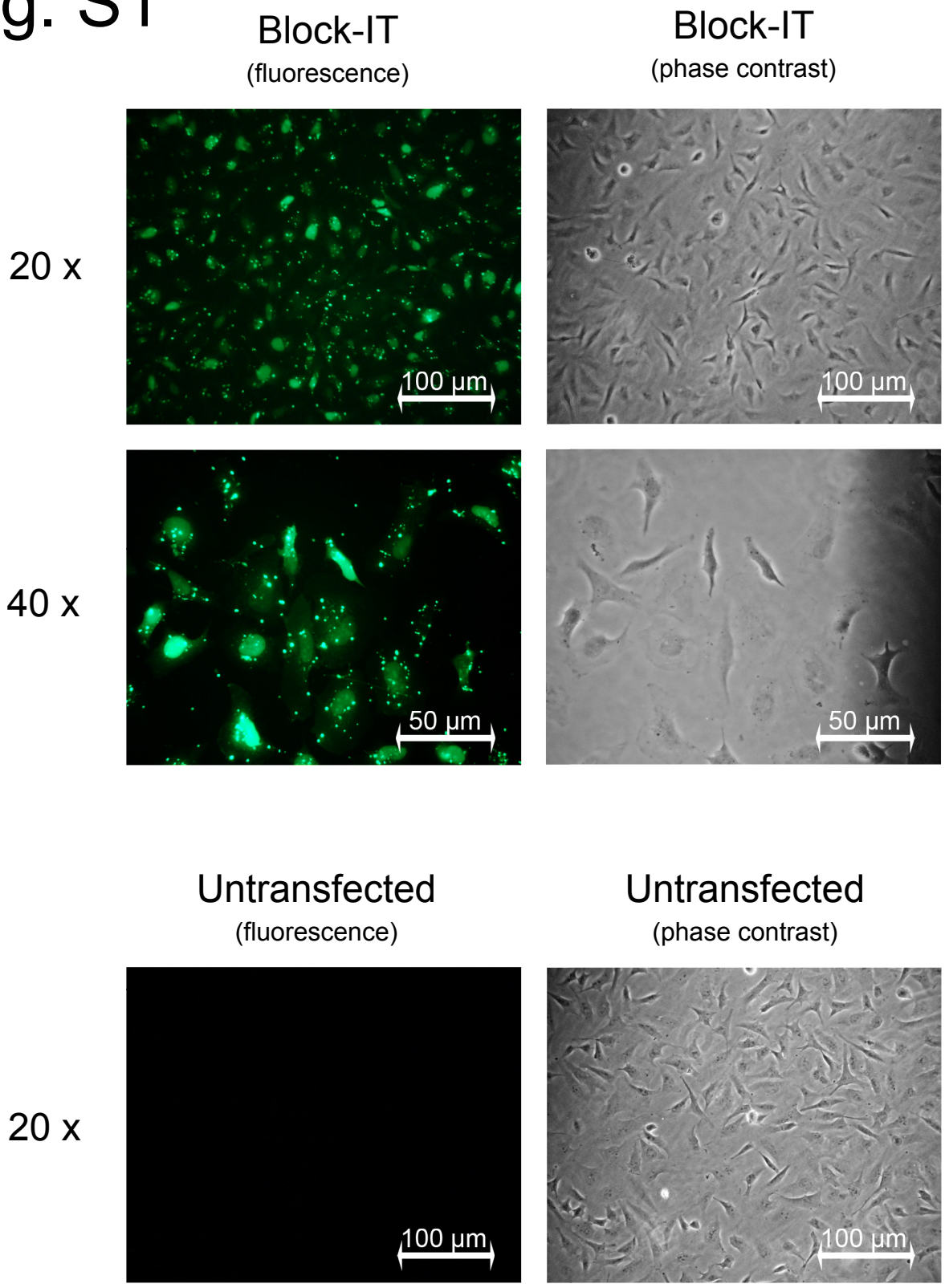
[60] Miranda, K. C. et al., A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* 126, 1203-1217 (2006).

[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).

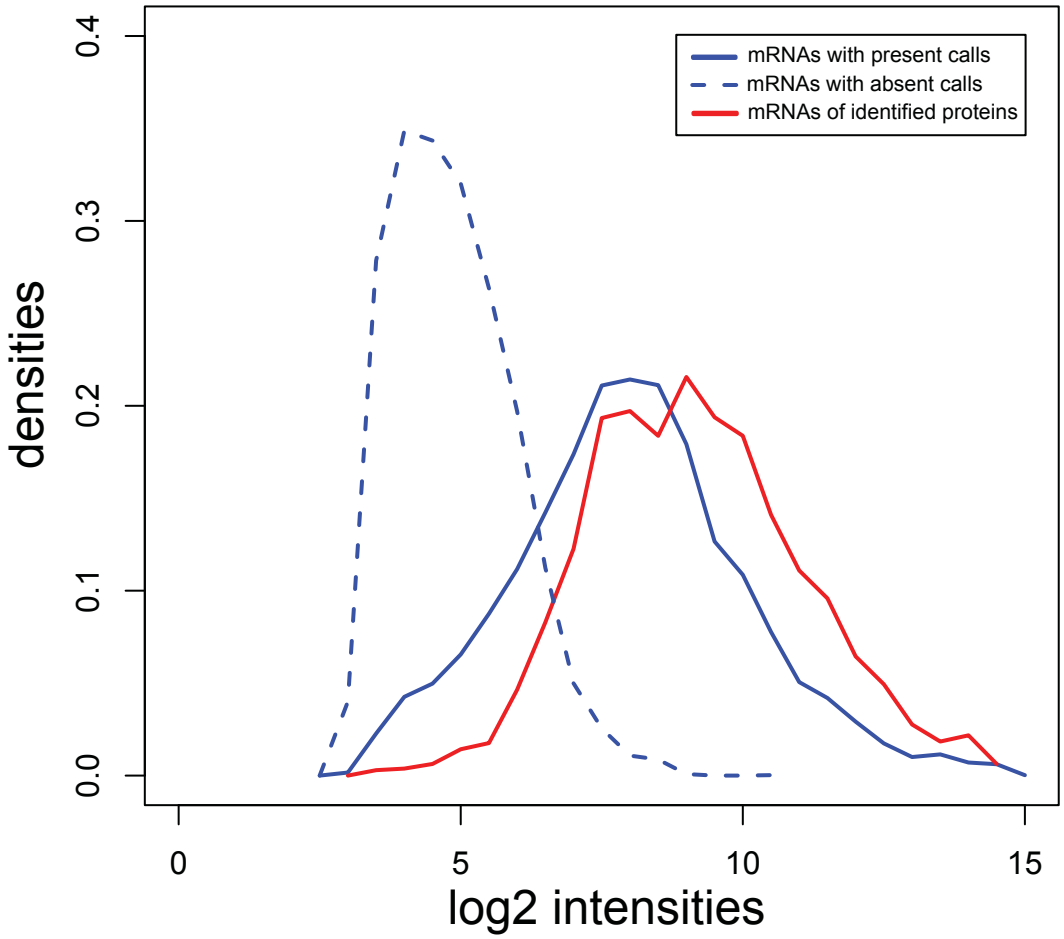
[63] John, B. et al., Human MicroRNA targets. *PLoS Biol* 2, e363 (2004).

Fig. S1



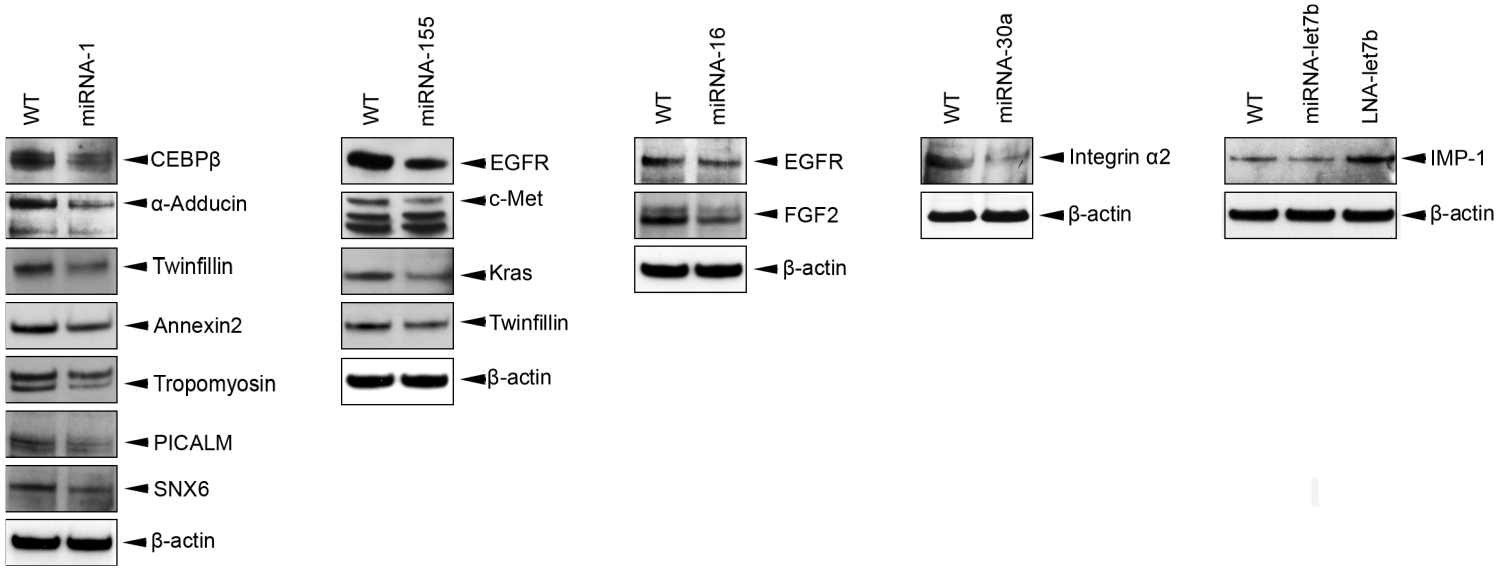
Analysis of transfection efficiency. HeLa cells were transfected with fluorophore-conjugated dsRNA (Block-IT, Invitrogen) according to our miRNA transfection protocol. Phase contrast images are presented as a reference.

Fig. S2



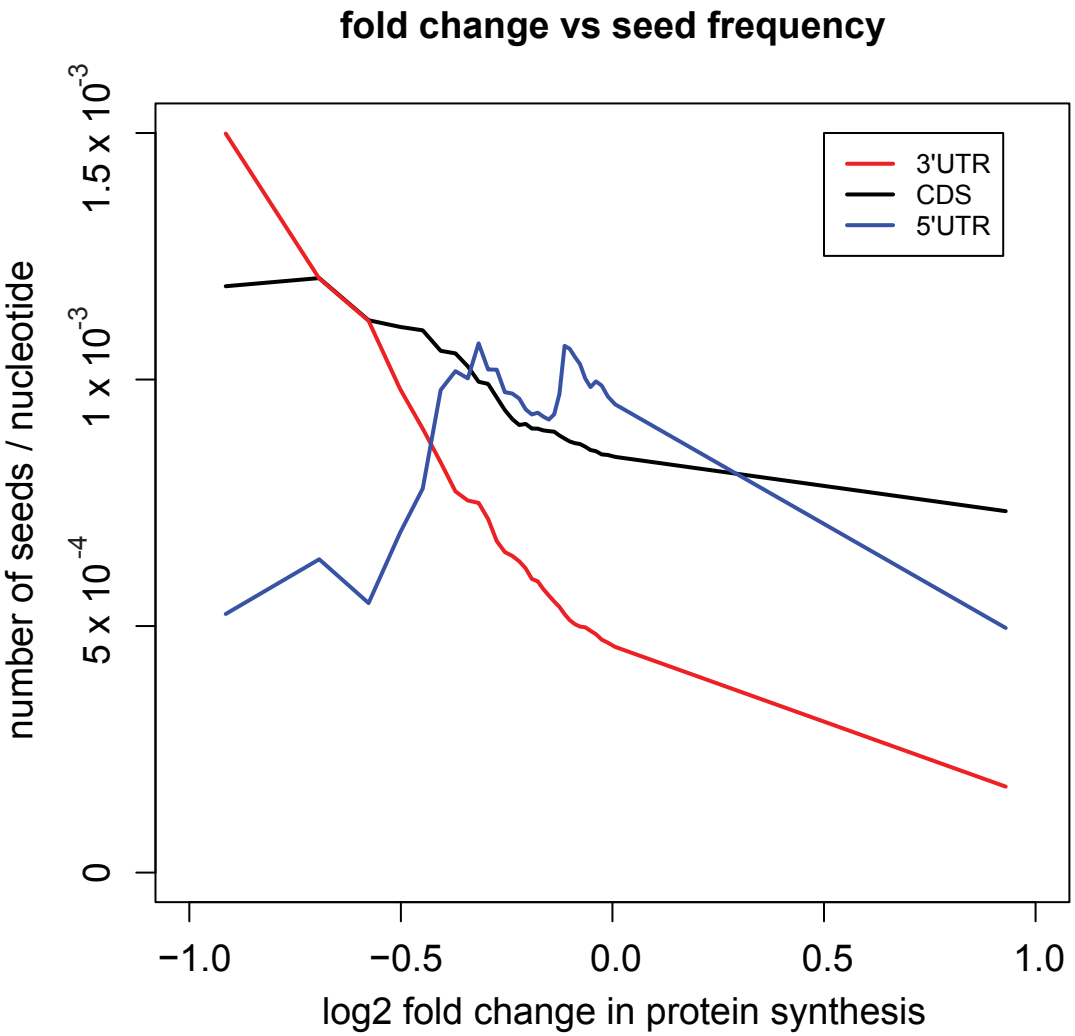
Dynamic range of detection. Signal intensities for all 29,729 mRNAs measured in a microarray experiment (blue) and the subset of 4778 mRNAs for which we could quantify corresponding proteins (red) are shown. Solid and dashed blue lines designate mRNAs with present or absent calls, respectively. Normalizing 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



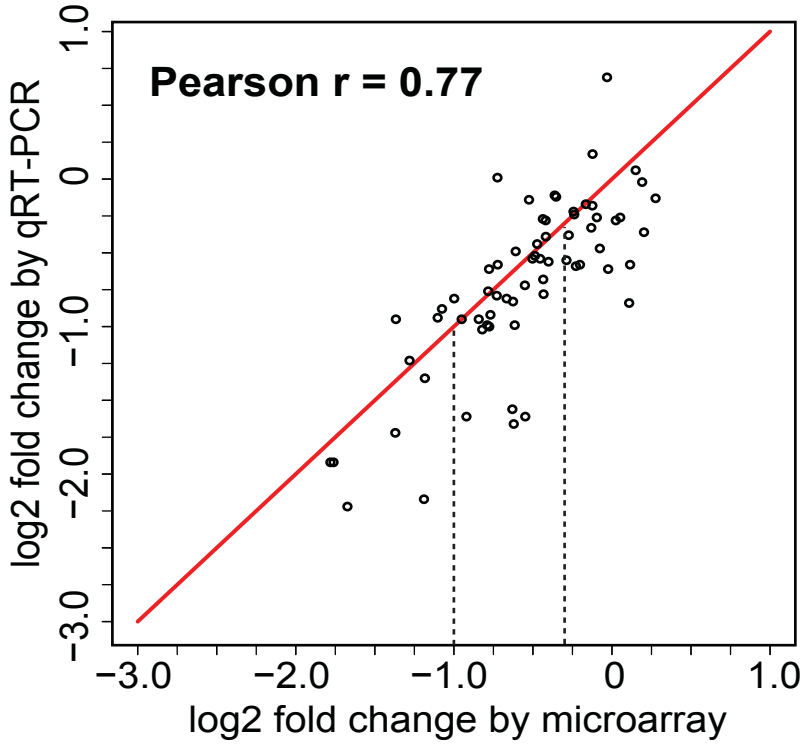
Verification of targets identified by pSILAC with western blots. HeLa cells were transfected with the miRNA, LNA or control treated and analysed by western blotting with antibodies against c-Met, EGFR, α-Adducin, CEBPβ, Twinfilin, Annexin2, Tropomyosin, PICALM, SNX6, Kras, FGF2, Integrin-α2 and IMP-1.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

Fig. S4



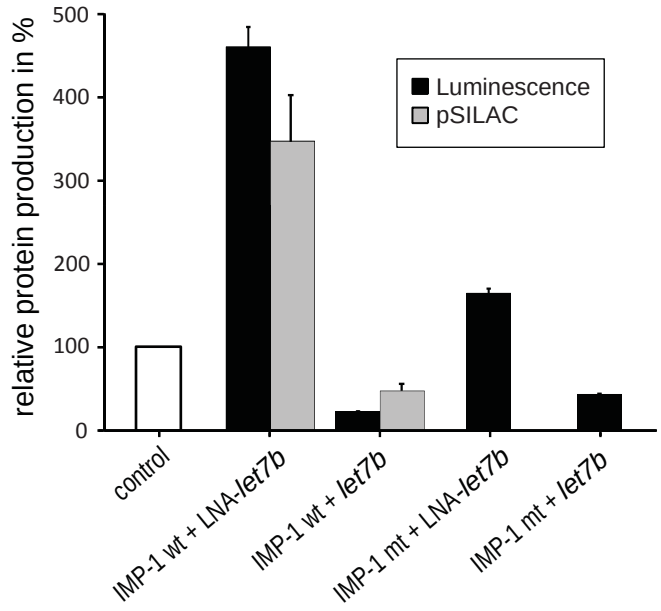
Seed frequencies in coding sequences (CDSs), 5' untranslated regions (UTRs) and 3' UTRs of mRNAs and changes in production of the corresponding proteins upon miR-16 overexpression. Seed frequencies for each mRNA were calculated by dividing the number of seeds by the number of nucleotides. Seed frequencies 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



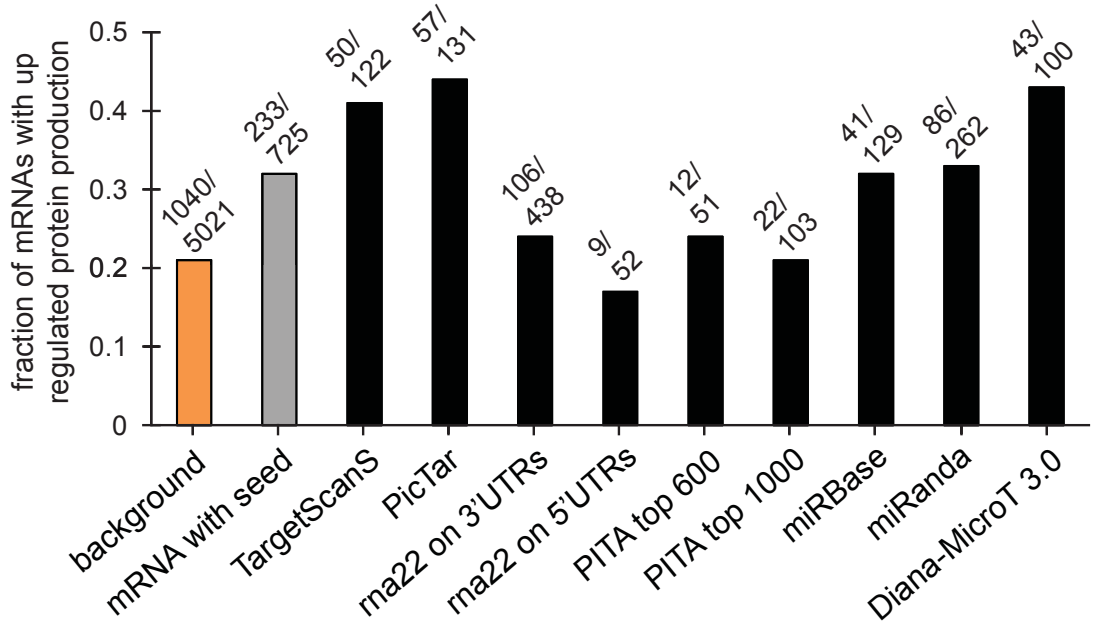
Accuracy of microarray measurements. A total of 69 quantitative RT-PCR (qRT-PCR) measurements were compared to Affymetrix microarray data obtained from the same samples. The correlation was very good in the range where the vast majority of genome-wide miRNA induced mRNA fold changes were observed (log2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