

# A functional assay for microRNA target identification and validation

Joop Gäken<sup>1,\*</sup>, Azim M. Mohamedali<sup>1</sup>, Jie Jiang<sup>1</sup>, Farooq Malik<sup>1</sup>, Doris Stangl<sup>1</sup>, Alexander E. Smith<sup>1</sup>, Constantinos Chronis<sup>1</sup>, Austin G. Kulasekararaj<sup>1</sup>, N. Shaun B. Thomas<sup>1</sup>, Farzin Farzaneh<sup>1</sup>, Mahvash Tavassoli<sup>2</sup> and Ghulam J. Mufti<sup>1</sup>

<sup>1</sup>Department of Haematological Medicine, King's College London, Rayne Institute, London SE5 9NU, UK and

<sup>2</sup>Department of Oral Pathology, Head and Neck Oncology Group, King's College London, Dental Institute, Guy's Hospital Campus, London SE1 9RT, UK

Received August 29, 2011; Revised December 20, 2011; Accepted January 23, 2012

## ABSTRACT

**MicroRNAs (miRNA) are a class of small RNA molecules that regulate numerous critical cellular processes and bind to partially complementary sequences resulting in down-regulation of their target genes. Due to the incomplete homology of the miRNA to its target site identification of miRNA target genes is difficult and currently based on computational algorithms predicting large numbers of potential targets for a given miRNA. To enable the identification of biologically relevant miRNA targets, we describe a novel functional assay based on a 3'-UTR-enriched library and a positive/negative selection strategy. As proof of principle we have used mir-130a and its validated target MAFB to test this strategy. Identification of MAFB and five additional targets and their subsequent confirmation as mir-130a targets by western blot analysis and knockdown experiments validates this strategy for the functional identification of miRNA targets.**

## INTRODUCTION

MicroRNAs (miRNA) regulate key biological events in both normal and cancer cells [reviewed in (1,2)]. Human miRNAs act mainly by inhibiting target mRNA translation and to a lesser extent by promoting the degradation of the target mRNA [reviewed in (3,4)]. To obtain insight into the function of miRNAs, considerable effort has gone into writing different computational algorithms that identify miRNA targets (5). However, a major drawback of these prediction models is a substantial false positive

rate and an inevitable bias due to reliance on the few known miRNA/target mRNA interactions (6). Furthermore, the available algorithms are based on several parameters of which homology between the 5'-seed sequence of the miRNA and a complementary region in the 3'-UTR of the target mRNA is an important component. However, perfectly matched miRNA seeds are neither necessary nor sufficient for all functional miRNA-target interactions. Data obtained by immunoprecipitation of miRNA containing ribonucleoprotein complexes have shown that only 30–45% of the recovered miRNAs contain perfectly matched seed sequences corresponding to the 3'-UTR (7,8). Furthermore, the genetically validated targets for *let-7* in *Caenorhabditis elegans* such as *lin-41*, *pha-4* and *let-60* contain wobbles or bulges in the seed region (9). Recently it has been shown that mir-24 regulates several genes by binding to seedless 3'-UTR miRNA recognition elements and that mir-146a regulates murine STAT1 in the absence of canonical target seed sequences (10,11).

The lack of sensitivity and specificity of the developed computational algorithms is clearly shown by the fact that for the 1048 human miRNAs identified (miRBase release 16 September 2010), experimentally validated targets have only been reported for 206 miRNAs (miRecords at <http://mirecords.umn.edu/miRecords/>). Therefore, the vast majority of confirmed miRNAs lack even a single validated target. Hence, there is a clear need for methodologies that identify and validate the functional targets of specific miRNAs. Here, we describe a novel, selection-based functional assay for the identification of miRNA targets. This strategy makes no assumptions based on previously identified sequences, but relies directly on down-regulation, by a miRNA, of a selectable

\*To whom correspondence should be addressed. Tel: +44 207 8485839; Fax: +44 207 8485905; Email: joop.gaken@kcl.ac.uk  
Present address:

Constantinos Chronis, Department of Biological Chemistry, UCLA School of Medicine, 615 Charles E. Young Drive South, BSRB 390D, Los Angeles, CA 90095, USA.

marker expressed in-frame with a library of 3'-RNA sequences. Cells that either do not express or express low levels of the miRNA of interest are transfected with a 3'-UTR library inserted downstream of a TKzeo fusion gene in the plasmid p3'TKzeo. Zeocin selection results in a population of cells that express the TKzeo fusion protein and are resistant to zeocin and sensitive to Ganciclovir (GCV). The zeocin-resistant cells are next transfected with the miRNA of interest cloned into a selectable expression vector and transduced cells are isolated. Subsequent GCV treatment selects for cells that have down-regulated the TKzeo fusion protein expression either by inhibition of translation or mRNA cleavage. The 3'-UTR sequences present downstream of the TKzeo fusion gene were isolated from the GCV resistant cells by PCR amplification and identified by sequencing. Using mir-130a as a model, we were able to isolate five target genes and validated these by western blot and mir knockdown experiments. The procedure was further validated through identification of targets for mir-10a.

## MATERIALS AND METHODS

### Plasmids and library construction

Plasmid p3'TKzeo was constructed by digestion of psectag2 (Invitrogen) with NheI, filling-in and DraIII digestion which left the vector backbone with the CMV promoter and the SV40 polyA. TKzeo was removed as an NcoI, filled-in SalI fragment from plasmid pUT102 (Cayla, France). Vector and insert were ligated in a 3-way ligation with a linker that introduced 2 SfiI sites (underlined) that differ in their interpalindromic sequence enabling directional cloning of inserts with the same SfiI sites.

TCGAGGCCATTAAGGCCGCGCCTCGGCCCA  
CTTC

Plasmid p3'TKzeo was verified by Sanger sequencing, digested with SfiI and gel cleaned.

A cDNA library derived from oligo dT primed total RNA from human brain directionally cloned into the same SfiI sites as above in the mammalian expression vector pEXPI was obtained (Clontech). The brain cDNA library was digested with SfiI and cDNAs <2.5 Kb were isolated. The cDNA was gel purified and ligated into SfiI-digested p3'TKzeo vector. The ligation mixture was electroporated into pre-aliquoted E-Shot DH10B T1 electrocompetent *Escherichia coli* (Invitrogen), according to the manufacturer's protocol. A total of 600 000 individual clones were expanded as single colonies on 20 square 400 cm<sup>2</sup> plates. The bacterial colonies were scraped off the plates, pooled, expanded and plasmid DNA was isolated (Qiagen Giga prep).

Mir-130a was cloned by PCR amplification of a 974-bp fragment from human genomic DNA containing the mir-130a pre-miRNA hairpin with the following primers: forward CGGATCCTTGAGGCCTAGAGCTCTGCT, reverse TGAATTCGAGATGGCATCAAAA introducing BamHI and EcoRI restriction sites. The PCR fragment was TOPO-TA cloned (Invitrogen) and

sequenced. Correct plasmid DNA was digested with BamHI and EcoRI and cloned into the retroviral vector pBabepuro.

Mir-10a was cloned similarly as a 998-bp fragment with the following primers: forward TGGATCCTCACAAAC TAAAGGCTTGGAG and reverse TGAATTCCAATAA CACTCCTTGGTAATCG.

Part of the MAFB 3'-UTR containing the mir-130a targetsite was cloned from human genomic DNA by PCR amplification with the following forward and reverse primers: GGCCATTAAGGCCGCGCTGCGCTA ATTGTAGGA GGCCGAGGCGGCCCTCGTTTCTG ATGCAGGACA.

These primers introduce SfiI sites for directional cloning into p3'TKzeo. The PCR product was TOPO-TA cloned sequenced and cloned into p3'TKzeo.

### Transfection and selection

The plasmid library (100 µg/flask) was transfected using the BES co-precipitation procedure into 10<sup>8</sup> MCF7 cells, in 24 triple layer flasks (NUNC) (4 × 10<sup>6</sup> cells per flask). Forty-eight hours post-transfection the cells were selected, for 10 days, in 500 µg/ml zeocin. The cells were trypsinized, pooled and expanded in the presence of 500 µg/ml zeocin. Of this mixed population (1.6 × 10<sup>7</sup> cells) were plated in four triple layer flasks and transfected with either pBabepuro or pBabepuro expressing mir-130a and selected in 1 µg/ml puromycin for 48 h.

Cells were then seeded at a density of 10<sup>6</sup> in 10-cm dishes and selected with 8 µM clinical grade GCV (Roche) for 10 days. Pooled populations of GCV selected cells containing mir-130a were expanded in the presence of 8 µM GCV and 1 µg/ml puromycin. Genomic DNA was isolated (Qiagen DNAeasy kit) from the GCV selected cells and the inserts amplified by PCR with p3'TKzeo specific primers flanking the cloning site (Forward AACTGCGTGCACCTTCGTG and Reverse A AGGCGATTAAGTTGGGTAA), cloned into the TopoTA vector and sequenced using M13F and M13R sequencing primers. Similarly, 8 × 10<sup>6</sup> zeocin-resistant MCF7 cells were transfected with pBabepuro-mir-10a and treated as above.

### ABI solid next-generation sequencing

To analyse the complexity of the library of zeocin-resistant MCF7 cells inserts were PCR amplified with the above primers. The PCR product was fragmented to an average length of 150 bp and ends repaired as per ABI Solid protocols. Adapters were then ligated to the library DNA containing sequences necessary for ABI solid sequencing. Ligation products of ~200 bp were then fractionated by acrylamide gel electrophoresis and PCR amplified according to ABI protocols using primers specific for the ligated adapter. Sequencing was performed on an ABI SOLiD<sup>tm</sup> v3 system with a read length of 35 bp. Number of reads was 8 233 395.

### Western blot analysis

Cells were lysed in RIPA buffer (Sigma) freshly supplemented with protease inhibitor cocktail (Sigma) on ice

for 30 min. The lysate was diluted with 2× Laemmli buffer (Sigma) and 10% β-mercaptoethanol was added. Samples were boiled for 5 min before gel loading. Proteins were separated on 10% SDS–polyacrylamide gels, then transferred to nitrocellulose membranes (GE Healthcare) and blocked with 5% non-fat dry milk for 1 h at room temperature. Immunoblotting was performed with the following primary antibodies: TPT1 (ab58362, mouse monoclonal IgG, 1:1000; Abcam), CYP27A1 (ab64889, rabbit polyclonal IgG 1:1000, Abcam), MAFB (ab56242, rabbit polyclonal IgG 1:2000, Abcam) and KIFAP3 (12700-1-AP, rabbit polyclonal IgG 1:1500; ProteinTech Group, Cambridge, UK). The membrane was washed with 3× TBS-0.05% Tween-20 buffer and incubated with a horseradish peroxidase-conjugated secondary antibody against the host IgG. Subsequently, the blot was developed using an enhanced chemiluminescent (ECL plus) detection system (GE Healthcare, UK) and exposed to ECL film (GE Healthcare, UK). The blot was stripped with ReBlot plus strong antibody stripping buffer (Millepore, UK) and re-probed with a γ-tubulin antibody (sc31787, goat polyclonal IgG 1:500; Santa Cruz Biotech) for loading control.

### QRT-PCR

Total RNA was isolated using Tri-reagent (Invitrogen) and quantified by absorbance at 260 nm. Gene-specific primers for mir-10a, mir-130a and RUN6B (All from Applied Biosystems) were used to generate cDNA from 5 ng of RNA using the TaqMan miRNA Reverse Transcription Kit (Applied Biosystems). QPCR was performed on this product using the Taqman Universal Mastermix-No UNG (Applied Biosystem) according to the manufacturer's protocol in an ABI 7900 HT Sequence Detection System. The delta Ct (relative expression) was calculated after normalizing the target to RNU6B.

### Mir-130a knockdown

The mir-130a knockdown experiments were performed by the reverse transfection procedure according to the manufacturer's instructions, with Lipofectamine RNAiMAX (Invitrogen).

MCF-7 and NIH/3T3 cells cultured in Dulbecco's modified Eagle's medium (DMEM PAA) supplemented with 10% Foetal bovine serum (PAA), 1× penicillin/streptomycin/glutamine (Sigma), 1 mM Pyruvate (Sigma). Prior to addition to the test wells, cells were transferred to fully supplemented media without antibiotics.

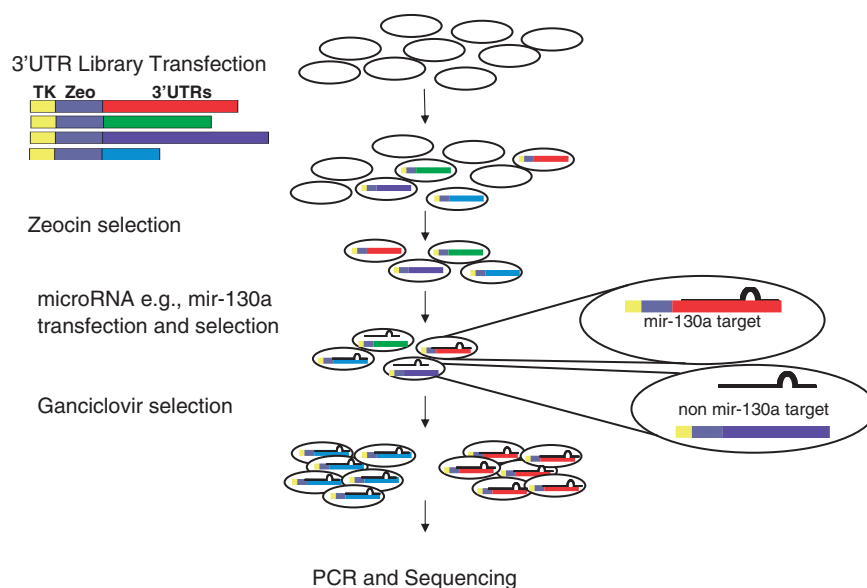
In six-well plates, 500 μl of DMEM medium containing Glutamine and Pyruvate, but no foetal calf serum or antibiotics was added to the wells, followed by 3.75 μl (25 nM final) of MIRIDIAN hairpin mir-130a inhibitors or control inhibitor (Dharmacon) to the appropriate wells followed by 5 μl of Lipofectamine RNAiMAX. The resultant mixture was incubated at room temperature for 15 min. After incubation, 3 × 10<sup>5</sup> cells in 2.5 ml of fully supplemented DMEM media without antibiotics were added to each well. All conditions were setup in duplicate.

Control conditions of cells only (without any additives) and cells +5 μl of Lipofectamine RNAiMAX, were also included in the experimental design. Cultures were incubated in humidified CO<sub>2</sub> incubator at 37°C for 48 h after which protein extracts were prepared for western blot analysis.

## RESULTS

### Library cloning and transfection

Cells that do not express the miRNA of interest are transfected with a library of genes cloned downstream of a TKzeo fusion gene that provides resistance to zeocin and sensitivity to GCV (see Figure 1 for a schematic representation of the target identification strategy).



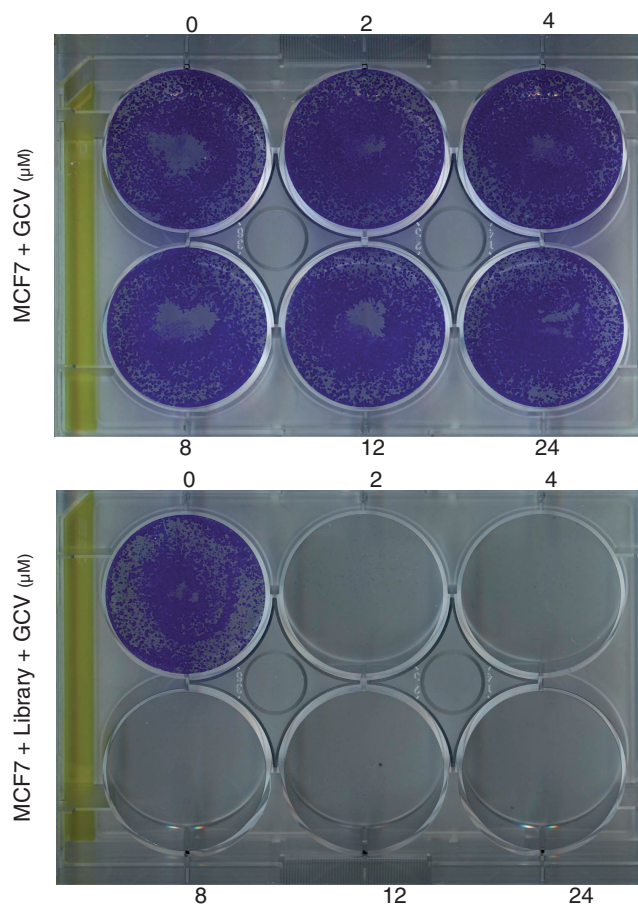
**Figure 1.** Schematic representation of the functional assay for miRNA target discovery.

Zeocin selection results in the isolation of transduced cells (i.e. the target cell library). The subsequent introduction of the miRNA of interest into the target cell library results in the down-regulation of TKzeo in cells containing a cDNA with a target site for the introduced miRNA abrogating the zeocin resistance but by the same token gaining resistance to GCV due to reduced expression of TKzeo.

To generate a cDNA library with the largest number of individual genes present, we cloned a brain cDNA library downstream of a TKzeo fusion gene. Brain and testis are the two tissues expressing the largest number of genes (12). cDNA was removed from a commercially available, Oligo dT primed, brain cDNA library and directionally cloned downstream of TKzeo in the plasmid p3'TKzeo (Supplementary Figure S1). Individual bacterial colonies (600 000) were expanded. Analysis of 72 individual clones confirmed the presence of different length cDNA in all 72 clones tested (results not shown). As proof of concept, we used mir-130a and mir-10a which have verified targets in the transcription factor *MAFB* (v-maf musculoaponeurotic fibrosarcoma oncogene homolog B) and HoxA1 (13) and spiked the 3'-UTR library with 10% of p3'TKzeo with a *MAFB* 3'-UTR. This library was transfected into MCF7 cells, which express very low levels of mir-130a and mir-10a, (Supplementary Figure S2) and selected in 500 µg/ml zeocin. This resulted in a library of approximately 40 000 independent MCF7 clones. PCR amplification and Next-generation sequencing of the inserts present in these clones resulted in 9873 unique hits representing 5626 different genes present in the library of approximately 40 000 MCF7 clones of which 5207 were protein coding genes and 419 were derived from non-coding mRNAs (see Supplementary Data).

### GCV sensitivity

Initially three constructs of a selectable marker fused in-frame to HSVtk were tested for resistance to the selectable marker and sensitivity to GCV. The constructs used were HyTK, TKpuro and TKzeo conferring resistance to hygromycin, puromycin and zeocin respectively. TKzeo conferred the highest sensitivity to GCV and was therefore selected for the library construction (Supplementary Figure S3). To analyse the GCV sensitivity of the zeocin-resistant cells normal MCF7 cells and zeocin-resistant cells were plated and incubated with different concentrations of GCV. After 10 days the plates were stained, clearly showing that untransduced MCF7 cells are resistant to GCV up to the highest concentration, 24 µM, tested (Figure 2). In contrast, zeocin resistant MCF7 cells were sensitive to GCV even at the lowest concentration, 2 µM, tested. Conversely, it was shown that down-regulation of TKzeo by binding of mir-130a to the *MAFB* UTR cloned downstream of TKzeo resulted in zeocin sensitivity of these MCF7 cells (Supplementary Figure S4). To test the stability of the sensitivity to GCV over time of the zeocin-resistant MCF7 cells, zeocin was removed at Day 0 and at Days 1, 30 and 60 after zeocin removal cells were selected in 8 µM GCV for 10 days. The results show that in cells that had been without zeocin selection pressure for 30 or 60 days false

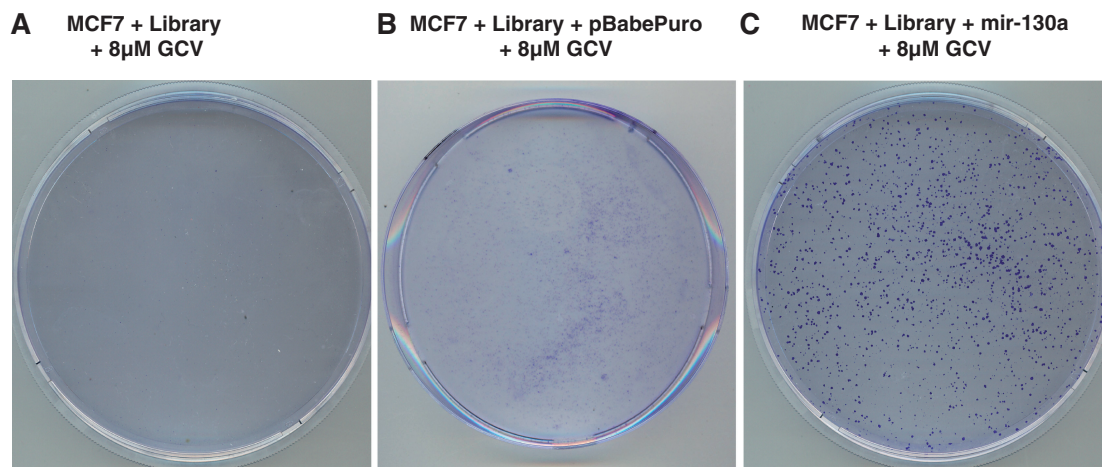


**Figure 2.** MCF7 cells were transfected with the plasmid library and selected in 500 µg/ml zeocin. Zeocin resistant and normal MCF7 cells were selected in different concentrations of GCV showing GCV sensitivity of the library transfected cells.

positive GCV-resistant clones appear. However, in cells that have only been without zeocin for 1 day no GCV-resistant colonies develop. Renewed zeocin selection at Day 60 for 10 days followed by GCV selection restored GCV sensitivity to the cells (Supplementary Figure S5).

### miRNA target identification

Zeocin-resistant MCF7 cells were transfected with pBabepuro-mir-130a and after 48 h selected in 1 µg/ml puromycin resulting in approximately 10 000 individual MCF7 clones. As control zeocin-resistant MCF7 cells were transfected with empty vector pBabepuro. After 96 h 8 µM GCV was added and cells were selected for 10 days in both puromycin and GCV which resulted in GCV-resistant clones in mir-130a transfected cells but not in control pBabepuro transfected cells (Figure 3). Genomic DNA was isolated from the GCV-resistant cells, PCR amplified, TOPO-TA cloned and 96 of the resulting bacterial clones were sequenced. Of these 96 clones, 70 contained a *MAFB* insert. In addition, we identified *TPT1* (tumour protein translationally controlled 1) in 3 clones, *PRR14* (proline rich 14) in 3, *KIFAP3* (kinesin-associated protein 3) in 3, *MITD1* (microtubule



**Figure 3.** GCV selection of library transfected cells with or without mir-130a. Cells ( $10^6$  MCF7 cells) transfected with p3'TKzeo+library (A) or MCF7 cells transfected with p3'TKzeo+library+pBabepuro empty vector (B) or transfected with p3'TKzeo+library+pBabepuro-mir-130a (C) were plated in 10-cm plates and selected in  $8\ \mu\text{M}$  GCV. After 10 days plates were stained with Coomassie blue.

**Table 1.** Identity and frequency of targets isolated for mir-10a and mir-130a

mir-130a			mir-10a		
Accession number	Name	Number of clones	Accession number	Name	Number of clones
NM_005461.3	MAFB	70	NM_001199214.1	STMN2	9
NM_003295.2	TPT1	3	NM_007044.3	KATNA1	4
NM_024031.2	PRR14	3	NM_001014809.1	CRMP1	3
NM_014970.3	KIFAP3	3	NM_005522.4	HoxA1	3
NM_138798.1	MITD1	2			
NM_000784.3	CYP27A1	2			

interacting and transport domain containing 1) in 2, *CYP27A1* (cytochrome P450 family 27 subfamily A polypeptide 1) in 2. The remainder of the tested clones had no insert in the Topo TA vector (Table 1). To further validate the procedure an identical small scale experiment was performed to identify targets for mir-10a, sequencing of 24 bacterial clones identified 4 putative targets, STMN2 (Stathmin like 2) in 9 clones, KATNA1 (Katanin p60) in 4, and CRMP1 (Collapsin response mediator protein 1) in 3. A further three clones contained HoxA1 (Homeobox A1) a validated target for mir-10a (13) (Table 1). Expression of both mature mir-130a and mir-10a in the transfected cells was measured by QRT-PCR 48 h after transfection and after puromycin selection (Supplementary Figure S6).

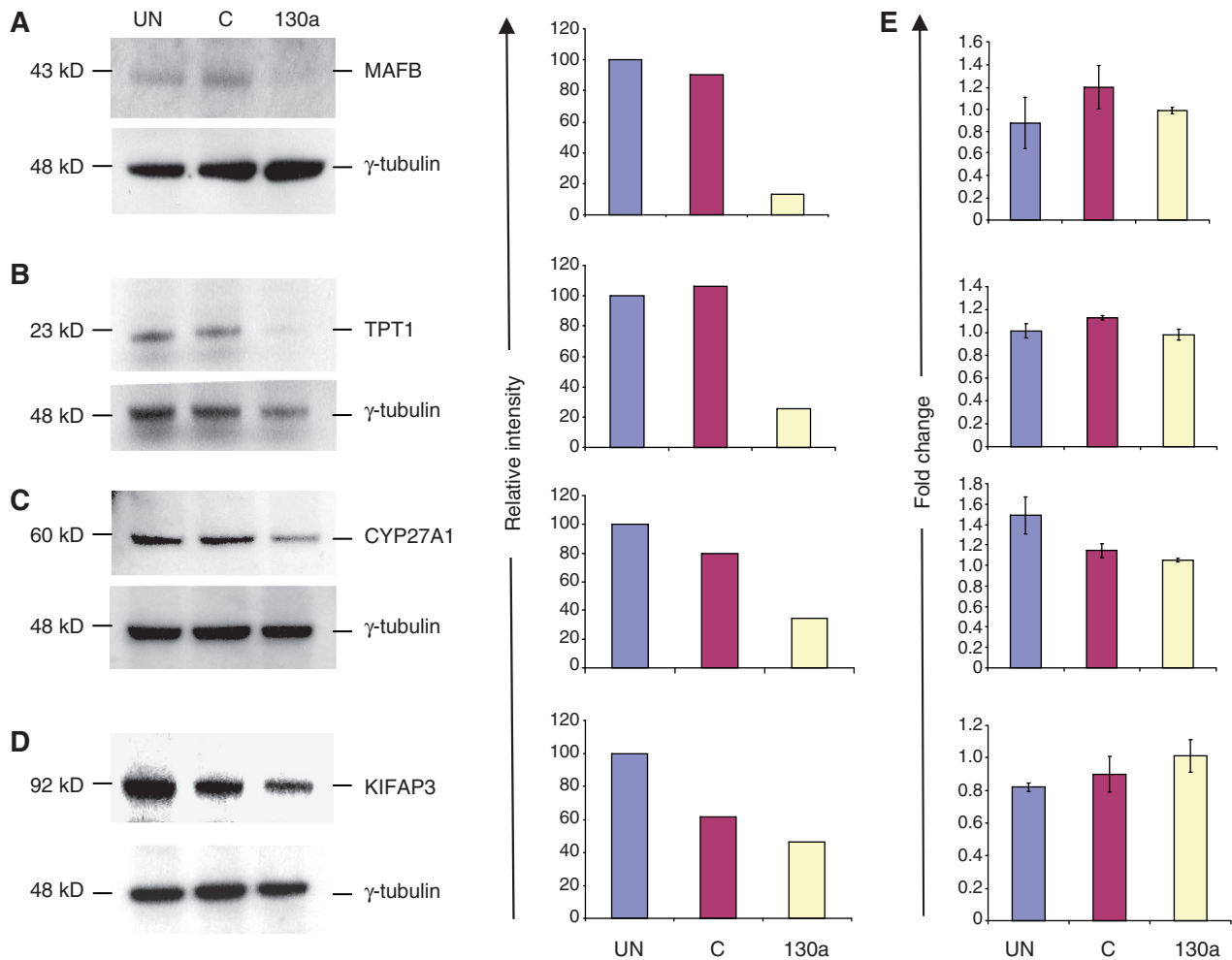
### Target validation

To validate the putative mir-130a targets, western blot analysis of endogenous protein expression was performed in MCF7 cells transfected with the empty vector pBabepuro or pBabepuro-mir-130a. Protein extracts obtained from MCF7 cells were probed with antibodies to TPT1 (Abcam), KIFAP3 (PTG Cambridge) and CYP27A1 (Abcam). MCF7 cells do not express endogenous MAFB, therefore, HepG2 cells (which do

express MAFB) were transfected with pBabepuro or pBabepuro-mir-130a and protein extract was used to probe for MAFB (Santa Cruz). Blots were stripped and re-probed with a  $\gamma$ -tubulin antibody (Santa Cruz). There were no commercial antibodies available for PRR14 and MITD1. The results showed clear down-regulation of MAFB, TPT1, KIFAP3 and CYP27A1 (Figure 4A–D). To determine whether these proteins were also regulated at the mRNA level, we analysed total RNA from the cell lines by QRT-PCR and found no difference in the levels of mRNA for MAFB and the novel mir-130a targets identified in this study, indicating that mir-130a modulates expression of these proteins at the translational level (Figure 4E).

Of the putative novel targets for mir-130a the highest level of down-regulation ( $\sim 80\%$ , Figure 4B) was observed for TPT1. TPT1 plays a role in many cellular processes such as cell cycle progression, apoptosis and carcinogenesis (14). *TPT1* expression is highly regulated both at the transcriptional and translational level and is ubiquitously expressed in most tissues, however TPT1 expression is low in brain (15), from which the 3'-UTR library was derived, indicating the robustness of this selection procedure.

To further investigate *TPT1* as a target for mir-130a, MCF7 cells expressing exogenous mir-130a were transfected with two MIRIDIAN hairpin inhibitors directed



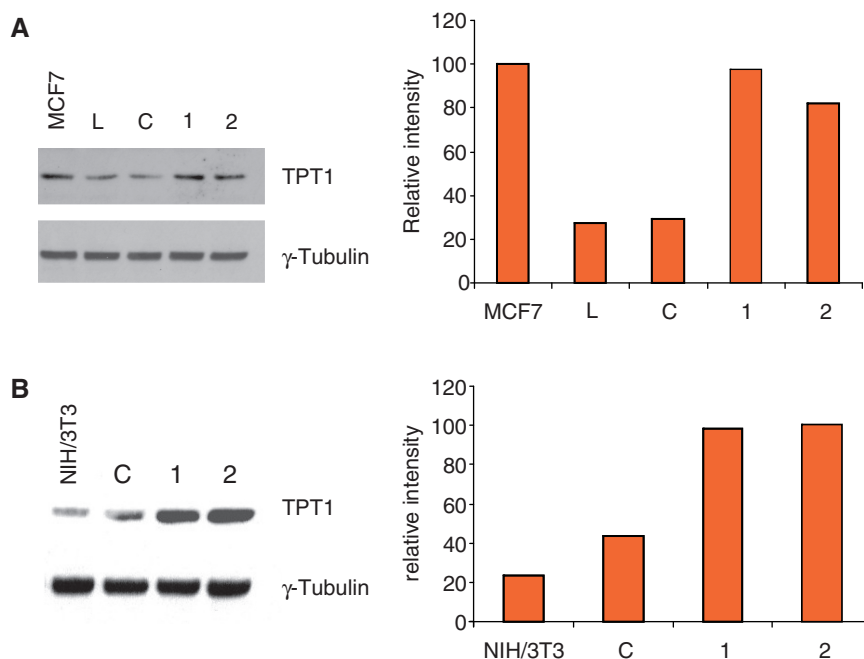
**Figure 4.** Western blot and QRT-PCR analysis of identified mir-130a targets. Endogenous protein expression levels were determined in untransfected, empty vector (pBabepuro) and mir-130a transfected cells (pBabepuro-mir-130a). Cells were transfected using the BES precipitation procedure. (A) HepG2 cell extract probed with MAFB antibody (Santa Cruz) showing down-regulation of MAFB in mir-130a expressing cells. (B–D) MCF7 cells probed with TPT1 antibody (Abcam), CYP27A1 antibody (Abcam) and KIFAP3 antibody (PTG Cambridge). Expression of the respective proteins was down-regulated in the presence of mir-130a. Blots were stripped and re-probed with a  $\gamma$ -tubulin antibody (Santa Cruz) as a loading control. Expression levels were quantified by densitometry as shown in the corresponding bar graphs. (E) Determination of mRNA levels by QRT-PCR. To investigate whether the down-regulation of mir-130a targets was due to mRNA degradation or inhibition of translation, the mRNA levels of MAFB, TPT1, CYP27A1 and KIFAP3 were quantified by QRT-PCR in triplicate on untransfected cells (UN), empty vector (pBabepuro) transfected cells (C) and pBabepuro-mir130a transfected cells (130a). The results show that mRNA levels remain unchanged indicating that the protein knockdown is due to inhibition of translation.

against mir-130a or a control hairpin inhibitor (Dharmacon). Cells were transfected using RNAiMAX transfection reagent (Invitrogen) and after 48 h analysed for protein expression. Western blot analysis demonstrated a 3- to 4-fold up-regulation of TPT1 expression in MCF7-mir-130a cells when treated with the hairpin inhibitors but not in cells treated with the control hairpin inhibitor or mock transfected cells (Figure 5A). Similarly NIH/3T3 cells, which express high levels of endogenous mir-130a (Supplementary Figure S2), were treated with the MIRIDIAN hairpin inhibitors. Western blot analysis 48 h after transfection showed a 3- to 4-fold up-regulation of TPT1 protein in these cells, validating *TPT1* as a target for mir-130a (Figure 5B). Western blot analysis of the same NIH/3T3 protein extracts showed that the inhibition of mir-130a also resulted in the up-regulation of two other

identified targets KIFAP3 and CYP27A1, although to a lesser extent, 2- and 0.4-fold, respectively (Supplementary Figure S7).

## DISCUSSION

To our knowledge, this study represents the first functional assay for the identification of miRNA target genes. This highly efficient assay does not merely rely on homology or binding to a putative target but on functional activity. Furthermore, this procedure will identify targets that are down-regulated either by mRNA degradation or inhibition of translation and also miRNA targets that are regulated by binding to ‘seedless’ recognition elements (10). Due to the nature of the assay, it will



**Figure 5.** Inhibition of mir-130a expression by MIRIDIAN hairpin inhibitors. (A) MCF7-mir-130a cells were either mock transfected (mir130aL) or transfected with MIRIDIAN hairpin inhibitors (Dharmacon). Cells were transfected with control hairpin inhibitor (mir130C), or two hairpin inhibitors (mir130a1, mir130a2) directed against mir-130a. After 48 h protein extracts were prepared and western blot analysis with anti-TPT1 was performed. Knockdown of mir-130a restores the TPT1 expression to ~80–90% of the control MCF7 expression. (B) NIH/3T3 cells, which express high levels of mir-130a, were treated the same as MCF7 cells above. Knockdown of endogenous mir-130a results in a 3- to 4-fold up-regulation of TPT1 protein.

only identify targets that are down-regulated to a large extent and are therefore likely to be amongst the most important targets of the miRNA under investigation. It is therefore likely that even in a large scale experiment the number of targets identified by this assay will be considerably smaller than predicted by computational algorithms, expression array or Ago2-IP procedures which identify hundreds of putative targets (8,16,17). Similarly the SILAC method results in the identification of hundreds of putative targets (18). This first SILAC based study for the identification of miRNA targets finds relatively mild repression of approximately 300 proteins after miRNA transfection in fairly small groups of 3000–3500 proteins. Furthermore, of the approximately 300 proteins identified only 60% contained a seed sequence for the miRNA under investigation (18).

This assay relies on the down-regulation of expression of a fusion construct by a miRNA resulting in the conversion of GCV resistance to GCV sensitivity of cells. Therefore, it is important to use a fusion construct conferring the highest sensitivity to GCV. We tested three different fusion constructs and found that TKzeo resulted in the highest sensitivity to low concentrations of GCV.

Furthermore, the miRNA of interest should not be expressed highly in the test cells. We examined a number of cell lines for expression of mir-130a and mir-10a, by QRT-PCR, and found that the breast epithelial carcinoma cell line MCF7, which grows as an adherent monolayer, showed one of the lowest levels of mir-130a and

mir-10a expression. Expression was also low in Jurkat and HL60 cells which are suspension cells. Adherent cells can be transfected with higher efficiency and therefore MCF7 cells were used to construct the target cell library. Expression profiling of MCF7 has shown that these cells only express 25% of the profiled miRNAs (19) and accordingly the established library of approximately 40 000 independent MCF7 clones expressing 5626 different cDNAs could be used to identify targets for many more miRNAs.

In addition to the validation of *MAFB* as a mir-130a target, this strategy identified a further five putative targets for mir-130a. Western blot analysis of endogenous protein levels for four of the identified targets demonstrated down-regulation of all these targets by mir-130a and similarly the knockdown of mir-130a in NIH/3T3 cells resulted in up-regulation of TPT1, KIFAP3 and CYP27A1.

At present, there are five experimentally validated mir-130a targets *TAC1* (20), *CSF1* (21), *MEOX2* (22), *HoxA5* (22) and *MAFB* (13). Of these validated targets we only isolated *MAFB* with this procedure. However next-generation sequencing showed that none of the additional validated mir-130a targets were present in the 5626 cDNAs comprising the starting library of approximately 40 000 clones.

A small-scale repeat experiment was performed to identify targets for mir-10a using the same library spiked with 10% *MAFB*. Sequencing 24 clones identified four targets (Table 1) for mir-10a one of which is *HoxA1* a

target validated by the same group who discovered MAFB as a mir-130a target (13). Importantly in this experiment we did not isolate MAFB as a false positive target for mir-10a although the library had been spiked with 10% MAFB.

We also investigated the stability of GCV sensitivity of the zeocin-resistant cells and found that a number of false positives appear over time in cultures that are not under selection pressure of zeocin. This is probably due to silencing of the CMV promoter which drives expression of the Tkzeo fusion gene resulting in GCV resistance due to the absence of HSV thymidinekinase expression. However, zeocin selection restores the GCV sensitivity and it is therefore important to select the cells with zeocin prior to miRNA introduction.

These findings show the fidelity and also the obvious limitations of this strategy; that is, it does identify miRNA targets that are present in the library but it cannot find targets that are absent from the target cell library. To overcome these limitations, a novel normalized cDNA library has been constructed in the p3'TKzeo vector (Sigma). This library contains cDNA derived from the 'Human Universal Reference Total RNA' (Clontech), which is a mixture of RNA derived from several human tissues, and cDNA derived from the 'Universal Human Reference RNA' (Agilent) which is a mixture of RNA from 10 human cell lines. Next-generation sequencing of this novel library showed the presence of 16 923 unique genes which is ~77% of the refseq sequences (Sigma).

Given the profound importance of the regulation of gene expression by miRNAs in both normal and malignant cells, the identification of their specific targets is an important objective for the functional analysis of the genome. The strategy described in this study should pave the way for the rapid identification and validation of large panels of miRNA targets. Combined with next-generation sequencing this strategy allows high-throughput, systematic identification of targets of specific miRNAs in a relatively short time.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–7, Supplementary Methods, Supplementary Data set.

## ACKNOWLEDGEMENTS

Thanks to Scott Hunicke-Smith and Vishwanath Iyer (University of Texas at Austin) for help with next-generation sequencing. The authors thank D. Darling and K. Ford (Department of Haematological Medicine King's College London) for helpful discussion. J.G. conceived the study, performed experiments, co-wrote manuscript. F.F., M.T. and G.J.M. conceived the study and co-wrote manuscript. A.M., J.J., F.M., D.S. and A.K. performed experiments and approved the manuscript. A.S., C.C. and N.S.B.T performed and

analysed Next-generation sequencing and approved the manuscript.

## FUNDING

Heptagon Fund London (J.G.); Leukaemia and Lymphoma (J.G., A.K., N.S.B.T., and C.C.); British Society for Haematology (J.G. and A.K.). Funding for open access charge: Departmental Funds.

*Conflict of interest statement.* Methods have been developed into a kit by Sigma. King College London, Division of Cancer Studies at King's and the authors Joop Gäken and Azim Mohamedali receive royalties dependent on sales.

## REFERENCES

- Ambros, V. (2004) The functions of animal microRNAs. *Nature*, **431**, 350–355.
- Ghildiyal, M. and Zamore, P.D. (2009) Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.*, **10**, 94–108.
- Filipowicz, W., Bhattacharyya, S.N. and Sonenberg, N. (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.*, **9**, 102–114.
- Carthew, R.W. and Sontheimer, E.J. (2009) Origins and mechanisms of miRNAs and siRNAs. *Cell*, **136**, 642–655.
- Bentwich, I. (2005) Prediction and validation of microRNAs and their targets. *FEBS Lett.*, **579**, 5904–5910.
- Didiano, D. and Hobert, O. (2006) Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat. Struct. Mol. Biol.*, **13**, 849–851.
- Easow, G., Teleanu, A.A. and Cohen, S.M. (2007) Isolation of microRNA targets by miRNP immunopurification. *RNA*, **13**, 1198–1204.
- Beitzinger, M., Peters, L., Zhu, J.Y., Kremmer, E. and Meister, G. (2007) Identification of human microRNA targets from isolated argonaute protein complexes. *RNA Biol.*, **4**, 76–84.
- Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D. and Slack, F.J. (2005) RAS is regulated by the let-7 microRNA family. *Cell*, **120**, 635–647.
- Lal, A., Navarro, F., Maher, C.A., Maliszewski, L.E., Yan, N., O'Day, E., Chowdhury, D., Dykxhoorn, D.M., Tsai, P. *et al.* (2009) miR-24 Inhibits cell proliferation by targeting E2F2, MYC, and other cell-cycle genes via binding to "seedless" 3'UTR microRNA recognition elements. *Mol. Cell.*, **35**, 610–625.
- Lu, L.F., Boldin, M.P., Chaudhry, A., Lin, L.L., Taganov, K.D., Hanada, T., Yoshimura, A., Baltimore, D. and Rudensky, A.Y. (2010) Function of miR-146a in controlling Treg cell-mediated regulation of Th1 responses. *Cell*, **142**, 914–929.
- Jongeneel, C.V., Delorenzi, M., Iseli, C., Zhou, D., Haudenschild, C.D., Khrebtkova, I., Kuznetsov, D., Stevenson, B.J., Strausberg, R.L. *et al.* (2005) An atlas of human gene expression from massively parallel signature sequencing (MPSS). *Genome Res.*, **15**, 1007–1014.
- Garzon, R., Pichiorri, F., Palumbo, T., Iuliano, R., Cimmino, A., Aqeilan, R., Volinia, S., Bhatt, D., Alder, H., Marcucci, G. *et al.* (2006) MicroRNA fingerprints during human megakaryocytopoiesis. *Proc. Natl Acad. Sci. USA*, **103**, 5078–5083.
- Bommer, U.A. and Thiele, B.J. (2004) The translationally controlled tumour protein (TCTP). *Int. J. Biochem. Cell. Biol.*, **36**, 379–385.
- Thiele, H., Berger, M., Skalweit, A. and Thiele, B.J. (2000) Expression of the gene and processed pseudogenes encoding the human and rabbit translationally controlled tumour protein (TCTP). *Eur. J. Biochem.*, **267**, 5473–5481.
- Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S. and Johnson, J.M. (2005)



- Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, **433**, 769–773.
17. Karginov, F.V., Conaco, C., Xuan, Z., Schmidt, B.H., Parker, J.S., Mandel, G. and Hannon, G.J. (2007) A biochemical approach to identifying microRNA targets. *Proc. Natl Acad. Sci. USA*, **104**, 19291–19296.
  18. Selbach, M., Schwanhäusser, B., Thierfelder, N., Fang, Z., Khanin, R. and Rajewsky, N. (2008) Widespread changes in protein synthesis induced by microRNAs. *Nature*, **455**, 58–63.
  19. Fix, L.N., Shah, M., Efferth, T., Farwell, M. and Zhang, B. (2010) MicroRNA expression profile of MCF-7 human breast cancer cells and the effect of green tea polyphenon-60. *Cancer Genomics Proteomics*, **7**, 261–277.
  20. Greco, S.J. and Rameshwar, P. (2007) MicroRNAs regulate synthesis of the neurotransmitter substance P in human mesenchymal stem cell-derived neuronal cells. *Proc. Natl Acad. Sci. USA*, **104**, 15484–15489.
  21. Sorrentino, A., Liu, C.G., Addario, A., Peschle, C., Scambia, G. and Ferlini, C. (2008) Role of microRNAs in drug-resistant ovarian cancer cells. *Gynecol Oncol.*, **111**, 478–486.
  22. Chen, Y. and Gorski, D.H. (2008) Regulation of angiogenesis through a microRNA (miR-130a) that downregulates antiangiogenic homeobox genes GAX and HOXA5. *Blood*, **111**, 1217–1226.