Published in final edited form as:

Nat Cell Biol. 2010 April ; 12(4): 372–379. doi:10.1038/ncb2037.

Genome-wide RNAi screen identifies miR-19 targets in Notch-induced acute T-cell leukaemia (T-ALL)

Konstantinos J. Mavrikis^{1,*}, Andrew L. Wolfe^{1,2,*}, Elisa Oricchio¹, Teresa Palomero^{3,4}, Kim de Keersmaecker⁴, Katherine McJunkin^{5,6}, Johannes Zuber⁵, Taneisha James^{5,**}, Kenneth Chang⁵, Aly A. Khan^{2,7}, Christina S. Leslie⁷, Joel S. Parker⁸, Patrick J. Paddison⁹, Wayne Tam¹⁰, Adolfo Ferrando^{3,4}, and Hans-Guido Wendel^{1,***}

¹ Cancer Biology & Genetics Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021

² Weill Cornell Graduate School of Medical Science, New York, NY 10065

³ Department of Pathology, Columbia University, NY 10032

⁴ Institute for Cancer Genetics, Columbia University, NY 10032

⁵ Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

⁶ Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

⁷ Computational Biology, Memorial Sloan-Kettering Cancer Center, New York, NY 10021

⁸ Expression Analysis Inc., Durham, NC 27502

⁹ Fred Hutchinson Cancer Research Center, Seattle, WA 98109

¹⁰ Department of Pathology and Laboratory Medicine, Joan and Sanford I. Weill Medical College of Cornell University, New York, NY 10025

MicroRNAs (miRNAs) have emerged as novel cancer genes. In particular, the 17~92 cluster of miRNAs is highly expressed in haematopoietic cancers and promotes lymphomagenesis *in vivo*^{1,2}. Clinical use of these findings hinges on isolating the oncogenic activity within the 17~92 cluster and defining its relevant target genes. Here we show that miR-19 is sufficient to promote leukaemogenesis in *Notch1* induced T-cell lymphoblastic leukaemia (T-ALL) *in vivo*. Consistent with the pathogenic importance of this interaction, we report a novel translocation targeting the 17~92 miRNA cluster coinciding with a second rearrangement that activates *Notch1* in T-ALL. To identify the miR-19 targets responsible for its oncogenic action, we conducted a large-scale short-hairpin RNA (shRNA) screen for genes whose knockdown could phenocopy miR-19. Strikingly, the results of this screen were enriched for miR-19 target genes, and include *Bim* (*Bcl2L11*)^{1,3}, AMP-activated kinase (*Prkaa1*), and the tumour suppressor phosphatases *Pten* and *PP2A* (*Ppp2r5e*). Hence, an unbiased, functional genomics

*** Correspondence and requests for materials should be addressed to: Hans-Guido Wendel, Cancer Biology & Genetics Program, Sloan-Kettering Institute, 1275 York Ave, New York, NY 10021, Phone: 646-888.2526, Fax: 646-422.0197, Wendelh@mskcc.org.

* KJM and ALW have contributed equally.

** Current address: Department of Gene and Cell Medicine & Black Family Stem Cell Institute, Mount Sinai School of Medicine, New York, NY 10029

Author contributions K.J.M., A.L.W. contributed equally; K.J.M., A.L.W., E.O. experimental design and analysis; K.d.K., T.P., A.F. T-ALL translocation analysis; K.M., J.Z., T.J., K.C., screen and analysis; A.A.K., C.S.L., J.S.P. data analysis; P.J.P. generated shRNA library; W.T. clinical specimens; H.G.W. designed the study and wrote the paper.

Author information The authors declare no competing financial interests.

approach reveals a coordinate clamp down on several regulators of PI3K-related survival signals by the leukaemogenic miR-19.

MiRNAs are ubiquitous regulators of biological processes involved in normal development, differentiation, and diseases including cancer. They act by regulating gene expression at transcriptional and translational levels⁴. Several methods have been devised to identify genes that are regulated by miRNAs, e.g. programs like Targetscan can identify potential miRNA targets based on conserved sequence complementarity^{5,6,7}. Experimental tools can define the actual effects of miRNAs on gene expression, e.g. expression profiles and comparative proteomics (SILAC)^{8,9}, or measurements of incorporation into the miRNA machinery (HITS-CLIP)¹⁰. However, most likely only a limited number of miRNA targets is directly responsible for a specific phenotype¹¹. An unbiased way of identifying these functionally *relevant* miRNA targets is needed to understand the biological basis of miRNA activity.

The oncogenic 17~92 cluster of miRNAs is of eminent importance in human hematopoietic cancers^{1,2,12}. It has two paralogues and together these clusters encode 15 miRNAs with overlapping functions in development³. While miR-19 is important in Burkitt's lymphoma^{13, 14}, which particular miRNA(s) deliver the oncogenic activity of these clusters is not established. Moreover, while comprehensive lists of potential targets of the 17~92 cluster are available, and some individual candidates have been confirmed experimentally^{3,15,16}, the mechanism responsible for its oncogenicity remains unknown.

To identify the oncogenic activity within the 17~92 cluster and its paralogues we used an *in vitro* assay of hematopoietic transformation¹⁷. We tested all miRNAs representing families of seed sequences, specifically miR-17, miR-18a, miR-19b-1 (miR-19), miR-20a, miR-106a, miR-106b and miR-25 (Figure 1a)³. The assay is based on the interleukin-3 (IL3) dependence of FL5-12 lymphocytes, which undergo apoptosis when removed from IL3. We partially transduced FL5-12 cells with the individual miRNAs and GFP or empty vector (MIG), and monitored these mixed populations by FACS for changes in their relative proportions (Figure 1b). FL5-12 cells expressing miR-19 were rapidly enriched over parental cells upon IL3 depletion, and none of the other miRNAs had a similar protective effect (Figure 1c) (miR-19: $p < 0.002$, all other miRNAs and vector $p > 0.05$ by t-test). Accordingly, miR-19 showed direct protection from cell death in FL5-12 cells and had little effect on proliferation (Suppl. Figure 1). Hence, within the 17~92 cluster and its paralogues, miR-19 has a distinct ability to enhance lymphocyte survival *in vitro*.

Next, we assessed the expression of miR-19 in human lymphatic malignancies. Compared to lymphocytes from tonsils, we found a 5-17 fold increase in miR-19 expression in T-ALL, and less for other miRNAs in the 17~92 cluster (Figure 2a, Suppl. Fig. 2 and Suppl. Table 1). MiR-19 was also highly expressed in B-ALL and aggressive lymphomas. Consistent with previous profiling studies^{18,19}, we found lower levels in indolent follicular lymphomas. Retroviral expression of miR-19 in the FL5-12 cells produced levels comparable to some tumour specimens (red bar in Figure 2a).

Karyotype analysis of an adult patient with T-ALL revealed the co-existence of two clonal translocations involving both TCRA/D loci in chromosome band 14q11. The first rearrangement, t(9;14)(q34;q11) affects the *NOTCH1* gene locus in chromosome band 9q34, and 5'RACE confirmed the expression of a constitutively active form of *NOTCH1* (exons 29–34) (Suppl. Figure 3)^{20,21}. The second rearrangement - t(13;14)(q32;q11) - also involving the TCRA/D locus is novel and pinpoints a new TALL oncogene at 13q32. FISH analysis using genomic clones RP11-980D6 and RP11-97P7 reveals a split signal between chromosome 13 and the derivative chromosome 14 and maps the breakpoint close to the 17~92 cluster t(13;14)

(q32;q11) (Figure 2b, c and Suppl. Fig. 4). Thus, the 17~92 cluster is the target of a novel genomic rearrangement in T-ALL, where miR-19 is highly expressed.

Most cases of T-ALL show mutations or translocations that result in increased expression or constitutive activity of a truncated *Notch1* gene (Notch-ICN)^{20–22}. To determine whether miR-19 could enhance *Notch1*-induced T-ALL *in vivo*, we co-transduced hematopoietic progenitor cells (HPCs) from fetal livers with Notch-ICN and miR-19 or empty vector²³. We then transplanted the HPCs into irradiated recipients and monitored them for leukaemia (Figure 2d). Mice receiving HPCs expressing both Notch-ICN and miR-19 succumbed to ALL in ~ 2 months. At the same time, 80% of the mice receiving Notch-ICN remained healthy ($p=0.0003$) (Figure 2e). Pathology confirmed the diagnosis of T-ALL with abundant lymphoblasts, marrow infiltration, splenomegaly, lymphomas, and expression of T-cell markers (Figures 2f–h, Suppl. Figure 5 and Suppl. Table 2). T-ALLs expressing miR-19 or vector looked identical (not shown). We saw similar activity of miR-19 in the *EμMyc* model of Burkitt's lymphoma (Suppl. Figure 6)^{13,14}. Screening a pool of miRNAs containing all members of the 17~92 cluster and its paralogues for tumorigenesis in this model, revealed enrichment of cells transduced with miR-19 (Suppl. Figure 7). Hence, miR-19 behaves as an oncogene in murine models of lymphoma and *Notch1*-induced in T-cell lymphoblastic leukaemia (T-ALL).

We now wanted to determine which molecular targets were responsible for miR-19's oncogenicity. Initially we combined computational target identification and gene expression analyses. Targetscan predicts 938 human and 744 murine miR-19 targets (Suppl. Tables 3 and 4). Gene expression profiles in FL5-12 cells transduced with miR-19 or vector showed only modest changes in expression (mean fold change -0.2 ± 0.39) (Figure 3). Overall, expression levels of all predicted miR-19 targets were more reduced than other genes ($p < 2e-04$, Kolmogorov-Smirnov test). Among these were also potentially interesting candidates like the *Pten* tumour suppressor. However, there was no significant enrichment of miR-19 targets among genes showing more pronounced (1.5 or 2 SD) reductions in expression ($p > 0.46$ at 2 SD; $p > 0.077$ at 1.5 SD, Fisher's exact test) (Figure 3 and Suppl. Table 5). Thus, expression analysis and target prediction readily confirm global effects of miR-19 and may lead to promising candidates, but they do not provide strong enough filters to identify the most relevant miR-19 targets.

Therefore, we wondered if an unbiased genetic screen might be an alternative means to functionally identify miR-19 targets involved in lymphocyte survival. Specifically, a short-hairpin RNA (shRNA) screen^{24,25} for genes whose knockdown phenocopies miR-19 in lymphocytes should also identify the genes responsible for miR-19 action. We conducted extensive control experiments using an shRNA against *Bim* (*Bcl2L11*), a known target of the 17~92 cluster^{1,2}, and also increasing dilutions of miR-19 to establish the exact behaviour of miR-19 and a known target in our screening assay (Suppl. Figures 8a–d). We then devised a screening protocol involving transduction of FL5-12 cells with pools of ~1,000 shRNAs, and two cycles of IL3 depletion or continued passage in complete media (Figure 4a). We used custom half-hairpin arrays to measure changes in the abundance of shRNAs in treated (–IL3) versus untreated (+IL3) samples at two time points (T1 and T2, equal one or two cycles of IL3 depletion, respectively)²⁶. Unsupervised clustering showed good reproducibility between biological replicates (A–C; mean correlation $r=0.60 \pm 0.17$), and revealed a progressive shift in shRNAs across subsequent cycles of IL3 depletion (Figure 4b). Statistical tools helped us identify biologically significant signals. A Significance Analysis of Microarrays (SAM) identified changes in individual shRNAs (Figure 4c, see Suppl. Table 6 for the full analysis). A Gene Set Analysis (GSA) defined groups of shRNAs targeting the same genes. The GSA identified 14 genes, each targeted by at least 2 and up to 5 different shRNAs, and the top 'hit' was the alpha subunit of AMP-activated kinase (*Prkaa1*) targeted by 5 independent shRNAs (Figure 4d and Suppl. Table 7 for the full analysis).

Next, we individually validated candidate genes in the same experimental system. We included nearly all genes identified from the GSA analysis, and also tested protein-coding genes from the SAM analysis above an arbitrary threshold (≥ 1.65 fold increase, $p < 0.05$). In total, we re-tested > 70 genes and typically 3 shRNAs against each (all genes included in the validation are highlighted in Suppl. Tables 6 and 7). Ultimately, shRNAs against eight genes positively validated and produced a survival benefit in FL5-12 cells depleted from IL3 (in triplicate experiments $p(\text{vector/not shown}) > 0.2$ and $p(\text{for all shRNAs}) < 0.05$; t-test) implying a positive validation rate of $\sim 10\%$ (Figure 4e). Strikingly, in the human genome five of these genes harboured the miR-19 seed sequences. Besides the pro-apoptotic *Bim*¹, these were the tumour suppressor gene *Pten*¹, the alpha subunit of AMP-activated kinase (*Prkaa1*), the epsilon isoform of PP2A (*Ppp2r5e*), and the dedicator of cytokinesis-5 gene (*Dock5*). In addition, we validated three other genes that did not contain miR-19 binding sites in their 3'UTRs, the FoxO transcription factors *FoxO1* and *FoxO3*, and *Bnip3*, a regulator of Rheb/mTOR and Bcl2 binding protein. Taken together, the results of this unbiased genetic screen for shRNAs that behave like miR-19 showed a highly significant enrichment for genes that contain miR-19 binding sites ($p < 7.2 \times 10^{-7}$ in Fisher's exact test; Suppl. Figure 9a). Unlike the human *DOCK5* gene, the mouse gene is not a miR-19 target, and calculation of the enrichment statistics for the murine genome confirmed similarly high significance levels ($p < 3.2 \times 10^{-5}$ by Fisher's exact test; Suppl. Figure 9b). The prediction of miR-19 sites is based on Targetscan and did not take into account potential non-classical miRNA binding sites, e.g. those in the coding sequence (CDS). However, some of the eight genes also harbour potential CDS sites and these are shown in Suppl. Table 13. Thus, our unbiased, large-scale shRNA screen identified five human and four murine genes whose knockdown phenocopies miR-19 action in lymphocyte survival and that harbour miR-19 binding sites in their 3'UTRs.

Next, we wanted to test whether miR-19 actually regulates the expression of these genes. Reporter assays confirmed direct 3'UTR inhibition mediated by the predicted binding sites for these genes (Suppl. Figure 10 and Suppl. Table 9 and 12). Using qRT-PCR we measured the effect of miR-19 on mRNA levels of *Pten*, *Ppp2r5e*, *Prkaa1* and *Bim* in miR-19 expressing murine T-ALL cells (Suppl. Figure 11) and in FL5-12 cells (Figure 5a), where we consistently observed up to 2 fold reductions ($p(\text{vector vs. miR-19}) < 0.05$ for all four genes, t-test) (Suppl. Table 8). Conversely, an antagomir against miR-19 caused an increase in mRNA levels *in vitro* ($p(\text{vector vs. anti-19}) < 0.05$ for all four genes) (Figure 5b, Suppl. Table 8), and both miR-19 and the antagomir had measurable effects on protein levels. For example, miR-19 produced a clear reduction in PPP2R5E, PRKAA1, BIM, and resulted in overall activation of PI3K signalling as indicated by increased phosphorylation of AKT and the ribosomal S6 protein (Figure 5c, Suppl. Figure 12). While we did not detect an effect of miR-19 on PTEN protein, its mRNA was clearly decreased, and the antagomir produced increases in all miR-19 target proteins including PTEN (Figure 5b and c). As expected neither miR-19, nor the antagomir had an effect on *Bnip3*, which does not contain a 3'UTR site for miR-19 (Figure 5c). The larger effects of the antagomir compared to miR-19 expression may indicate that these genes are tonically suppressed by the endogenous miR-19 in proliferating cells. Hence, we confirm that the expression of *Bim*, *Pten*¹, *Prkaa1*, and *Ppp2r5e* is regulated by miR-19 in lymphocytes.

Next, we wanted to assess the contribution of these target genes to miR-19 action. The pro-apoptotic BIM protein opposes BCL2 and is an important regulator of lymphocyte survival²⁷. To measure if miR-19 had Bim-independent effects, we conducted complementation studies with FL5-12 cells engineered to stably express *Bcl2* (FL5-12/*Bcl2* cells). As expected the *Bim* shRNA conferred no additional benefit, whereas miR-19, and the shRNAs against *Pten* and *Pp2r5e* showed continued enrichment in the presence of *Bcl2* (Figure 5d). Conversely, an antagomir against miR-19 produced anti-proliferative effects and FL5-12/*Bcl2* cells expressing the antagomir were lost upon continued passage (Figure 5e). Similarly, miR-19 maintained its protective effect in the presence of Metformin-activated AMP kinase

or enforced expression of Pten or Pp2r5e (Suppl. Figure 13). This indicates that multiple target genes contribute to miR-19 action *in vitro*. *In vivo* we find that individual knockdown of some targets is sufficient to accelerate leukaemogenesis, this is most striking for Pten (n=4, vector n=10, p < 0.001) and Bim (n=6; p < 0.001), an shRNA against Prkaa1 produces a modest effect (n = 5; p=0.3), while knockdown of Pp2r5e alone is unable to drive leukaemogenesis (n=4) (Figure 5f). Together, our data identify multiple miR-19 targets, which together act on PI3K-related signals and affect lymphocyte survival *in vitro* and leukaemogenesis *in vivo* (Figure 5g).

Our results provide new insights into the oncogenic activity of the miR-17~92 cluster in T-cell leukaemia. MiR-19, expressed at levels seen in human tumours, enhances lymphocyte survival and is sufficient to cooperate with *Notch1* in T-cell lymphoblastic leukaemia (T-ALL) *in vivo*. Similarly, miR-19 can enhance *c-Myc* induced B-cell lymphomagenesis (Suppl. Figure 6)^{13,14}. Given the conservation of the entire 17~92 cluster it is likely that these miRNAs carry out important functions in other tissues or biological processes¹². The interaction of *Notch1* and miR-19 is a relevant pathogenic mechanism in T-ALL. This is most strikingly demonstrated by the co-existence in the same leukaemic clone of two translocations activating the *NOTCH1* gene - t(9;14)(q34;q11) - and targeting the 17~92 cluster - t(13;14)(q32;q11), respectively. It is likely that other mechanisms, for example transcriptional activation by *c-Myc* and *Notch1* are more common causes of the increased miR-19 expression we observed in T-ALL specimens. Together, miR-19 is an oncogene that contributes to the molecular pathogenesis of T-ALL.

We identify a pattern of gene regulation by miR-19. Using an unbiased shRNA screening approach we find eight genes whose knockdown recapitulates miR-19 effects *in vitro*. Intriguingly, in the human genome five of these 'hits' harbour miR-19 binding sites and are regulated by miR-19. It is likely that these genes are among the key effectors of miR-19 action in lymphocyte survival, and that the regulation of multiple genes contributes to the overall activity of miR-19. These include *Bim*, a pro-apoptotic *Bcl2* protein and target of the 17~92 cluster³, as well as AMP-activated kinase (*Prkaa1*), the tumour suppressor phosphatases Pp2a (subunit *Ppp2r5e*) and *Pten*¹¹ (Figure 5g). Other miRNAs encoded in the 17~92 cluster are predicted to target some of the same genes, suggesting a potential mechanism for the functional cooperation between these clustered miRNAs¹¹. These results suggest that miR-19 coordinates a PI3K pathway related program of cell survival in lymphocytes that contributes to leukaemogenesis.

Our results validate a new functional genomics approach to identify miRNA target genes. A genetic screen provides an unbiased tool to pinpoint miRNA targets genes that may be individually responsible for a specific phenotype, such as lymphocyte survival in this study. Some limitations should be noted, for example the screen is based on shRNAs targeting individual genes, as a consequence we may miss weak miRNA targets whose activity will only be apparent in the context of others. Also shRNAs and miRNAs act through somewhat distinct mechanisms^{6,28}, and current technology cannot guarantee efficient protein knockdown by all shRNAs in a large library and this may. Short-hairpin RNA screens are an evolving technology, and limitations of array detection and shRNA validation likely account for the overall low (10%) validation rate. Despite the apparent complexity of gene regulation by miRNAs, our data indicate that phenotype-based genetic screens can identify key targets of miRNA action and thus complement other computational^{7,29} and genomic/proteomic^{8,10} methods of target identification. Notably, the cross-comparison of these methods will enable new insights into the biological activity of important miRNAs.

Methods

Cell culture, viability, proliferation assays and vectors constructs

FL5-12 murine lymphocytes, cell cycle, apoptosis studies, viral transductions were as described^{30,17}. All vectors are based on MSCV, and include the miRNA expression vector encoding miRs17, 18a, 19b-1 (miR-19), 106a, 106b, 25², Notch-ICN (a gift from W. Pear)²³, *Bcl2*³¹, the individual shRNA vectors (protein knockdown by shRNAs Suppl. Figure 14, vector constructs Suppl. Figure 15, Suppl. Table 10). The miRNA pool contained 48 miRNAs (all 17~92 and additional control miRNAs; the complete list is available upon request). The antagomirs MZIP19a-PA-1, MZIP19b-PA-1, and scrambled control (MZIP000-PA-1) were from System Biosciences(Suppl. Figure 16).

Generation of mice

The mouse T-ALL and EμMyc lymphoma models have been reported^{23,31}. Data were analyzed in Kaplan-Meier format using the log-rank (Mantel-Cox) test for statistical significance. The *p53* LOH PCR and surface marker analysis were as described³¹.

Western blot analysis

Immunoblots were performed from whole cell lysates as described³¹. Antibodies were against Prkaa1 (2532, 1:1000, Cell Signaling), Bim/Bcl2L11 (AAP-330, 1:1000, Assay Designs), FoxO1 (94545, 1:1000, Cell Signaling), FoxO3a (9467, 1:1000, Cell Signaling) Phospho-FoxO3a (94665, 1:1000, Cell Signaling), Ppp2r5e (NB 100-845, NOVUS), Pten (9559, 1:1000, Cell Signaling), Tubulin (1:5000; Sigma, B-5-1-2) and Actin (1:5000; Sigma, AC-15), Bnip3 (3769, 1:1000, Cell Signaling), phosphorylated S6 (2215, 1:1000, Cell Signaling), phosphorylated Akt (4058, 1:1000, Cell Signaling), the Dock5 antibody was a gift from Alan Hall (MSKCC).

Real time quantitative PCR

Total RNA and miRNA enriched RNA was extracted using the Allprep DNA/RNA/Protein and miRNeasy Mini. Pathological diagnosis was by expert haematopathologists at Weill Cornell University. cDNA synthesis and qRT-PCR and analysis by the $\Delta\Delta$ Ct method as described³⁰. Taqman Gene Expression Assays: Bcl2l11 (Mm00437796_m1), Dock5 (Mm00555757_m1), FoxO1 (Mm00490672_m1), FoxO3 (Mm01185722_m1), Ppp2r5e (Mm00803759_m1), Prkaa1 (Mm01296695_m1), Pten (Mm01212532_m1), and Mouse GAPD (GAPDH) (4352932, Applied Biosystems), miR-19b (000396, Applied Biosystems) expression was normalized to RNU6B (001093, Applied Biosystems); primer sequences for miRNA detection are available upon request.

Luciferase assays

The Dock5, Ppp2r5e, Prkaa1, and Bim 3'-UTR fragments (Suppl. Table 11) were generated by PCR and cloned into the psi-CHECK-2 vector (Promega). The assays were performed as described¹. The binding site mutants were generated by site-directed mutagenesis (Suppl. Table 12).

Karyotype and Fluorescence in situ hybridization (FISH) analysis

Metaphase chromosome preparations made from primary patient lymphoblasts were subjected to karyotype analysis following standard procedures. Genomic clones RP11-97P7 and RP11-980D6 located at 13q32 were from Invitrogen. DNA was labelled by nick-translation using spectrum orange dUTP fluorochrome (Vysis, Downers Grove, IL). A spectrum green-labelled RB1 probe was used as control for chromosome 13q hybridization. FISH was performed by standard methods on cells used for cytogenetic analysis. Hybridization signals

were scored on at least 20 metaphase spreads on DAPI-stained slides using the Cytovision Imaging system attached to a Nikon Eclipse 600 microscope (Applied Imaging, Santa Clara, CA).

5' RACE Amplification of aberrant *NOTCH1* transcripts

mRNA was extracted from primary patient lymphoblasts using the Nucleotrap mRNA extraction kit from Clontech, and 5'RACE was performed with the SMART RACE kit (Clontech) using a oligonucleotide primer complementary to the sequence of exon 29 of the *NOTCH1* gene (5'-TCGTCCATGAGGGCACCGTCTGAAG -3').

Pooled shRNA library screen

The pooled shRNA screen and half-hairpin array detection has been described, the library has been cloned from the original PSM2 constructs into the MRP vector and it contains ~ 14,000 distinct shRNAs, pooled into 1,000 shRNAs per pool^{25,26,30}. FL5-12 cells were transduced at low multiplicity of infection in triplicates and each subjected to 2 cycles of IL3 depletion and rescue (T1 and T2). Samples were collected after viability had recovered for DNA isolation, PCR amplification of integrated shRNAs or the library as a reference, labelled and hybridized. Data generated (Nimblescan software) from image scans (Axon 4000BScanner) were imported into R version 2.4 for processing and analysis. Each Nimblegen 12-plex custom array consists of 12,033 half-hairpin probes (sequences are available on request). Most were represented by shRNAs in the library, and the remainder used for background estimation. Negative control spots showed lower intensities than the experimental probes on each array (Suppl. Figure 15c). Signal values less than the background were replaced with the background estimate to dampen large ratios due to low signal. The two channels for each array were then normalized using loess normalization and log ratios of the normalized intensity were utilized in further analyses.

Data analysis for pooled shRNA screens

The analyses were performed using the Bioconductor linear models for microarray (limma)³², Significance Analysis of Microarrays (SAM)³³ and Gene Set Analysis (GSA)³⁴ libraries in R. Correlations between biological replicates were calculated for each time point, principal component analyses (PCA) to determine the magnitude of experimental effects and biological replicates. Biological replicates did not constitute a major source of variation in the experiment. The mean correlation for biological replicates was 0.60 (+/-0.17) and likely reflected random variations as populations drifted in culture. Biologically significant signals were identified using the SAM software to select probes reflecting differences in relative abundance of shRNAs. We used the unpaired, two class algorithm, which assumes independent samples of each feature across the conditions. Q values are calculated empirically and used for identifying significant features. Overall, a gene was scored as being a potential candidate if its shRNA probe was found to fulfil these criteria: ≥ 1.65 fold change (FC), $p < 0.05$, and current gene bank annotation with a corresponding protein. Further testing was performed using GSA to identify candidates based on data from multiple hairpins targeting the same gene, as each gene is targeted by 2–4 hairpins their fold changes are summed. This statistic was compared to the same statistic from exhaustive permutation to arrive at empirical p-values for each gene. All gene sets showing positive fold-change in the GSA analysis were included in validations (except olfactory receptor481). These cut-offs are relatively arbitrary and were chosen as a conservative criterion to minimize false positives.

Computational analysis of shRNA OFF-target effects

We performed several computational analyses to investigate whether the results of the screen might be attributable to off-target effects. We considered two hypotheses: (1) that shRNAs overrepresented in the screen have sequence similarity to miR-19 itself and hence mimic

miR-19 activity via sites in 3' UTRs; (2) that results of the screen are explainable by microRNA-like off-target effects on a few key genes. We ranked all shRNAs by ungapped sequence similarity of (a) the shRNA "seed" (positions 2–8) to the miR-19 seed region and (b) the shRNA to the full-length miR-19a and miR-19b sequences. No shRNAs in the library had the same 7-mer seed region as miR-19; among the shRNAs whose seed region matched the miR-19 7-mer seed at 6 positions (12 shRNAs with 6-mer seeds matching positions 3–8, 2 shRNAs with 6-mer seeds matching positions 2–7), none were in the top 100 shRNAs overrepresented in the screen. The maximal sequence similarity between shRNAs in the library and the full-length miR-19a/b sequences was a 13 base ungapped match; 23 shRNAs had this degree of similarity to miR-19a and/or miR-19b, but none of them occurred in the top 100 overrepresented shRNAs from the screen. Moreover, we found poor Spearman rank correlation ($|\rho| < .01$) between the ranking of shRNAs by overrepresentation in the screen and these similarity comparisons to miR-19. Next, we considered whether shRNAs that were highly ranked in the screen were more likely to have predicted off-target effects on PTEN, PRKAA1, BCL2L11, PPP2R5E, DOCK5, FOXO1, FOXO3 and BNIP3, compared to the full library. If this were true, overrepresentation of an shRNA in the screen might be explainable through off-target silencing of these key genes. To predict potential microRNA-like off-target effects of shRNAs, we scanned 3' UTRs in the entire mouse genome for matches to 7-mer seed sequences (positions 2–8, no conservation filter) of shRNAs in the library. A gene with a 7-mer seed match for a given shRNA was considered to be a predicted off-target for this shRNA. We then tested if, for any of the 8 genes, the set of shRNAs that were predicted to "off-target" the gene were enriched in the top K overrepresented shRNAs from the screen, as compared to the full shRNA library. We used values of K ranging from 2 to 250, and we did not observe statistical significance by a Fisher's exact test for any value of K ($p > 0.05$ in all cases for each gene). These controls indicate that the shRNAs are not similar to miR-19 itself, and that the screen result could not be explained by off-target effects of top ranked shRNAs on key genes. We computed the enrichment statistic with respect to the number of unique targets of the ~12000 shRNAs in the library and the subset of these targets that were also predicted miR-19 targets using Fisher's exact test to account for the small number of 'hits'. This analysis was done for both the human and the murine genome. Target predictions were as per Targetscan 5.1.

Expression array analysis

RNA isolation, quality control, cRNA synthesis, labelling hybridization, scanning according to standard protocols and manufacturer (Affymetrix) and analyzed using Partek® Genomic Suite™ 6.4. For statistical analysis, we centred the data using the mean log2 (expression change) of genes and normalized to have unit variance in log2 (expression change) across all genes. This normalization results in a modified Z-transformation of the data (Z score).

KS statistic

To compare the expression changes for miR-19 targets versus all genes, we compared their distributions of log(expression change) values using a one-sided Kolmogorov-Smirnov (KS) statistic, which assesses whether the distribution of expression changes for one set is significantly shifted downwards (down-regulated) compared to the distribution for the other set. The KS statistic computes the maximum difference in value of the empirical cumulative distribution functions (cdfs):

$$\sup_x (F_1(x) - F_2(x)),$$

where $F_j(x) = \frac{1}{n_j} \sum_{i=1}^{n_j} I_{x_i \leq x}$ is the empirical cdf for gene set $j = 1, 2$, based on n_j (Z-transformed) log(expression change) values. We used the Matlab function `kstest2` to calculate the KS test statistic and asymptotic p-value.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Lin He, S.W. Lowe, and G. Hannon for access to shRNA and miRNA screening technologies. W. Pear, J. Huse, E.C. Holland for plasmids. H. Zhu for technical assistance, the MSK animal facility and RARC, A. Viale of the MSK Genomics Core, H. Zhao of cBIO program, J. Schatz and J. Massague for editorial advice, V. Murty for cytogenetic and FISH analyses. This work is supported by grants to HGW from the American Cancer Society, the Geoffrey Beene Cancer Center, Leukemia Research Foundation, The Louis V. Gerstner Foundation, May and Samuel Rudin Foundation, the NY Community Trust, the William and Blanche Foundation, Andrew Seligson Memorial Clinical Fellowship (JZ), NYStar (PJP), AF is supported by R01CA120196, the WOLF Foundation, the Rally across America Foundation, the Leukemia and Lymphoma Society (grants 1287-08, 6237-08), and AF is a Leukemia & Lymphoma Society Scholar.

References

1. Xiao C, et al. Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes. *Nat Immunol* 2008;9:405–414. [PubMed: 18327259]
2. He L, et al. A microRNA polycistron as a potential human oncogene. *Nature* 2005;435:828–833. [PubMed: 15944707]
3. Ventura A, et al. Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. *Cell* 2008;132:875–886. [PubMed: 18329372]
4. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004;116:281–297. [PubMed: 14744438]
5. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009;19:92–105. [PubMed: 18955434]
6. Grimson A, et al. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* 2007;27:91–105. [PubMed: 17612493]
7. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell* 2003;115:787–798. [PubMed: 14697198]
8. Baek D, et al. The impact of microRNAs on protein output. *Nature* 2008;455:64–71. [PubMed: 18668037]
9. Selbach M, et al. Widespread changes in protein synthesis induced by microRNAs. *Nature* 2008;455:58–63. [PubMed: 18668040]
10. Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 2009;460:479–486. [PubMed: 19536157]
11. Xiao C, Rajewsky K. MicroRNA control in the immune system: basic principles. *Cell* 2009;136:26–36. [PubMed: 19135886]
12. Mendell JT. miRiad roles for the miR-17-92 cluster in development and disease. *Cell* 2008;133:217–222. [PubMed: 18423194]
13. Mu P, et al. Genetic dissection of the miR-17~92 cluster of microRNAs in Myc-induced B-cell lymphomas. *Genes Dev* 2009;23:2806–2811. [PubMed: 20008931]
14. Olive V, et al. miR-19 is a key oncogenic component of mir-17-92. *Genes Dev* 2009;23:2839–2849. [PubMed: 20008935]
15. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* 2005;435:839–843. [PubMed: 15944709]

16. Petrocca F, et al. E2F1 -regulated microRNAs impair TGFbeta-dependent cell-cycle arrest and apoptosis in gastric cancer. *Cancer Cell* 2008;13:272–286. [PubMed: 18328430]
17. Plas DR, Talapatra S, Edinger AL, Rathmell JC, Thompson CB. Akt and Bcl-xL promote growth factor-independent survival through distinct effects on mitochondrial physiology. *J Biol Chem* 2001;276:12041–12048. [PubMed: 11278698]
18. Malumbres R, et al. Differentiation stage -specific expression of microRNAs in B lymphocytes and diffuse large B-cell lymphomas. *Blood* 2009;113:3754–3764. [PubMed: 19047678]
19. Calin GA, Croce CM. Investigation of microRNA alterations in leukemias and lymphomas. *Methods Enzymol* 2007;427:193–213. [PubMed: 17720486]
20. Palomero T, et al. Activating mutations in NOTCH1 in acute myeloid leukemia and lineage switch leukemias. *Leukemia* 2006;20:1963–1966. [PubMed: 17008890]
21. Ellisen LW, et al. TAN -1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell* 1991;66:649–661. [PubMed: 1831692]
22. Weng AP, et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* 2004;306:269–271. [PubMed: 15472075]
23. Pear WS, et al. Exclusive development of T cell neoplasms in mice transplanted with bone marrow expressing activated Notch alleles. *J Exp Med* 1996;183:2283–2291. [PubMed: 8642337]
24. Paddison PJ, et al. A resource for large -scale RNA-interference-based screens in mammals. *Nature* 2004;428:427–431. [PubMed: 15042091]
25. Silva JM, et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science* 2008;319:617–620. [PubMed: 18239125]
26. Chang K, Elledge SJ, Hannon GJ. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods* 2006;3:707–714. [PubMed: 16929316]
27. Bouillet P, et al. Proapoptotic Bcl -2 relative Bim required for certain apoptotic responses, leukocyte homeostasis, and to preclude autoimmunity. *Science* 1999;286:1735–1738. [PubMed: 10576740]
28. Paddison PJ, Caudy AA, Sachidanandam R, Hannon GJ. Short hairpin activated gene silencing in mammalian cells. *Methods Mol Biol* 2004;265:85–100. [PubMed: 15103070]
29. John B, et al. Human MicroRNA targets. *PLoS Biol* 2004;2:e363. [PubMed: 15502875]
30. Mavrakis KJ, et al. Tumorigenic activity and therapeutic inhibition of Rheb GTPase. *Genes Dev* 2008;22:2178–2188. [PubMed: 18708578]
31. Wendel HG, et al. Survival signalling by Akt and eIF4E in oncogenesis and cancer therapy. *Nature* 2004;428:332–337. [PubMed: 15029198]
32. Yang, YH.; Paquet, A.; Dudoit, S. R package version 1.12.0. 2006.
33. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;98:5116–5121. [PubMed: 11309499]
34. Efron B, Tibshirani R. On testing the significance of sets of genes. *Annals of Applied Statistics* 2007;1:107–129.

Figure 1. miR-19 enhances cytokine independent survival *in vitro*

a, Genomic organization of the 17~92 cluster (including "oncomir-1") and its paralogues, miRNAs shown in identical colour share common seed sequences; **b**, Schematic of the competition assay for cytokine independent survival of immortalized FL5-12 lymphocytes; **c**, Representative FACS profiles showing enrichment of miR-19/GFP expressing FL5-12 cells upon IL3 depletion, while populations transduced with vector or the other miRNAs remain unchanged (all experiments in triplicates).

Figure 2. miR-19 is a novel T-ALL oncogene

a, qRT-PCR measurement of miR-19 expression in a panel of human lymphoid malignancies: (T-/B-ALL) T- and B-cell acute lymphatic leukaemia; (FL) follicular lymphoma; (DLBCL) diffuse large B-cell lymphoma; (BL) Burkitt's lymphoma; (HD) Hodgkin's disease; (Tonsil) lymphocytes from reactive tonsils; (F) indicates FL5-12 cells, both parental (black bar) and miR-19 transduced (red bar) shown are mean $\pm$ SD; **b**, Double colour FISH analysis of t(13;14)(q32;q11) using a RB1 probe (green) in 13q14 and genomic clones RP11-97P7 and RP11-980D6 overlapping the 17~92 locus in 13q32 (red). **c**, Graphic representation of FISH results; **d**, Mouse model of *Notch*-induced T-ALL; **e**, Kaplan-Meier analysis of leukaemia free survival after HPC transplantation (Red: Notch-ICN + miR-19; n=6; black: Notch-ICN + vector; n=9); **f-h**, Representative microphotographs of *Notch*/miR-19 induced ALL; **f**, Leukaemic blasts on blood film; **g**) Effacement of the bone marrow by miR-19/GFP expressing leukaemic cells; **h**) Splenomegaly and lymphomas. The pathologic appearance of *Notch1* induced leukaemia is identical (not shown).

Figure 3. Gene expression analysis of parental and miR-19 transduced FL5-12 cells

a, Heat-map illustration of the unsupervised clustering analysis reveals differences in gene expression between parental (FL/Vector) and miR-19 expressing FL5-12 cells (FL/miR-19); **b**, Comparison of the expression change of predicted miR-19 targets represented on the array (336 genes, red line) versus all represented genes (8065 genes, black line) ($p < 2e-04$; KS-test); **c**, Histogram of genes whose expression is down regulated by $> 1SD$ in FL5-12/miR-19 cells compared to parental cells: miR-19 target genes are not over represented in the among genes showing more pronounced (1.5 or 2 SD) reductions in expression. Expression array readily detects global down regulation by miR-19, but is not a sufficient filter to define key miR -19 targets.

Figure 4. Genetic screen for shRNAs that phenocopy miR-19 in lymphocyte survival

a, Design schematic of the pooled shRNA screen, where large populations of FL5-12 cells are transduced with library pools and subjected to IL-3 depletion; **b**, Heat-map of the unsupervised clustering analysis of the half-hairpin array results. (T1, T2) indicate time points after 1 and 2 cycles of IL3 depletion; (-IL3) treated and (+IL3) untreated groups; (A-C) are the replicates (replicate B at T2/+IL3 was removed for technical reasons); **c**, Statistical analysis of Microarray data (SAM) indicating fold change (log₂) for individual shRNAs and the threshold for validation studies (red line); **d**, Gene Set enrichment analysis (GSA) identifies enrichment of sets of shRNAs targeting 14 genes (red circle); **e**, Representative FACS profiles from triplicate validation experiments showing the enrichment of FL5-12 cells expressing the indicated shRNAs and GFP upon IL3 depletion. Vector alone showed no change (not shown).

Figure 5. The identified genes are actual and relevant targets of miR-19

a, qRT-PCR for the indicated genes on cDNA prepared from vector (V; black bars) or miR-19 transduced (19; shaded bars) FL5-12 cells. Expression levels (mean $\pm$ SD) are normalized to the vector controls (relative expression); **b**, qRT-PCR comparing FL5-12 cells transduced with vector (V; black bars) and antagomirs to miR-19 (a19; shaded bars) analyzed as above; **c**, Immunoblot on lysates from vector, miR-19 or antagomir (Anti-19) transduced FL5-12 cells probed for the indicated proteins (see also Suppl. Fig. 17); **d**, Representative FACS profiles of triplicate experiments on FL5-12 cells co-expressing Bcl2 (FL5-12/Bcl2) and the indicated shRNAs or miR-19 and GFP. Cells are shifted from complete (+IL3) to IL3 deficient media (-IL3) to assess Bcl2 independent effects on lymphocyte survival; **e**, FACS analysis of FL5-12/Bcl2 cells transduced with antagomir (Anti-19) or vector and GFP and grown in complete media; **f**, *In vivo* knockdown of individual target genes accelerates *Notch1*-induced T-ALL *in vivo*. Kaplan-Meier analysis of survival after HPC transplantation (Black: vector n=10; green: shPten n=4; red: shBim n=6; blue: shPrkaa n=5; yellow: shPp2r5e n=4); **g**, Diagram indicating the multilevel control of PI3K survival signals by miR-19.