nature chemical biology

ARTICLE

PUBLISHED ONLINE: 22 DECEMBER 2014 | DOI: 10.1038/NCHEMBIO.1713

miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19–miR-106a interaction

Jochen Imig[1,5], Andreas Brunschweiger[1,4,5], Anneke Brümmer[2,4], Boris Guennewig[1,4], Nitish Mittal[2], Shivendra Kishore[2,4], Panagiota Tsikrika[1], André P Gerber[3], Mihaela Zavolan[2]* & Jonathan Hall[1]*

Identifying the interaction partners of noncoding RNAs is essential for elucidating their functions. We have developed an approach, termed microRNA crosslinking and immunoprecipitation (miR-CLIP), using pre-miRNAs modified with psoralen and biotin to capture their targets in cells. Photo-crosslinking and Argonaute 2 immunopurification followed by streptavidin affinity purification of probe-linked RNAs provided selectivity in the capture of targets, which were identified by deep sequencing. miR-CLIP with pre-miR-106a, a miR-17-5p family member, identified hundreds of putative targets in HeLa cells, many carrying conserved sequences complementary to the miRNA seed but also many that were not predicted computationally. miR-106a overexpression experiments confirmed that miR-CLIP captured functional targets, including H19, a long noncoding RNA that is expressed during skeletal muscle cell differentiation. We showed that miR-17-5p family members bind H19 in HeLa cells and myoblasts. During myoblast differentiation, levels of H19, miR-17-5p family members and mRNA targets changed in a manner suggesting that H19 acts as a 'sponge' for these miRNAs.

MiRNAs are short noncoding RNAs (ncRNAs) that regulate gene expression after transcription[1–3]. miRNAs interact with target mRNAs in the miRNA-induced silencing complex (miRISC), which contains a protein of the Argonaute (Ago) family, leading to translational repression, mRNA decay or both[4–6].

The canonical interaction of a miRNA with an mRNA involves base pairing between the 5′ miRNA end ('seed') and a complementary sequence in the target's 3′ untranslated region (3′ UTR). However, noncanonical modes of binding have also been described, for example, those involving the 3′ regions of miRNAs[1,7–9], and there is variability between miRNAs in the prevalence of canonical versus noncanonical binding[10]. Together, these studies suggest that there are yet-undiscovered mechanisms of post-transcriptional regulation through miRNAs. They also highlight a need for new experimental methods to complement existing computational target prediction methods. Established genome-wide methods to identify miRNA-bound RNAs are RNA immunoprecipitation (RIP) and crosslinking precipitation (CLIP), in which the RNAs in RNA-induced silencing complexes (RISCs) are identified after the immunoprecipitation of participating proteins, usually Ago2 (refs. 8,11,12). However, neither of these methods reveals which miRNA interacted with each of the isolated binding sites. In two recent studies, biotinylated miRNA mimics were used to capture mRNA targets[13,14]. However, these methods have the potential to generate false positives from probe-target interactions outside the miRISC[15] and false negatives owing to the loss of target RNAs during purification[16]. Solutions to these potential shortcomings are immunoprecipitation of the mRNA ribonucleoprotein complexes[17] and crosslinking of the miRNA-RNA partners[13,18] before sample work-up. Recently, a crosslinking approach (cross-linking, ligation and sequencing of hybrids, also known as CLASH) was described in which crosslinked miRNA–target site chimeras are generated and sequenced[10,19]. Ligated miRNA-mRNA chimeras were also subsequently identified in published Ago-CLIP data sets[19]. The main drawbacks of this approach are its relatively low efficiency of hybrid capture and inability to provide quantitative information concerning the strength of miRNA-target interactions.

We developed a new approach, termed miR-CLIP, that we used to identify the 'targetome' of miR-106a in HeLa cells. miR-106a is a member of the miR-17-5p seed family, which includes miR-106a/b, miR-20a/b and miR-93 (ref. 20). The central feature of miR-CLIP is the capture probe, which carries internally positioned photoreactive psoralen and biotin groups. A miR-106a probe was transfected into cells, and after brief irradiations, probe-linked RNAs were isolated by sequential Ago2 immunoprecipitation and streptavidin affinity purifications. Deep sequencing of cDNAs generated from these RNAs identified approximately 600 reproducibly enriched miR-106a potential targets, many of which contained highly conserved TargetScan-predicted binding sites. Among these was a maternally imprinted long noncoding RNA, H19 (refs. 21–23), which is highly expressed during embryogenesis in skeletal muscle and in certain cancers[24,25]. We demonstrated that H19 interacts physically with the Ago2 protein as well as with miR-17-5p family members in HeLa cells and differentiating primary myoblasts. Our data indicate that miR-106a stabilizes H19 expression and that in early differentiation of muscle cells, H19 competes for both let-7 and miR-106a with the mRNA targets of these miRNAs.

RESULTS

Design of chemical probes for miR-CLIP

We sought to capture RNA binding to miRNAs in cells using psoralen- and biotin-modified oligoribonucleotides. The crosslinking of psoralen-conjugated oligonucleotides to complementary strands has been demonstrated in different systems[18,26].

We selected the human miR-106a from the miR-17-5p family to establish the required chemistry. We carried out a copper-catalyzed

[1]Institute of Pharmaceutical Sciences, ETH Zurich, Zurich, Switzerland. [2]Computational and Systems Biology, Biozentrum, University of Basel, Basel, Switzerland. [3]Department of Microbial and Cellular Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey, UK. [4]Present addresses: Faculty of Chemistry and Chemical Biology, TU Dortmund University, Dortmund, Germany (A. Brunschweiger); Department of Integrative Biology and Physiology, University of California–Los Angeles, Los Angeles, California, USA (A. Brümmer); Garvan Institute of Medical Research, Darlinghurst, Australia (B.G.); Centogene, Rostock, Germany (S.K.). [5]These authors contributed equally to this work. *e-mail: mihaela.zavolan@unibas.ch or jonathan.hall@pharma.ethz.ch.

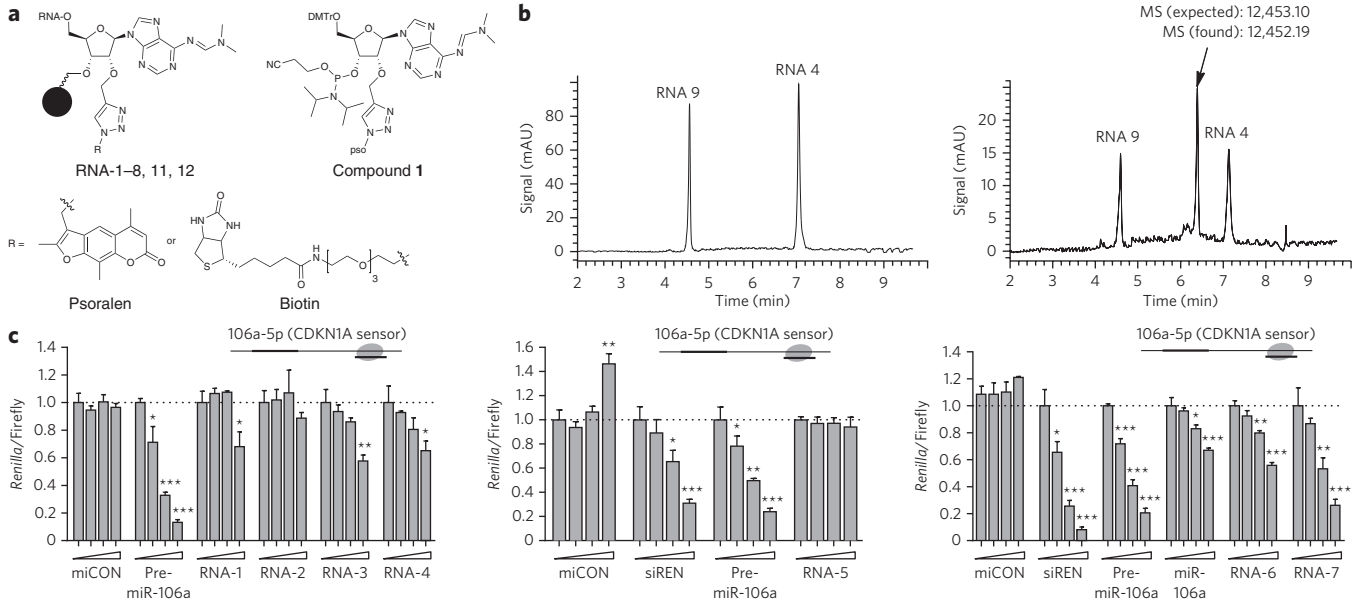

Figure 1 | Synthesis and characterization of functionalized miRNA mimics. (**a**) Left, position-specific post-synthetic labeling of 2′-*O*-propargyladenosine-modified (pre-) miR-106a mimics by reaction with azido-functionalized psoralen (pso) or biotin. Filled circle represents attachment to solid support. Right, phosphoramidite building block **1** for introduction of psoralen into RNAs 6, 7, 11 and 12. (**b**) *In vitro* photo-crosslinking of RNA-4 to a 15-nt counterstrand RNA-9. Left, HPLC chromatogram of the annealed duplex before irradiation; right, HPLC chromatogram after 5 min of irradiation at 365 nm. (**c**) The silencing activity of miR-106a mimics RNAs 1–7, together with controls miCON and siRen (a siRNA that specifically targets the *Renilla* reporter gene and is used as an indicator of transfection efficiency) after cotransfection with luciferase CDKN1A sensors reporters into HeLa cells at concentrations ranging from 0–36 nM. Error bars indicate ± 1 s.d. (n = 3).

alkyne-azide cycloaddition (CuAAC) of 4′-azidomethyltrioxsalen with 2′-*O*-propargyl adenosines (**Supplementary Results, Supplementary Note**) at positions 1–4 of miR-106a or pre-miR-106a on solid support (RNAs 1–8; **Fig. 1a**, **Table 1** and **Supplementary Table 1**). A preloaded solid support was used to introduce biotin at the miRNA 3′ end (RNA-5), whereas to introduce biotin and psoralen at internal positions, we used CuAAC chemistry using a modified adenosine phosphoramidite (**1**; **Fig. 1a** and **Supplementary Note**). Hence, miR-106a was functionalized with psoralen and biotin groups at positions 4 and 18, respectively, yielding RNA-6 and RNA-7 (**Table 1**).

To confirm that the psoralen could reach a hybridized target RNA, we developed an *in vitro* crosslinking assay. RNA-4 was annealed to complementary RNA-9 (**Table 1**) and irradiated at 365 nm for 5 min. A new product was formed in approximately 50% yield with a mass corresponding to that of a crosslinked duplex (**Fig. 1b** and **Supplementary Fig. 1**). Similar results were obtained with RNAs 1–3. However, analysis of LC/MS chromatograms suggested that intrastrand crosslinking in probe strands also occurred with some miRNAs. The results suggested that the psoralen group was appropriately positioned to crosslink the probe to cellular targets.

To demonstrate that the psoralen-modified miRNAs were functional, we hybridized RNAs 1–4 with a passenger strand (RNA-10; **Table 1**) and tested them for their suppression of reporter mRNAs bearing target sequences (**Fig. 1c** and **Supplementary Fig. 2**). RNA-1 and RNA-2 showed poor activity, whereas RNA-3 and RNA-4 were active, indicating that the silencing mechanism tolerated the psoralen at positions 3 and 4 (**Fig. 1c**). RNA-5, the 3′-biotinylated analog of RNA-4, was inactive, consistent with a report that biotin modification at the miRNA 3′ terminus may prevent uptake into miRISC[15]. However, RNA-6 and the pre-miRNA analog RNA-7 bearing biotin at internal positions were only slightly less active than unmodified miR-106a and pre-miR-106a, respectively. To explore the generality of the probe design, we tested duplex miR-CLIP probes for two other miRNAs, let-7g (RNA-11), a member of the let-7 family, and miR-34a (RNA-12). The probes contained psoralen at position 7 and 5, respectively, and biotin at position 17 and 13, respectively (**Table 1**). RNA-11 and RNA-12 were annealed to their respective passenger strands, RNA-13 and RNA-14 (**Table 1**), and tested for inhibition of luciferase reporters that we have used previously[27]. Both showed significant activity (**Supplementary Fig. 2**), suggesting that a functional psoralen or biotin substitution pattern should be available for most miRNAs.

Establishment of the miR-CLIP protocol

To capture targets of miR-106a-5p in cells under mild conditions, i.e., without eliciting a strong biological response to miR-106a

Table 1 | Oligonucleotide sequences.

Entry	Sequences (5′ to 3′)
RNA-1	**A**^{pso}AAAGUGCUUACAGUGCAGGUAG
RNA-2	A**A**^{pso}AAGUGCUUACAGUGCAGGUAG
RNA-3	AA**A**^{pso}AGUGCUUACAGUGCAGGUAG
RNA-4	AAA**A**^{pso}GUGCUUACAGUGCAGGUAG
RNA-5	AAA**A**^{pso}GUGCUUACAGUGCAGGUAG^{bio}
RNA-6	AAA**A**^{pso}GUGCUUACAGUGC**A**^{bio}GGUAG
RNA-7	AAA**A**^{pso}GUGCUUACAGUGC**A**^{bio}GGUAGCUUUUUGAGA UCUACUGCAAUGUAAGCACUUCUUAC
RNA-8	AAA**A**^{pso}GUGCUUACAGUGCAGGUAGCUUUUUGAGAU CUACUGCAAUGUAAGCACUUCUUAC
RNA-9	ACUGUAAGCACUUUU
RNA-10	CUGCAAUGUAAGCACUUCUUAC
RNA-11	UGAGGU**A**^{pso}GUAGUUUGU**A**^{bio}CAGUU
RNA-12	UGGC**A**^{pso}GUGUCUU**A**^{bio}GCUGGUUGU
RNA-13	CUGUACAGGCCACUGCCUUGC
RNA-14	CAAUCAGCAAGUAUACUGCCCU

Structures of psoralen (pso) and biotin (bio) are in **Figure 1**.

overexpression, we used 3-nM concentrations of RNA-7, which were mostly ineffective in the reporter assays (**Fig. 1c** and **Supplementary Fig. 2**). Biotinylated and 365-nm UV-light cross-linked adducts were captured directly on streptavidin beads. Samples from RNA-7 and control RNA-8 (lacking biotin) treatments were analyzed by quantitative real-time PCR (qRT-PCR) for enrichment of the previously validated targets CDKN1A[28] and RB1 (ref. 29). Because only low enrichment for CDKN1A (P = 0.035) and no enrichment for RB1 were recorded (**Supplementary Fig. 3**), we introduced a more stringent, two-step irradiation and tandem purification process (miR-CLIP; **Fig. 2a**). It consisted of a first irradiation at 254 nm aiming to crosslink RNAs to proteins in miRISC, followed by an irradiation at 365 nm to crosslink psoralen-bearing miRNAs to their targets. Irradiations were followed by Ago2 immunoprecipitation (**Supplementary Fig. 4**)[30] and then streptavidin-based affinity isolation of complexes. We observed significant enrichments of CDKN1A (ninefold, P = 0.0017) and the predicted targets ARID4B and ANKFY1 (six- and fourfold, respectively, (P = 0.0157 and P = 0.024, respectively) for RNA-7 treatment (**Fig. 2b**).

The targetome of miR-106a-5p

To identify miR-106a targets, we sequenced miR-CLIP–captured RNAs from HeLa cells. Four cDNA libraries (**Supplementary Fig. 5**), each in two replicates, were submitted for deep sequencing: an RNA sample from compound RNA-7–transfected cells ('Input'); a sample of the Ago2-immunopurified RNA from cells treated with transfection reagent ('Mock'), a sample of the Ago2-immunopurified RNA from compound RNA-7 treatment ('Ago2-IP'), and the sample from miR-CLIP-purified RNA ('miR-CLIP'; **Fig. 2a**). Deep sequencing yielded up to 2 × 10^7 good-quality reads per library, which were uploaded to the CLIPZ server (http://www.clipz.unibas.ch/)[31] for mapping and annotation (**Supplementary Table 2**). Of the mapped reads, 30–50% mapped to mRNAs, a result similar to that commonly observed in Ago2-CLIP experiments[32]. The abundance of individual transcripts was highly reproducible in miR-CLIP replicates (Spearman coefficients of >0.91; **Supplementary Fig. 6**). A global comparison of data from the Ago2 immunoprecipitation from RNA-7–treated and mock-treated cells revealed that there were no large-scale changes (correlation coefficient: 0.98) in Ago2 RNA recruitment upon miR-106a overexpression; i.e., the data were not skewed by treatment with low concentrations of RNA-7 (**Supplementary Fig. 7**). A gene ontology term analysis of the top 600 reproducibly enriched miR-CLIP targets revealed a strong representation of genes involved in transcription regulation, cytoskeleton remodeling and cell adhesion (**Supplementary Fig. 8**; $P < 2.0 \times 10^{-5}$), consistent with the oncogenic properties of this miRNA family.

We then examined the representation of miR-106a targets that were predicted on the basis of evolutionary conservation with TargetScan (TargetScan PCT, where PCT stands for probability of conserved targeting). For miR-106a-5p, TargetScan lists more than 2,600 highly conserved mRNA targets. Compared to nontargeted transcripts with comparable abundance in the Input RNA, predicted miR-106a targets showed significantly higher abundance in Ago2-IP and especially in miR-CLIP (**Fig. 3a**), much more than that measured for predicted targets of any other miRNAs (**Fig. 3a**). Also as expected from Ago2 immunoprecipitation, the proportion of miRNA (miR-106 and let-7) targets was highest among transcripts with the highest enrichment in Ago2-IP compared to Input (**Supplementary Fig. 9**). This suggests that the immunoprecipitation protocol was effective. Importantly, miR-106a targets were significantly enriched in miR-CLIP compared to nontargets over the entire range of their abundance in the Ago2 immunoprecipitation (**Fig. 3a**), whereas the biotin-based pulldown of miR-106a did not enrich predicted targets of other miRNAs (**Fig. 3a**). Altogether,

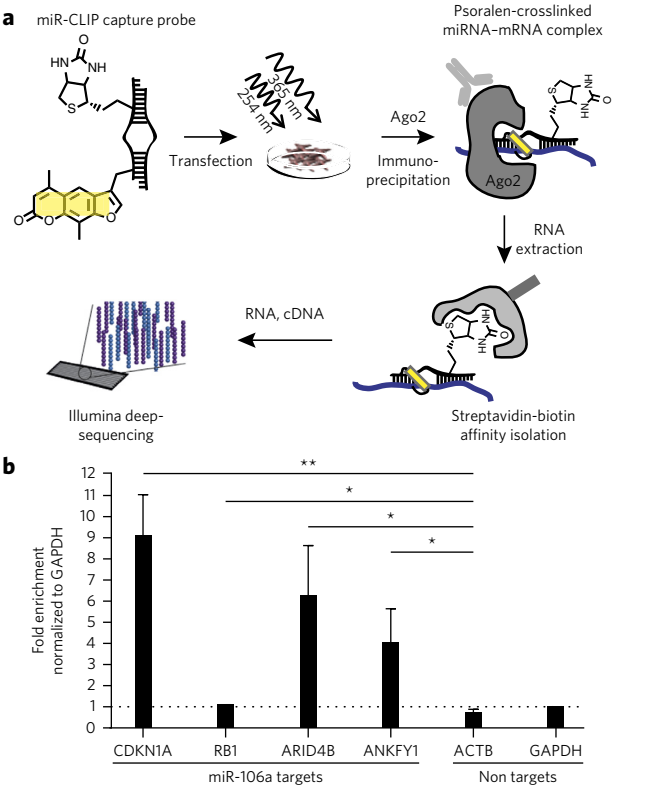

Figure 2 | Development of the miR-CLIP method. (**a**) Sketch of the miR-CLIP method. miRNA probes bearing psoralen and biotin or controls were transfected into HeLa cells. Ago2 immunoprecipitation was performed 24 h after transfection, followed by streptavidin affinity purification of the probe-linked RNAs. The isolated RNA was converted into cDNA and subjected to deep sequencing. (**b**) Enrichment of targets through the miR-CLIP method using RNA-7 versus control RNA-8: mRNA levels were determined by qRT-PCR (n = 3). *P < 0.05 or **P < 0.01, two-tailed Student's t-test. Error bars indicate ± 1 s.d.

these results indicated that miR-CLIP tandem purification led to the isolation of the miR-106a targetome.

To explore the general applicability of miR-CLIP, we repeated the entire protocol using miR-CLIP probe RNA-11, which corresponds to the let-7g miRNA. The 1800 TargetScan Pct-predicted let-7g targets were more significantly enriched in Ago2-IP and miR-CLIP than the predicted targets of most other miRNAs with comparable abundance in Input (**Supplementary Fig. 10**). Also, a significant enrichment of let-7g targets was found in the miR-CLIP compared to the Ago2-IP sample over a wide range of transcript abundances.

Consistent with the expectation that miR-106a recognizes its targets predominantly through its seed region[1,10], we found that the reverse complements of 6-mers, 7-mers and 8-mers derived from the miR-106a seed sequence were more abundant compared to similar-length motifs derived from the miRNA's 3′ end. Furthermore, these motifs were strongly enriched in the 3′ UTRs of purified transcripts in the miR-CLIP library compared to shuffled 3′ UTR sequences (**Supplementary Fig. 11**). Notably, the reverse complement of 6-mer motifs derived from the miRNA's 3′ end were almost as strongly enriched as those 6-mers derived from the miRNA's 5′ end, possibly indicating interaction between the 3′ end of the miRNA with some targets.

To investigate the functionality of interactions, we analyzed changes in expression of the miR-CLIP-106a targets after transfection with RNA-7 under the same mild conditions used for the miR-CLIP protocol (**Fig. 3b** and **Supplementary Fig. 12**). Targetscan PCT–predicted miR-106a targets with 8-mer and 7-mer

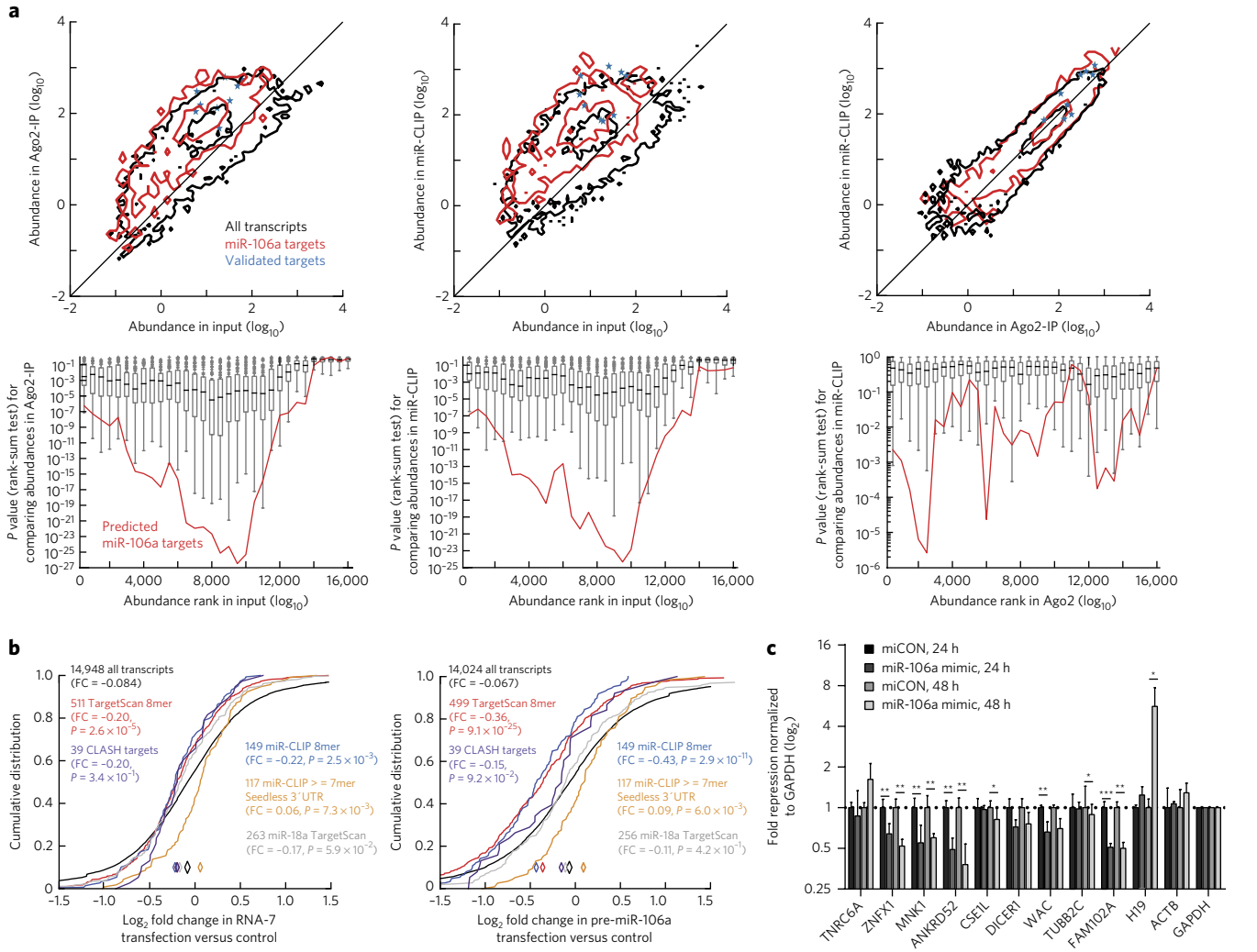

Figure 3 | miR-CLIP miR-106a targetome identification. (**a**) Upper panels, contour plots of transcript abundance levels ($n = 2$). Blue asterisks, validated miR-106 targets; red lines, TargetScan-predicted miR-106a targets; black lines: all transcripts. Contours show 95% and 40% of the data of each transcript set. Lower panels, *P* values from Wilcoxon's rank-sum test comparing abundances of TargetScan-predicted miR-106a targets (red) to nontargets in a sliding window (sliding by 500 transcripts) of 1,000 transcripts sorted by abundance from high (left) to low (right); distributions of *P* values for predicted transcripts of all other miRNAs shown as box plots. IP, immunoprecipitation. (**b**) Cumulative distributions of $\log_2$ fold changes (FC) in transcript abundances of RNA-7 (left) and pre-miR-106a (right) transfections. Diamonds mark the median fold change of each subset denoted the figure legends. The distribution of fold changes in all transcripts is compared to five subsets of transcripts: predicted miR-106a TargetScan targets with 8mers seed match and the top 1,500 miR-CLIP targets with 8mer seed match and without a seed match, but with a match of at least 7 nucleotides starting at or beyond position 8 of the miRNA sequence (located in the 3′ UTR). Comparison to unrelated miR-18a predicted and top miR-106a CLASH targets are shown as a control. *P* values are calculated with Wilcoxon's rank sum test and indicated. (**c**) qRT-PCR validation of the top ten abundance ranked miR-CLIP miR-106a targets after 24 h and 48 h. $^*P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$, two-tailed Student's *t*-test ($n = 4$). Error bars indicate ± 1 s.d.

seed matches (7mer-m8 and 7mer-1a) were significantly repressed after transfections with RNA-7 (**Fig. 3b**). miR-CLIP-106a transcripts with 8-mer and 7-mer seed matches behaved in a similarly significant manner (**Fig. 3b**). Downregulation of both TargetScan PCT–predicted and miR-CLIP-106a–identified targets was also observed after transfection of an unmodified chemically synthesized pre-miR-106a, indicating that RNA-7 emulates its unmodified counterpart well (**Fig. 3b** and **Supplementary Fig. 12**). Although the fold changes in target expression were typical for this type of experiment[10], RNA-7 induced smaller changes than pre-miR-106a, possibly reflecting different reagent potencies. Notably, among the 1,500 transcripts that were most abundant in miR-CLIP, we found a set of transcripts with no seed match to miR-106a but instead a matched stretch of ≥7 nt starting at or after position 8 of the miRNA. These transcripts were significantly elevated after both transfections compared to all transcripts, particularly when the matches were located in the target 3′ UTRs (**Fig. 3b**). These results indicate that miR-CLIP captures bona fide functional targets, including transcripts that may be regulated in a nonconventional manner.

We also compared the behavior of miR-CLIP-106a targets with that of the targets identified for miR-106a[10] by the CLASH method as well as with that of the 21 ligated miR-106a–mRNA chimeras observed in Ago-CLIP experiments (eight of which were captured in our experiment). The miR-106a targets determined with CLASH (**Fig. 3b**) and Ago-CLIP (data not shown) were also downregulated, on average, upon miR-106a overexpression, though these data sets were too small to identify statistically significant changes.

Next, we determined the overlaps of the top 600 miR-CLIP-106a transcripts in terms of their abundance and their enrichment (and its significance) in miR-CLIP over Ago2-IP, with the top 600 miR-106a targets predicted by different computational methods (**Supplementary Table 3**). We consistently obtained the

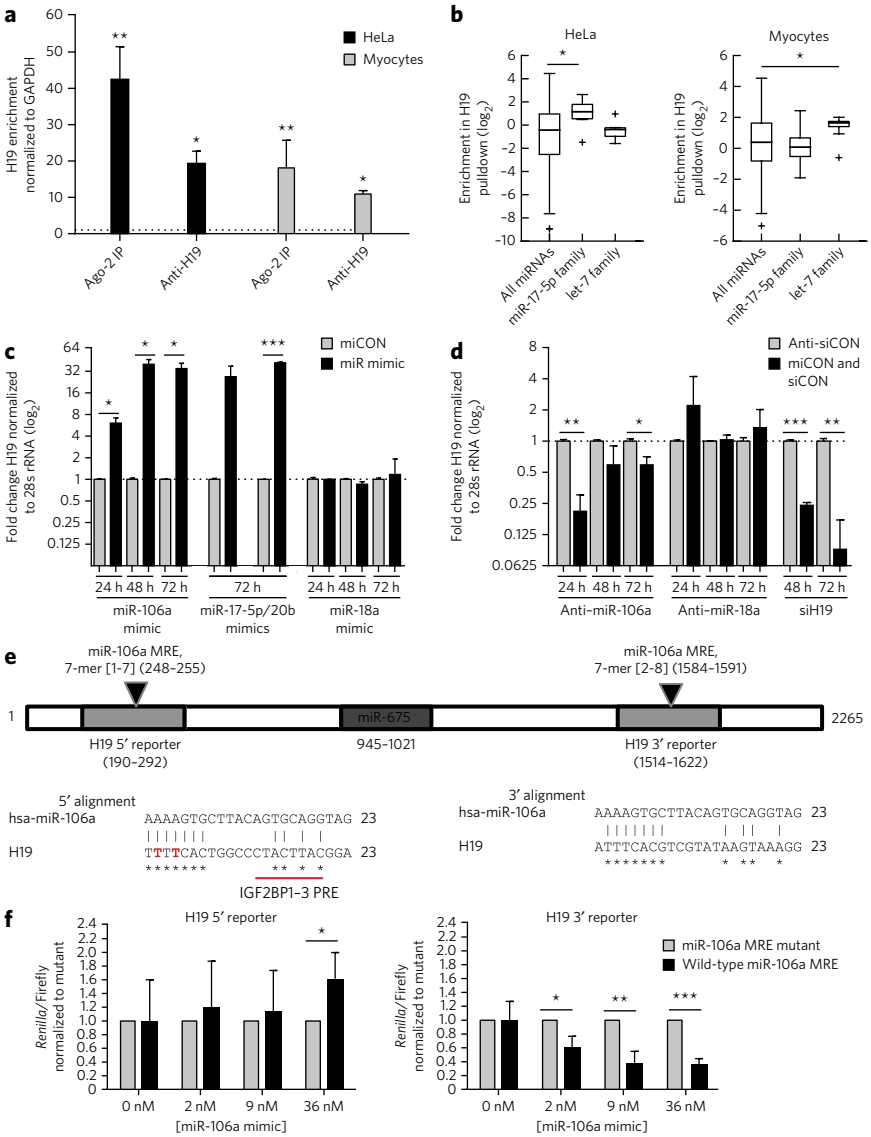

Figure 4 | Endogenous H19 associates with Ago2 and miR-17-5p seed family miRNAs in HeLa and myocytes. (**a**) H19 transcript coprecipitated with Ago2 and biotinylated α-H19 antisense probes ($n = 3$); enrichments were measured by qRT-PCR analysis. (**b**) Distribution of miRNA abundance of miR-17-5p and let-7 miRNA families coprecipitated with the biotinylated anti-H19 probes in H19 pulldown versus control (box plots) in HeLa cells and myocytes, assessed by deep sequencing. *$P < 0.05$. (**c**,**d**) Primary myoblasts were forced to differentiate after reverse transfection with 200 nM of miRNA mimics and anti-miRNAs, controls (miCON or siCON) or an H19 siRNA (siH19). Changes in H19 transcript levels were determined by qRT-PCR analysis ($n = 2$). *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$, two-tailed unpaired Student's t-test. (**e**) Structure of H19 transcript (BC04007) and miR-106a MRE containing luciferase reporters. The sequence alignment of H19–miR-106a below shows miR-106a to their two potential MREs on the H19 transcript. Highlighted in red are nucleotides with diagnostic IGF2BP2 PAR-CLIP T-to-C conversions, and underlined in red are two potential IGF2BP canonical motifs (CATT in IGF2BP3 and CATC/A in IGF2BP1–3). Numbers in square brackets indicate nucleotides in the seed region. (**f**) H19 target regulation and MRE functionality by miR-106a. The luciferase assay was performed on 5′- (left) and 3′- (right) H19 miR-106a MRE-containing reporter constructs with 0-nM, 2-nM, 9-nM and 36-nM concentrations of miR-106a mimic in HeLa cells (5′ construct, $n = 4$; 3′ construct $n = 3$). Relative *Renilla*/firefly ratios were normalized to values for miR-106a MRE seed mutated reporter constructs. *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$, two-tailed unpaired Student's t-test. Error bars indicate ±1 s.d.

largest overlap when ranking transcripts by their abundance (number of reads per nucleotide) in miR-CLIP. We used this measure to select ten putative targets for further validation: their abundances in miR-CLIP-106a were at least 2.5-fold higher ($P < 0.05$) than those in the Ago2-IP (RNA-7–treated) library (**Supplementary Data Set 1**). Six genes (*TNRC6A*, *ZNFX1*, *MINK1*, *ANKRD52*, *WAC* and *FAM102A*) were predicted to be targets by TargetScan, whereas *DICER1* has previously been validated[33]. We transfected HeLa cells with a miR-106a mimic (30 nM) and measured mRNA levels by qRT-PCR. *ZNFX1*, *MINK1*, *ANKRD52* and *FAM10ZA* were strongly repressed, whereas *CSE1L*, *WAC* and *TUBB2C* were repressed at one time point only (**Fig. 3c**). Target repression was greatest for *ANKRD52*, which has four conserved, predicted target sites in its 3′ UTR. Altogether, these data indicate that miR-CLIP-106a captured functional targets of miR-106a.

miR-106a-5p binds and regulates the lincRNA H19

To our surprise, we found that RNA-7 captured the lincRNA H19 (**Supplementary Fig. 13** and **Supplementary Data Set 1**). Furthermore, miR-106a overexpression caused an approximately sixfold increase in H19 after 48 h (**Fig. 3c** and **Supplementary Fig. 14**). Although the physiological role of H19 is still unknown, H19 expression increases markedly during muscle cell differentiation[24], which prompted us to investigate the miR-106a–H19 interaction in primary human myoblasts. Some lincRNAs have been linked to miRNA-containing regulatory networks[34–37]. Indeed, H19 is reported to associate with Ago2 (ref. 38) and is predicted by miRcode[39] (**Supplementary Table 4**) to be targeted by 19 miRNA families (including miR-17-5p and miR-18) on 27 miRNA recognition elements (MREs). The H19 sequence harbors two predicted 7-nt-long MREs for miR-106a that are highly conserved across primates and apes. To verify a direct interaction between RISC and H19, we performed Ago2 immunoprecipitation experiments. The H19 transcript was significantly enriched in HeLa cells and differentiated myocytes (~40-fold and ~20-fold, respectively) in Ago2 immunoprecipitation compared to those treated with a control antibody, consistent with previous findings[40] (**Fig. 4a**). In a complementary approach, we designed three 3′-biotinylated 2′-*O*-methyl-RNAs (RNAs 15–17; **Supplementary Table 1**) complementary to different regions of H19 and used them as a mixture to pull down endogenous H19 RNA from cell lysates. Compared to control beads bearing no oligonucleotides, 20-fold and 10-fold enrichments of H19 were obtained from HeLa and myocytes, respectively (**Fig. 4a**). We then performed small RNA sequencing from distinct fractions. In HeLa cells, the miR-17-5p family was significantly enriched in the H19 pull-down ($P = 0.032$), whereas the let-7 family was not ($P = 0.86$) (**Fig. 4b**). In contrast, let-7 miRNAs were found significantly bound to H19 in myocytes consistent with a previous report[40], whereas the miR-17-5p family was not ($P = 0.043$ and 0.79, respectively). Together, the data provide evidence for a cell type–specific physical interaction of H19 with Ago2 as well as with selected miR-17-5p and let-7 family members.

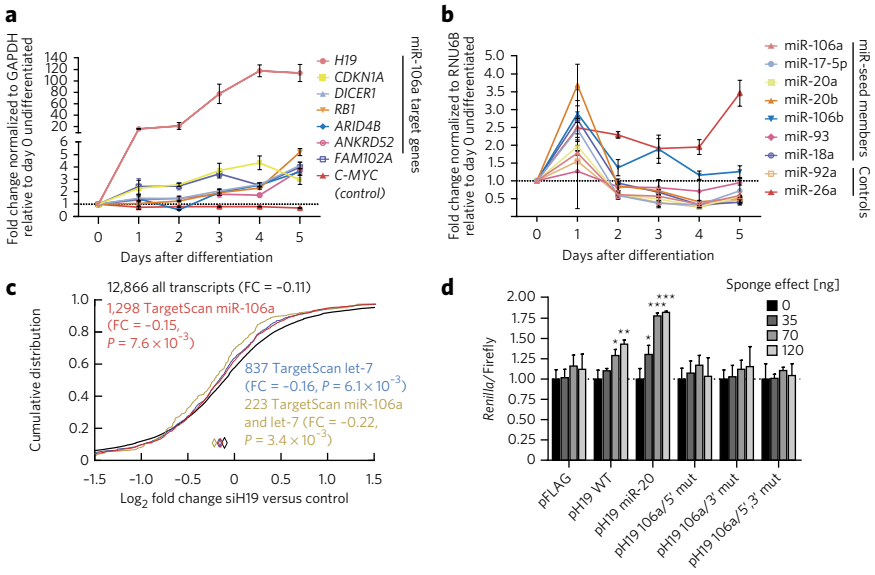

Figure 5 | miR-106a family member and target gene expression profiling in muscle cells with H19 sponge effects. (**a**,**b**) Levels of H19, miR-106a target genes and miR-106a family members from days 1–5 in differentiated myocytes. The expression levels were measured by qRT-PCR or TaqMan qRT-PCR. *C-MYC*, miR-18a and miR-92a served as negative controls, and myo-miR miR-26a served as a positive control ($n = 2$). (**c**) SiH19 treatment of myocytes leads to repression of miR-106a targets. Cumulative distribution of changes in transcript abundances 1 d after siH19 transfection and differentiation initiation in myocytes. Distributions of fold changes of predicted miR-106a, let-7 or combined miR-106a and let-7 target transcripts are compared with all transcripts using Wilcoxon's rank-sum test; median fold changes (diamonds) and calculated *P* values are indicated in the legend ($n = 2$). (**d**) Sponge effect of H19 by luciferase assays in HeLa cells. Cotransfection of different pH19 expression constructs with a fully complementary CDKN1A sensor reporter (80 ng) at indicated concentrations ($n = 3$). Error bars indicate ± 1 s.d.

To assess the effect of miR-106a overexpression on H19, we transfected HeLa cells and myoblasts before initiation of differentiation with miRNA mimics. In HeLa cells, H19 levels increased from 4-fold to 13-fold 48–72 h after transfection with miR-106a (**Supplementary Fig. 14**); miR-17-5p and miR-20b produced similar effects, whereas miR-18a was ineffective. Conversely, treatment of cells with anti-miR-106a decreased H19 by about 60%. Possibly, this anti-miR cross-hybridizes with other miR-17-5p family members because of their sequence similarity (**Supplementary Fig. 14**). Transfections of miR-17-5p family members induced H19 in primary myocytes as well (**Fig. 4c**), whereas miR-18a was ineffective, and let-7g repressed H19 by about 90% (**Supplementary Fig. 14**), as reported[40]. Blocking endogenous miR-106a in differentiating myoblasts also led to significant repression of H19 on days 1 and 3 ($P = 0.0077$ and $P = 0.0455$, respectively; **Fig. 4d**). In summary, in two different cell types, H19 is induced or stabilized by exogenous delivery of miR-17-5p family members, whereas blocking endogenous miR-106a yields the opposite effect.

To investigate whether the two putative miR-106a MREs (**Fig. 4e**) have a role in H19 regulation, we cloned the two sites, including approximately 100 nt of flanking sequence separately into luciferase psiCHECK vectors (5′-MRE and 3′-MRE) in an effort to separate these putative regulatory sites from each other and from other potential *trans*-acting factors, such as RNA-binding proteins (RBPs). Reporter constructs containing four mutated nucleotides in the miRNA seed–targeting regions served as negative controls (**Supplementary Fig. 15**). Reporter assays in myoblasts were thwarted by a poor transfection efficiency of the DNA vectors (in contrast to small RNAs; **Supplementary Fig. 16**). Instead, HeLa cells were transfected with reporter plasmids and miR-106a. For the 5′-MRE reporter, induction of ~60% compared to the control vector was observed, but only at the highest treatment concentration, possibly suggesting a weak miRNA-mediated stabilization effect. In contrast, strong concentration-dependent inhibition of 3′-MRE by miR-106a was seen (**Fig. 4f**). Thus, both sites may represent points of functional interaction of miR-106a–associated RISC, albeit in different manners. However, the data did not explain the induction of H19 by miR-17-5p family members.

Examples of target stabilization by miRNAs have been reported previously[41,42]. In these cases, it was suggested that nearby *cis*-acting elements act as docking sites for RBPs that could compete or cooperate with the miRNAs for binding. In this context, Ago2 is known to interact with insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs), which are reported to stabilize their RNA targets, including H19 (refs. 11,43,44). Database mining in publicly available Photoactivatable-ribonucleoside-enhanced CLIP (PAR-CLIP) data sets revealed two potential IGF2BP recognition elements (CATT and CATG) together with a diagnostic T/C mutation indicating a crosslinking event five nucleotides upstream of miR-106a 5′ MRE (**Fig. 4e**). This suggests a possible interplay of IGF2BP with miR-106a during H19 regulation. We therefore examined a putative stabilizing effect of IGF2BPs 1–3 on H19 levels in primary myocytes using a combination of siRNAs targeting those IGF2BP members. SiIGF2BP1 and siIGF2BP2 reduced their targets by 50–75%, whereas siIGF2BP3 was inactive (**Supplementary Fig. 17**). Inhibition of IGF2BP1 and IGF2BP2 alone markedly reduced H19 levels, suggesting that these RBPs are required to stabilize H19 RNA in myocytes.

H19 RNA inhibits miR-17-5p family

NcRNAs such as pseudogenes and long noncoding RNAs may sequester miRNAs from their target-coding transcripts, thereby functioning as competing endogenous RNAs[45,46]. To investigate the miR-106a–H19 interaction in this context, we monitored five miR-106a targets (*CDKN1A*, *DICER1*, *RB1*, *ANKRD52* and *FAM102A*) as markers during physiological modulation of H19 levels. First, we confirmed their regulation by miR-106a in myocytes (**Supplementary Fig. 18**). Whereas H19 expression was again enhanced by 10- to 12-fold, *CDKN1A*, *DICER1*, *RB1*, *ARID4B*, *ANKRD52* and *FAM102A* mRNAs were all reduced by transfection of a high concentration (100 nM) of miR-106a. *C-MYC* mRNA was selected as a negative control as it has no conserved target sites for miR-106a. Next, we used the progression of myoblast differentiation as an intrinsic H19 overexpression system. Myoblasts were induced to differentiate into myocytes, and the levels of the selected RNAs were measured between days 1 and 5 (**Fig. 5a**). We observed a strong differentiation-dependent induction of H19 up to 100-fold on day 4, as expected[47]. Similarly, the levels of miR-17-5p targets (markers) gradually increased during the time course (1.5- to 5-fold). We also measured the levels of miR-106a/b, miR-17-5p, miR-20a/b and miR-93 as well as that of miR-18a, miR-26a and miR-92a (**Fig. 5b**). 'Myo-miR' miR-26a was induced approximately twofold, as described[48]. The levels of the other miRNAs increased on average twofold, albeit with a high variability, one day after differentiation initiation, which remains to be explained. However, they decreased sharply on day 2 and remained low through day 5. These observations

are consistent with a model where high levels of H19 act as a sponge for these miRNAs, preventing them from targeting mRNAs, which are thus increasingly expressed.

In the event that elevated H19 functioned as a sponge for the miR-17-5p family, we expected H19 RNAi to release H19-bound miRNAs, whose targets might then be repressed. Thus we transfected an siRNA targeting H19 (siH19) into myoblasts and initiated their differentiation to myocytes. The already elevated H19 level was reduced to <10% after 3 d (**Supplementary Fig. 18**). Indeed, five from the six targets were suppressed after 72 h. We further analyzed the effect of H19 suppression on the global expression of predicted miR-17-5p (and let-7) targets. Consistent with our hypothesis, after 24 h, a small but significant repression of all Targetscan-predicted miR-17-5p and let-7-targets was observed compared to all transcripts (**Fig. 5c**). This repression was even more pronounced for mRNAs predicted to be targeted by both miR-17-5p and let-7. Predicted target sets of unrelated miRNAs were not significantly repressed (miR-19, miR-24 and miR-21; P = 0.10, 0.76 and 0.10, respectively). Taken together, these observations are consistent with H19 functioning as a sponge for miR-17-5p and let-7 families.

To add insight into how H19 may serve as a sponge for the miR-17-5p family, we tested a series of plasmids containing various miR-17-5p MREs, including a wild-type pH19 plasmid and a construct harboring four miR-17-5p family MREs in place of the natural let-7 binding sites (pH19-miR-20)[40]. In addition, we mutated the two putative miR-17-5p MREs in H19 separately (pH19-106a/5′mut and pH19-106a/3′mut) and together (pH19-106a/5′3′mut). The five MRE plasmids were transfected into HeLa cells, and we measured their effects on the CDKN1a reporter sensor. Indeed, wild-type pH19 transfection significantly increased reporter expression in a concentration-dependent manner by up to 40% (P = 0.0037; **Fig. 5d**). pH19-106a/5′3′mut was completely inactive, as were the individual single mutants. As expected, the pH19-miR-20 construct produced the strongest effect with an increased luciferase activity of up to 80%. These findings are again in line with the hypothesis that H19 competes for targets of the miR-17-5p family.

DISCUSSION

Computational predictions have been instrumental in uncovering principles of miRNA-dependent mRNA regulation and in guiding experimental efforts[49]. Conversely, experiment-based approaches to identify miRNA-mRNA interactions have fueled the development of computational methods and revealed new modes of interactions. For example, CLIP and CLASH have already exposed the preponderance of noncanonical miRNA binding sites[7,10,32]. A few papers have described efforts to capture miRNA targets using chemically modified mimics[14,18]. However, this approach is fraught with technical challenges. Facing these, we developed a new method, miR-CLIP, that uses a miRNA doubly modified with psoralen and biotin (**Fig. 2a**). We showed in three examples that they functioned similarly to their natural counterparts (**Fig. 1c** and **Supplementary Fig. 2**). We transfected low concentrations of miR-CLIP probes into cells, and, after crosslinking, we used Ago2 immunoprecipitation followed by streptavidin affinity capture to isolate RNAs. For miR-106a and let-7g, the two-step purification provided greater enrichment of high-confidence targets than a single-step protocol (**Fig. 3a** and **Supplementary Fig. 10b**). We then confirmed on the transcriptome level that the miR-CLIP-106a targetome is functional (**Fig. 3b**).

miR-CLIP mimics may not be easily accessible for many laboratories; however, there is an incentive to explore obvious avenues of reagent simplification because alternative efficient methods for capturing targets of individual miRNAs are lacking. The CLASH-type approach comes closest to miR-CLIP and has the advantage of having binding site–level resolution, a feature that could be added to miR-CLIP. However, its low efficacy can make it difficult to evaluate the quality of the identified targets[49]. Extending miR-CLIP to additional miRNAs, particularly those predicted to have predominantly nonseed modes of interaction, and other classes of noncoding RNAs may reveal new biological functions of these RNAs and help to refine prediction algorithms.

miR-CLIP identified the lincRNA H19 as a target of miR-106a. To our surprise, H19 levels were induced rather than decreased upon ectopic delivery of miR-106a into cells (**Fig. 3c** and **Supplementary Fig. 14**), possibly through a conserved miR-17-5p MRE located close to a previously identified binding site for a protein from the IGF2BP family[11]. Thus, IGF2BP and miR-106a may act synergistically to stabilize H19. However, full clarification of this potential interplay requires further investigation. H19 was recently shown to sequester (sponge) let-7 during muscle cell differentiation[40]. In our system, we observed that levels of H19, miR-17-5p family members and their targets change during myoblast differentiation in a fashion suggesting that H19 sequesters miR-17-5p family members and thereby suppresses pro-proliferative (anti-differentiating) mRNAs during the early phases of differentiation (**Fig. 5**).

In summary, we see miR-CLIP as a robust and potentially broadly applicable approach that can be used in conjunction with others to unveil regulators and components of noncoding RNA-containing networks.

Received 7 September 2014; accepted 29 October 2014; published online 22 December 2014

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Gene Expression Omnibus. All of the sequencing data has been uploaded under accession code GSE62681.

References

1. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
2. Lewis, B.P., Burge, C.B. & Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
3. Flynt, A.S. & Lai, E.C. Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat. Rev. Genet.* **9**, 831–842 (2008).
4. Chekulaeva, M. *et al.* miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.* **18**, 1218–1226 (2011).
5. Fabian, M.R. *et al.* miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. *Nat. Struct. Mol. Biol.* **18**, 1211–1217 (2011).
6. Braun, J.E., Huntzinger, E., Fauser, M. & Izaurralde, E. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell* **44**, 120–133 (2011).
7. Chi, S.W., Hannon, G.J. & Darnell, R.B. An alternative mode of microRNA target recognition. *Nat. Struct. Mol. Biol.* **19**, 321–327 (2012).
8. Chi, S.W., Zang, J.B., Mele, A. & Darnell, R.B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **460**, 479–486 (2009).
9. Xia, Z. *et al.* Molecular dynamics simulations of Ago silencing complexes reveal a large repertoire of admissible 'seed-less' targets. *Sci Rep* **2**, 569 (2012); erratum **2**, 909 (2012).
10. Helwak, A., Kudla, G., Dudnakova, T. & Tollervey, D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* **153**, 654–665 (2013).
11. Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
12. Beitzinger, M., Peters, L., Zhu, J.Y., Kremmer, E. & Meister, G. Identification of human microRNA targets from isolated argonaute protein complexes. *RNA Biol.* **4**, 76–84 (2007).
13. Ørom, U.A., Nielsen, F.C. & Lund, A.H. MicroRNA-10a binds the 5′UTR of ribosomal protein mRNAs and enhances their translation. *Mol. Cell* **30**, 460–471 (2008).
14. Lal, A. *et al.* Capture of microRNA-bound mRNAs identifies the tumor suppressor miR-34a as a regulator of growth factor signaling. *PLoS Genet.* **7**, e1002363 (2011).

15. Guo, Y.E. & Steitz, J.A. 3′-Biotin-tagged microRNA-27 does not associate with Argonaute proteins in cells. *RNA* **20**, 985–988 (2014).
16. Thomas, M., Lieberman, J. & Lal, A. Desperately seeking microRNA targets. *Nat. Struct. Mol. Biol.* **17**, 1169–1174 (2010).
17. Nonne, N., Ameyar-Zazoua, M., Souidi, M. & Harel-Bellan, A. Tandem affinity purification of miRNA target mRNAs (TAP-Tar). *Nucleic Acids Res.* **38**, e20 (2010).
18. Baigude, H., Ahsanullah, Li, Z., Zhou, Y. & Rana, T.M. miR-TRAP: a benchtop chemical biology strategy to identify microRNA targets. *Angew. Chem. Int. Edn Engl.* **51**, 5880–5883 (2012).
19. Grosswendt, S. *et al.* Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Mol. Cell* **54**, 1042–1054 (2014).
20. He, L. *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **435**, 828–833 (2005).
21. Brannan, C.I., Dees, E.C., Ingram, R.S. & Tilghman, S.M. The product of the *H19* gene may function as an RNA. *Mol. Cell. Biol.* **10**, 28–36 (1990).
22. Li, Y.M. *et al.* The *H19* transcript is associated with polysomes and may regulate *IGF2* expression in *trans*. *J. Biol. Chem.* **273**, 28247–28252 (1998).
23. Bartolomei, M.S., Zemel, S. & Tilghman, S.M. Parental imprinting of the mouse *H19* gene. *Nature* **351**, 153–155 (1991).
24. Gabory, A., Jammes, H. & Dandolo, L. The *H19* locus: role of an imprinted non-coding RNA in growth and development. *Bioessays* **32**, 473–480 (2010).
25. Gibb, E.A., Brown, C.J. & Lam, W.L. The functional role of long non-coding RNA in human carcinomas. *Mol. Cancer* **10**, 38 (2011).
26. Higuchi, M., Kobori, A., Yamayoshi, A. & Murakami, A. Synthesis of antisense oligonucleotides containing 2′-*O*-psoralenylmethoxyalkyl adenosine for photodynamic regulation of point mutations in RNA. *Bioorg. Med. Chem.* **17**, 475–483 (2009).
27. Guennewig, B. *et al.* Synthetic pre-microRNAs reveal dual-strand activity of miR-34a on TNF-α. *RNA* **20**, 61–75 (2014).
28. Gibcus, J.H. *et al.* MiR-17/106b seed family regulates p21 in Hodgkin's lymphoma. *J. Pathol.* **225**, 609–617 (2011).
29. Jiang, Y. *et al.* miR-106a–mediated malignant transformation of cells induced by anti-benzo[a]pyrene-trans-7,8-diol-9,10-epoxide. *Toxicol. Sci.* **119**, 50–60 (2011).
30. Beitzinger, M. & Meister, G. Experimental identification of microRNA targets by immunoprecipitation of Argonaute protein complexes. *Methods Mol. Biol.* **732**, 153–167 (2011).
31. Khorshid, M., Rodak, C. & Zavolan, M. CLIPZ: a database and analysis environment for experimentally determined binding sites of RNA-binding proteins. *Nucleic Acids Res.* **39**, D245–D252 (2011).
32. Khorshid, M., Hausser, J., Zavolan, M. & van Nimwegen, E. A biophysical miRNA-mRNA interaction model infers canonical and noncanonical targets. *Nat. Methods* **10**, 253–255 (2013).
33. Coley, W. *et al.* Absence of DICER in monocytes and its regulation by HIV-1. *J. Biol. Chem.* **285**, 31930–31943 (2010).
34. Poliseno, L. *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038 (2010).
35. Cesana, M. *et al.* A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* **147**, 358–369 (2011).
36. Hansen, T.B. *et al.* Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384–388 (2013).
37. Memczak, S. *et al.* Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**, 333–338 (2013).
38. Weinmann, L. *et al.* Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs. *Cell* **136**, 496–507 (2009).
39. Jeggari, A., Marks, D.S. & Larsson, E. miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics* **28**, 2062–2063 (2012).
40. Kallen & Amanda, N. *et al.* The imprinted H19 LncRNA antagonizes Let-7 microRNAs. *Mol. Cell* **52**, 101–121 (2013).
41. Shimakami, T. *et al.* Stabilization of hepatitis C virus RNA by an Ago2–miR-122 complex. *Proc. Natl. Acad. Sci. USA* **109**, 941–946 (2012).
42. Wang, Y. *et al.* Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev. Cell* **25**, 69–80 (2013).
43. Nielsen, F.C., Nielsen, J. & Christiansen, J. A family of IGF-II mRNA binding proteins (IMP) involved in RNA trafficking. *Scand. J. Clin. Lab. Invest. Suppl.* **234**, 93–99 (2001).
44. Weinlich, S. *et al.* IGF2BP1 enhances HCV IRES-mediated translation initiation via the 3′UTR. *RNA* **15**, 1528–1542 (2009).
45. Salmena, L., Poliseno, L., Tay, Y., Kats, L. & Pandolfi, P.P. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* **146**, 353–358 (2011).
46. Denzler, R., Agarwal, V., Stefano, J., Bartel, D.P. & Stoffel, M. Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. *Mol. Cell* **54**, 766–776 (2014).
47. Milligan, L. *et al. H19* gene expression is up-regulated exclusively by stabilization of the RNA during muscle cell differentiation. *Oncogene* **19**, 5810–5816 (2000).
48. Dey, B.K., Gagan, J., Yan, Z. & Dutta, A. miR-26a is required for skeletal muscle differentiation and regeneration in mice. *Genes Dev.* **26**, 2180–2191 (2012).
49. Hausser, J. & Zavolan, M. Identification and consequences of miRNA-target interactions—beyond repression of gene expression. *Nat. Rev. Genet.* **15**, 599–612 (2014).

Acknowledgments

This work was supported by the Swiss National Science Foundation (in part) with a joint Sinergia grant (CRSII3_127454) to A.P.G., M.Z. and J.H. as well as a grant to J.H. (CRS205321_124720). We thank B. Schoser (Ludwig-Maximilians-Universität, Munich) for myoblasts, M. Lucic and Y. Wang for help with assays and M. Zimmermann and U. Pradère for synthesis of RNAs. We thank H. Towbin and J.A. Zagalak for helpful discussions and Luca Gebert for graphical layout. We are grateful to Y. Huang (Yale Stem Cell Center) for sharing pH19 expression plasmids.

Author contributions

All authors designed experiments and analyzed data. A. Brunschweiger designed, synthesized and characterized probes. J.I. performed cell assays with myoblasts. B.G. and P.T. performed reporter assays. A. Brunschweiger, J.I. and S.K. performed pull-down experiments. N.M. generated deep-sequencing libraries. A. Brümmer and M.Z. analyzed sequencing data. J.I., A. Brunschweiger, A. Brümmer, M.Z. and J.H. wrote the paper.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information and chemical compound information is available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to M.Z. or J.H.

ONLINE METHODS

Postsynthetic modification of 2′-propargyl–substituted miRNAs by CuAAC. The CPG with the alkynyl-modified RNA was suspended in 0.48 ml of H_2O/MeOH (1:1) in an Eppendorf tube. Subsequently, the azide (1 µmol in 100 µl of DMF), TBTA (500 nmol/0.27 mg in 20 µl of DMF), Na-ascorbate (500 nmol/0.1 mg in 10 µl of H_2O) and $CuSO_4$ (50 nmol/0.0125 mg in 10 µl of H_2O) were added to the suspension in this order. All solutions were prepared as stock solutions directly prior to use. The 4′-azidomethyltrioxsalen was generated *in situ* before cycloaddition through reaction of 4′-chloromethyltrioxsalen (1 µmol/0.27 mg) with 1.2 eq. of NaN_3 (1.2 µmol/0.8 mg) for 1 h at RT. The reaction mixture was shaken overnight at 45 °C in an Eppendorf shaker. The CPG was filtered off and washed subsequently three times with 0.5 ml each of DMF, 0.1 N aqueous EDTA, DMF, MeCN and CH_2Cl_2. Then, it was dried and submitted to standard oligonucleotide workup procedures.

***In vitro* photo crosslinking experiments.** miRNA-psoralen conjugate RNA-4 and its counterstrand RNA-9 were dissolved in a final volume of 200 µl of annealing buffer (2.5 mM Na_2HPO_4, 5 mM NaH_2PO_4, 100 mM NaCl and 0.1 mM Na_2EDTA) at a final concentration of 1 µM. For annealing, the solution was heated to 85 °C, held at that temperature for 5 min and cooled down to room temperature over a period of 2 h. The compounds were irradiated in an open 24-well plate for 5 min at 0 °C (365 nm, 225 J/min; distance of the solution from the lamp: 5 cm). Then, the sample was directly purified by RP-HPLC using the gradient for DMT-off purification. Newly formed peaks were isolated and analyzed by LC/MS.

Cell culture. HeLa (ATCC, CCL-2) cells from LGC (Molsheim, FR) were maintained in Dulbecco's modified Eagle's medium (Life Technologies) supplemented with 10% FBS (FBS; Life Technologies) and 1× antibiotics/antimycotics (Sigma) at 37 °C, 5% CO_2. Primary myoblasts were derived from anonymous donor biopsies at Ludwig-Maximilians-Universität Munich according to the Helsinki declaration. Myoblasts were directly resuspended in differentiation medium for differentiation to myocytes (change from SMG-Medium (PromoCell) supplemented with PromoCell SupplementMix, Gentamicin 50 µg/ml Glutamax (1×, Gibco), FCS 10% (Gibco) to DMEM (Life Technologies), 2% Horse serum (Sigma), Gentamicin (50 µg/ml)). For H19 or Ago2 pulldowns, myoblasts were forced to differentiate to increase endogenous H19 levels.

Affinity purification of miR-106a–crosslinked RNAs. The biotinylated miR-106a–psoralen conjugate RNA-7 and the nonbiotinylated control compound RNA-8 were each diluted in 2 ml of Optimem (Life Technologies, 3 nM final concentration), reverse-transfected with RNAiMax (Life Technologies) to 15-cm dishes of HeLa cells (~1×10^7 = 80% confluency) and incubated for 24 h. Preparation of the beads: 40 µl of magnetic streptavidin beads (Dynabeads Streptavidin C1, Life Technologies) per 15-cm dish was washed according to the manufacturer's protocol and blocked with 1 ml of binding and washing buffer (B&W buffer, 0.5 mM EDTA, 1.0 mM NaCl, 50 U/ml RNAsin, 0.001% NP40, pH 7.5) containing salmon sperm DNA (100 µg/ml), BSA (100 µg/ml) and heparin (0.1 µg/ml) for 1 h at 4 °C. Thereafter, beads were washed five times with binding and washing buffer and pre-equilibrated in 50 µl of NP40 lysis buffer (50 mM HEPES pH 7.5, 150 mM KCl, 0.5% IGEPAL, 0.5 mM DTT, 2 mM EDTA, 50 U/ml RNAsin (complete protease inhibitor, Roche).

Irradiation of the cells and affinity purification. Cells were washed once with 10 ml of PBS, placed on ice and irradiated at 365 nm with 150 mJ (BLX-254 Bio-Link crosslinker (VilberLourmat) equipped with 365-nm lamps). After irradiation, cells were scraped in 1.5 ml ice-cold PBS, pelleted at 200*g* at 4 °C for 5 min and lysed in 1 ml of NP40 lysis buffer for 15 min on ice. The lysate was cleared at 14,000*g* at 4 °C for 15 min and incubated with streptavidin beads for 30 min at 4 °C with gentle agitation. Beads were washed once with NP40 lysis buffer, treated with DNAse (PCR grade, Roche) for 15 min at RT and washed five times with 1 ml of high-salt buffer (B&W buffer).

RNA isolation. The beads were shaken twice with 50 µl of a solution of 95% formamide/10 mM aqueous EDTA, pH 8.2, for 2 min at 65 °C. The solutions were combined, filled with 100 µl of water and 20 µl of 3 N aqueous sodium acetate and purified by chloroform/phenol (Life Technologies) extractions. The aqueous phase was isolated, and RNA was precipitated with 1 ml of 100% EtOH at −20 °C overnight with the addition of 20 µg/ml Glycoblue (Life Technologies). Then, the RNA pellet was washed with 70% EtOH, centrifuged at 14,000*g* for 10 min at 4 °C and dissolved in 20 µl of water. After reverse transcription (High-Capacity Reverse Transcriptase Kit, Life Technologies) of the isolated RNA, enrichment of target genes as normalized to GAPDH was measured by qRT-PCR (FastStart Universal SYBR Green Master, Roche Applied Sciences) on a Roche Light-Cycler 480 following the manufacturer's protocol.

miR-CLIP two-step purification protocol. Transfection of the biotinylated miR-106a–psoralen conjugate 7 and the nonbiotinylated control compound RNA-8 was performed as described above.

Preparation of Prot G beads and Ago2 immunoprecipitation. 80 µl of Prot G Dynabeads (Life Technologies) per 15-cm dish was washed twice with 1 ml of citrate-phosphate buffer (25 mM citric acid, 66 mM Na_2HPO_4, pH 5.0). The Ago2 antibody (40 µg per 80 µl of Prot G Dynabeads) was immobilized in a total volume of 500 µl of citrate-phosphate buffer by gentle rolling for 1 h at 4 °C. Then, the beads were washed three times with 1 ml of NP40 lysis buffer and blocked for 1 h at 4 °C with 1 ml of NP40 lysis buffer containing BSA (10 µg/ml). The beads were washed three times with 1 ml of NP40 lysis buffer and resuspended in 50 µl therein.

Irradiation of the cells and tandem purification. Cells were washed once with 10 ml of PBS, placed on ice and irradiated twice at 254 nm (CL-1000 UltraviolettCrosslinker, UVP) with 100 mJ, followed by irradiation at 365 nm with 150 mJ. Then, cells were harvested as described above and incubated with the Ago2 (clone 9A11; Ascenion, Munich, Germany) antibody-coupled Prot G beads for 1 h at 4 °C with gentle rolling. The beads were washed five times with each 1 ml of IP wash buffer (50 mM HEPES, pH 7.5, 300 mM KCl, 0.05% NP40, 0.5 mM DTT, complete protease inhibitor (Roche)). After the last washing step, 200 µl digest buffer (100 mM Tris-HCl, pH 7.5, 150 mM NaCl, 12.5 mM EDTA) containing 240–440 µg proteinase K (recombinant PCR grade solution, Roche, Rotkreuz, Switzerland) was added to each sample and digested at 65 °C for 15 min. RNA from the proteinase K digest was isolated by chloroform/phenol (Life Technologies) extraction and precipitation as above. For streptavidin affinity purification, the RNA was redissolved in 20 µl of binding and washing buffer, and affinity purification was performed as above except that the streptavidin beads were resuspended in 50 µl of binding and washing buffer after the blocking step. Input RNA, RNA isolated from Ago2 IP of either RNA-7 or mock transfected HeLa cells (ribo-depleted only one time) and miR-CLIP–purified RNA (not ribo-depleted) were submitted for Illumina sequencing. A sample of miR-CLIP–purified RNA was analyzed by qRT-PCR to control for successful enrichment of a set of validated miR-106a targets versus miR-CLIP-purified RNA from control RNA-8–transfected cells. Before cDNA library preparation[32] for Illumina sequencing, total input RNA was depleted of ribosomal RNA two times. The samples of total input RNA (before and after rRNA depletion) as well as the two Ago2 IP samples and the Ago2/Strep samples were analyzed using an Agilent Bioanalyzer (**Supplementary Fig. 5**) to confirm the quantitative removal of rRNA and no obvious degradation.

Luciferase assays. 1×10^4 HeLa cells were seeded in DMEM (Life Technologies) in 96-well plates. First, transfection of a negative control (Ambion Silencer negative control no. 5: miCon), pre-miR-106a, miR-106a, RNAs 1–7 or a positive control (siRNA against *Renilla* luciferase: siRen) was done at different concentrations (0 nM, 2 nM, 9 nM, 36 nM) after 24 h. Another 24 h later, we performed a second transfection of 20 ng dual luciferase reporter plasmid (Promega psiCHECK2), including a 5p sensor site reverse complementary to the corresponding miRNA in the indicated mRNA surrounding it, and relative luciferase activity was measured on Berthold Mithras LB940 Luminometer. H19 reporter plasmids were generated by DNA synthesis and cloned into psiCHECK2 vector (H19 transcript, accession no. BC04007; the 5′ fragment encompasses nucleotides 190–292, and the 3′-fragment encompasses nucleotides 1541–1622). Four mutations were introduced into their respective miR-106a MRE seed regions: TtoC at position 2, CtoT at position 4, AtoG at position 5 and CtoA at position 6. Additionally, the following mutations were introduced into each wild-type H19 transcript and into the 5′ reporter to produce IGF2BP single and IGF2BP/miR-106a double mutants: ACGTACG was substituted for CATTCATC five bases upstream of the last miR-106a seed base. These reporters served as seed knockout and IGF2BP PRE mutation negative controls (a schematic overview of the H19 transcript and the cloned reporter constructs is given in **Fig. 4e**). All values were normalized to the level of transfection efficiency measured through the expression of a luciferase not targeted by the miRNA and additionally to the corresponding 0 nM transfection (representing 1 at the *y* axis). The error bars represent the s.d. of at least three independent transfections (statistical analysis by unpaired, two-tailed Student's *t*-test). The QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent)

was used to mutate pH19 at the desired sites. pH19 was mutated in the 3′, 5′ or 3′/5′ MREs for miR-106a. The 5′ miR-106a MRE (CACTTTT) and 3′ MRE (GCACTTT) for miR-106a on the H19 sequence were mutated to generate plasmids pH19-5′-miR-106a-mut (CTGCAGT), pH19-3′-miR-106a-mut (GCAGTTT) or pH19-5′,3′-miR-106a-mut (CTGCAGT/GCAGTTT). The generated mutations are restriction sites for the PstI enzyme, which recognizes the sequence CTGCAGT. 5,000 HeLa cells were co-transfected with 80 ng CDKNA1 sensor reporter and 35 ng, 70 ng or 120 ng of pFLAG and pH19 expression plasmids in 96-well format and measured after 18 h.

Binding motifs and modes of miR-106a targets identified by miR-CLIP with RNA-7. The analysis of the 3′ UTRs of transcripts that were highly abundant in miR-106a miR-CLIP revealed a high number of occurrences of matches to the seed region of miR-106a (**Supplementary Fig. 11**). Moreover, a relatively high number of matches could also be detected for subsequences located at the 3′ end of miR-106a (**Supplementary Fig. 11**). Both 5′ and 3′ complementary motifs were significantly enriched in the 3′ UTRs of miR-CLIP targets compared to randomized sequences with the same nucleotide composition (**Supplementary Fig. 11**). This may indicate an abundance of '3′-compensatory' or '3′-supplementary' binding or may be a consequence of the modification in the miRNA that favors the identification of sites with more extensive binding to the miRNA 3′ end.

Pulldown of H19 with 3′-biotinylated 2′-methoxyRNA capture probes[50]. HeLa cells or differentiating myoblasts (10-cm dish, ~8×10^6 cells per treatment) were lysed for 15 min on ice with 1.0 ml of NP40 lysis buffer (50 mM HEPES, pH 7.5, 150 mM KCl, 0.5% IGEPAL, 0.5 mM DTT, 2 mM EDTA, 50 U/ml RNAsin, complete protease inhibitor (Roche)). The lysates were centrifuged for 10 min at 14,000*g* at 4 °C. An input sample was removed (50 µl). Then, 50 µl of a stock solution containing three 3′-biotinylated 2′-methoxy-RNA capture probes **15**–**17** (combined concentration: 2,000 nM in water) was added to the lysates to a final concentration of 100 nM. The lysates and control samples without capture probes were added to 200 µl of streptavidin beads per sample.

Preparation of streptavidin beads. An amount of 200 µl of magnetic streptavidin beads (Dynabeads streptavidin C1, Life Technologies) per 10-cm dish was washed according to the provider's protocol and incubated in with 1.0 ml of NP40 buffer containing salmon sperm DNA (100 µg/ml) for 1 h at 4 °C with rotation. Prior to affinity purification, the beads were washed five times with binding and washing buffer and resuspended in 50 µl of NP40 buffer per 10-cm dish. The cell lysates containing the biotinylated 2′-methoxy RNA capture oligonucleotides and the control lysates were incubated for 2 h with the streptavidin beads with gentle rolling at 4 °C. Then, the beads were washed once with NP40 lysis buffer, treated with DNAse, and washed five times with each 1 ml of binding and washing buffer (high salt, 5 mM Tris-Cl (pH 7.5), 0.5 mM EDTA and 1 M NaCl). To isolate the RNA, the beads were incubated two times with each 50 µl of 95% formamide/5% 10 mM aq. EDTA at 65 °C for 2 min. The eluate was transferred to new tubes, and 100 µl of water and 20 µl of 3 N Na-acetate were added. The aqueous solution was washed with 220 µl chloroform-phenol (Life Technologies). The upper layer was isolated, 1 ml of EtOH was added, and the RNA was precipitated at 20 °C overnight with the addition of 20 µg/ml Glycoblue (Life Technologies).

Small RNA sequencing. Anti-H19 pulled-down RNA of three biological replicates was pooled in a 1:1:1 ratio and processed for small RNA sequencing. Additionally, four calibrator oligonucleotides (Cal01-04 5 fMol each) were added as a reference as described previously[11]. Briefly, RNA was dephosphorylated using FastAP (Fermentas) and radiolabeled as described above. Subsequently, RNA was separated with denaturing PAA (15%) gel electrophoresis, and fractions between 18 nt and 25 nt in size were excised from the gel. Further sample processing was done according to the protocol described above (miR-CLIP library preparation).

Analysis of sequencing data sets. Deep- sequencing data from two biological replicates were uploaded to the ClipZ server (http://www.clipz.unibas.ch/)[31] for mapping and annotation of sequence reads (**Supplementary Table 2**). For all samples, except for the doubly purified 'miR-CLIP' sample, more than 85% of reads were found to be of good quality (i.e., containing at most one ambiguously identified nucleotide), and more than 70% of the reads could be mapped. For the doubly purified miR-106a miR-CLIP samples, only ~40% of reads were of good quality, one possible reason being that the complex two-step RNA purification protocol and the harsh elution method disrupt the biotin-streptavidin interaction and result in amplification of contaminating RNA. Nevertheless, the overall read number and quality was sufficient to capture a miR-106a and let-7g targetome-enriched pool of RNAs compared to 'Ago2' alone. A total of 16,620 (for let-7g miR-CLIP 13,740) transcripts were identified in at least one of the samples, and for 10,319 (1,273 for let-7g) transcripts, reads were found in all eight samples (**Supplementary Data Sets 1** and **2**). The reproducibility of the biological replicates was high (Pearson (Spearman) correlation coefficients for transcript abundance levels in two replicates for miR-106a were 0.99 (0.92), 0.83 (0.95), 0.80 (0.96) and 0.86 (0.91) (**Supplementary Fig. 6**), and those for let-7g were 0.92 (0.88), 0.36 (0.56), 0.48 (0.53) and 0.53 (0.35) for 'Input', 'Mock', 'Ago2' and 'miR-CLIP', respectively).

To confirm that we enriched for miR-106a target transcripts in the miR-CLIP experiment using RNA-7 in the double compared to the single purification, we compared the fraction of predicted targets of miR-106a and let-7 targets as a function of the rank of the transcript enrichment in miR-CLIP relative to Ago2-IP. The miRNA targets were obtained based on TargetScan's probability of preferentially conserved targeting (PCT). We traversed the list of mRNAs sorted by their enrichment in a particular setting with a sliding window of 1,000 transcripts and computed the fraction of predicted targets of a given miRNA among the 1,000 transcripts in the window. **Supplementary Figure 9** shows the median enrichment on the *x* axis and the proportion of predicted miRNA targets on the *y* axis for each of these sliding windows. The results indicate that miR-106a and let-7a have a comparable fraction of targets among the transcripts that are enriched in the Ago2-IP compared to input RNA (**Supplementary Fig. 9**). The fraction of miR-106a targets increases relative to the fraction of let-7a miRNA targets among the transcripts with the highest enrichment in miR-106a miR-CLIP relative to Input (**Supplementary Fig. 9**) as well as among the transcripts with the highest enrichment in the miR-106a miR-CLIP compared to Ago2-IP (**Supplementary Fig. 9**). These results demonstrate that miR-106a miR-CLIP enriched miR-106 targets.

Significantly enriched transcripts in 'miR-CLIP' versus 'Ago2' were extracted using DESeq[51]. A cutoff at $P < 0.05$ resulted in 644 transcripts (mRNAs and lncRNAs) that were significantly enriched (**Supplementary Data Sets 1** and **3**).

Data sets used for comparisons. Predicted target transcripts of miR-106a and let-7 were obtained on the basis of TargetScan's probability for preferentially conserved targeting (PCT)[52]. The top 600 predicted miR-106a target transcripts of the following computational prediction methods were further used for comparisons: ElMMo[53], miRanda[54], PITA[55], RNA22 (ref. 56) and TargetScan's context score[57]. Experimental miR-106a target transcripts identified by CLASH were taken from the supplement of ref. 10.

Oligonucleotides. A complete list of used miR mimics, anti-miRs, cloning and SYBRgreen qRT-PCR primers are provided in **Supplementary Table 5**.

Statistical analysis. For qRT-PCR and Luciferase assay analysis a two-tailed, paired Student's *t*-test was applied (GraphPad *t*-test calculator).

50. Hassan, T. *et al.* Isolation and identification of cell-specific microRNAs targeting a messenger RNA using a biotinylated anti-sense oligonucleotide capture affinity technique. *Nucleic Acids Res.* **41**, e71 (2013).
51. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
52. Friedman, R.C., Farh, K.K., Burge, C.B. & Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92–105 (2009).
53. Gaidatzis, D., van Nimwegen, E., Hausser, J. & Zavolan, M. Inference of miRNA targets using evolutionary conservation and pathway analysis. *BMC Bioinformatics* **8**, 69 (2007).
54. Betel, D., Koppal, A., Agius, P., Sander, C. & Leslie, C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.* **11**, R90 (2010).
55. Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U. & Segal, E. The role of site accessibility in microRNA target recognition. *Nat. Genet.* **39**, 1278–1284 (2007).
56. 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).
57. Garcia, D.M. *et al.* Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat. Struct. Mol. Biol.* **18**, 1139–1146 (2011).

doi:10.1038/nchembio.1713