Supplementary Methods

RNAi screen

The shRNAs were expressed in a mir30 context and the precise vector construction is detailed in the referenced mir30 shRNA paper (Dickins et al., 2005). To preserve library complexity, a minimum of 200-fold coverage of shRNA libraries was maintained at each step in the screens. Leukemia cells were infected with one of five individual pools of 10,000 shRNAs. For each pool of 10,000 shRNAs, at least 50 million leukemia cells were infected to a final percentage of 5-10%. The number of mouse genes targeted is approximately 15,000. The average number of shRNAs per gene is 2. The average number of sequencing reads mapped to each 10K pool was 13 million.

Spin infections were performed in the presence of 4 ug/mL polybrene. 48 hours after infection, at least $2x10^6$ shRNA-infected cells were sorted based on GFP expression. Cells were expanded for in culture for 24-48 hours after sorting. After expansion, $2x10^6$ leukemia cells/mouse were injected into three non-irradiated syngeneic recipient mice by tail vein injection. At the time of injection, 4 million cells were collected to serve as the input sample. Disease progression was monitored by peripheral blood counts. Following the appearance of overt leukemia, approximately 12 days after tail-vein injection, leukemia cells were harvested from the blood of mice. For *in vitro* samples, $2x10^6$ infected leukemia cells were plated in triplicate, expanded, and propagated for twelve days. Genomic DNA was isolated by proteinase K digestion and isopropanol precipitation. The antisense strand of shRNAs was amplified from genomic DNA using primers that include 1-basepair mutations to barcode individual samples. All PCR steps were performed in a UV hood to avoid cross-contamination between samples.

Hairpins were amplified in multiple 50 µL reactions using HotStar Taq (Qiagen). After PCR amplification, samples were pooled and prepared for sequencing with Illumina's genomic adaptor kit. At least 41 bases of the PCR product were sequenced with an Illumina HiSeq 2000 machine. shRNAs with less than 100 reads in the input sample were excluded from further analysis, and read numbers for each shRNA were normalized to the total read numbers per sample to allow for cross-comparison between samples. PCR primer sequences are provided in Supplementary Table S7.

For GFP competition assays, pure populations of mCherry positive ALL cells were partially transduced with single shRNA constructs co-expressing GFP, and were infected to a final percentage of 40-60% GFP positive cells. $1x10^6$ cells were injected into 6wk C57/BL6 females, and at the time of terminal disease, as determined by body condition score, leukemia cells were harvested from peripheral blood of mice. For RT-qPCR, total RNA was isolated from ALL cells after retroviral shRNA infection and GFP sorting.

Validation library cloning

We automated the design of shRNAs against the entire mouse and human transcriptomes, and stored the results in a publically accessible database for future rapid use. Known RefSeq entries for mRNA transcripts were retrieved from the NCBI Core Nucleotides database (http://www.ncbi.nlm.nih.gov/nuccore). Each transcript sequence was input to a locally generated perl script designed to replicate the results of the web tool "siRNA Scales"[1] (http://gesteland.genetics.utah.edu/siRNA_scales/) while being amenable to batch input. The result of this was a list of predicted efficacious siRNAs for each transcript. Those with a score of

30 or lower (i.e., predicted to knock down the expression of the target mRNA by at least 70 percent) were retained. All resultant shRNA hairpin sequences were stored in a MySQL database that is publically available via a web interface at http://shrna.mit.edu/. Here, a user can enter the name of one or more human or mouse genes and rapidly retrieve a desired number of pre-designed shRNAs designed against each. Options are available to exclude shRNA that includes any arbitrary subsequence and were bulk synthesized by LC Sciences (LC Sciences OligoMix).

Oligos were PCR amplified using the following primers:

XhoI 5' primer; 5' CAGAAGGCTCGAGAAGGTATATTGCTGTTGACAGTGAGCG 3'

EcoRI 3' primer; 5' CTAAAGTAGCCCCTTGAATTCCGAGGCAGTAGGCA 3'.

High-fidelity Platinum Taq (Invitrogen) was used for shRNA amplification, and amplified shRNAs were batch cloned into a mir30 retroviral vector (Dickins et al., 2005). *In vivo* and *in vitro* screens using this batch-cloned validation library were performed as described above, but five *in vitro* and *in vivo* replicates were generated.