Supplemental Information

Supplemental Data

Supplemental Figure 1 Validated miRNA functions for each of the models of miRNA-directed gene expression regulation

For each of the mechanistic models described in Figure 1 of the manuscript, a miRNA – gene set association, validated in literature, is presented. The plot shows the Gene Set Enrichment Analysis results for each respective miRNA and the gene set listed above the plot. Below each plot is a schematic representation of validated interactions between the miRNA and different components within the gene set.

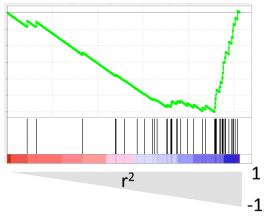
(A) miRNAs from the miR-17-92 cluster are negatively correlated to a gene set containing TGFβ-target genes. miR-17-92 miRNAs have been shown to target multiple components of the TGFβ-pathway (Mestdagh *et al*, 2010a). (B) miR-9 is negatively correlated to a set of CDH1 responsive genes. miR-9 has been shown to target CDH1, a component of the CDH1 – CTNNB1 axis (Ma *et al*, 2010). (C) miR-107 is negatively correlated to a set of Hypoxia activated genes that is enriched for targets of the HIF1 transcription factor. miR-107 has been shown to target HIF1B (Yamakuchi *et al*, 2010). (D) miR-17 is positively correlated to cell cycle genes. miR-17 is known to target CDKN1A, a negative regulator of the cell cycle (Fontana *et al*, 2008). (E) miR-181a is positively correlated to a set of MYC target genes. miR-181a is a direct transcriptional target of the MYCN transcription factor (Mestdagh *et al*, 2010b).

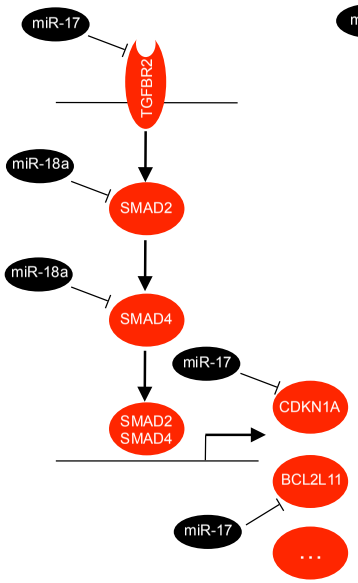
A

multi-component targeting

miR-17-92 ~ TGFβ

EARLY_RESPONSE_TO_TGFB1



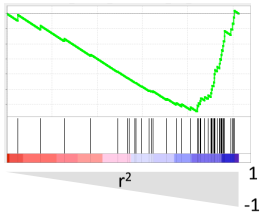


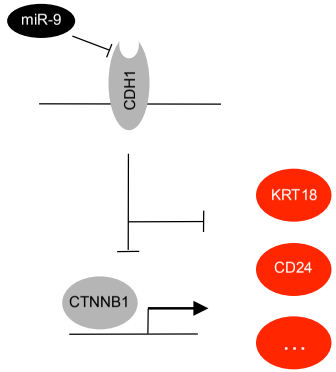
B

component targeting

miR-9 ~ E-cadherin

ONDER_CDH1_TARGETS_2_DN



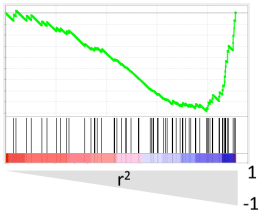


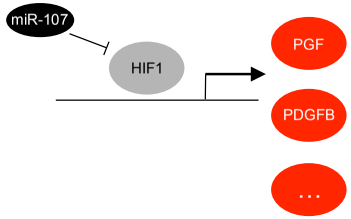
C

transcription factor targeting

miR-107 ~ HIF1

MANALO_HYPOXIA_UP



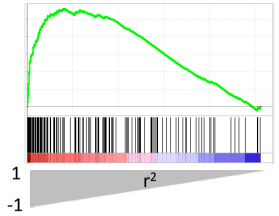


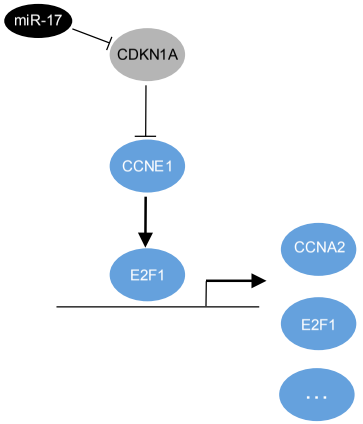
D

targeting negative regulator

miR-17 ~ CDKN1A

CELL_CYCLE_PROCESS



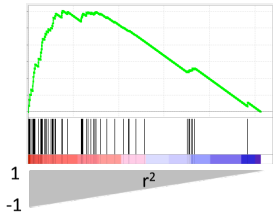


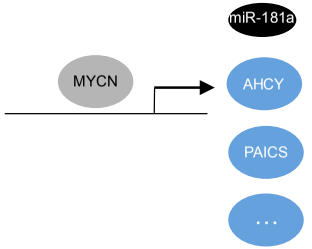
E

common transcriptional activator/repressor

MYCN ~ miR-181a

SCHUHMACHER_MYC_TARGETS_UP





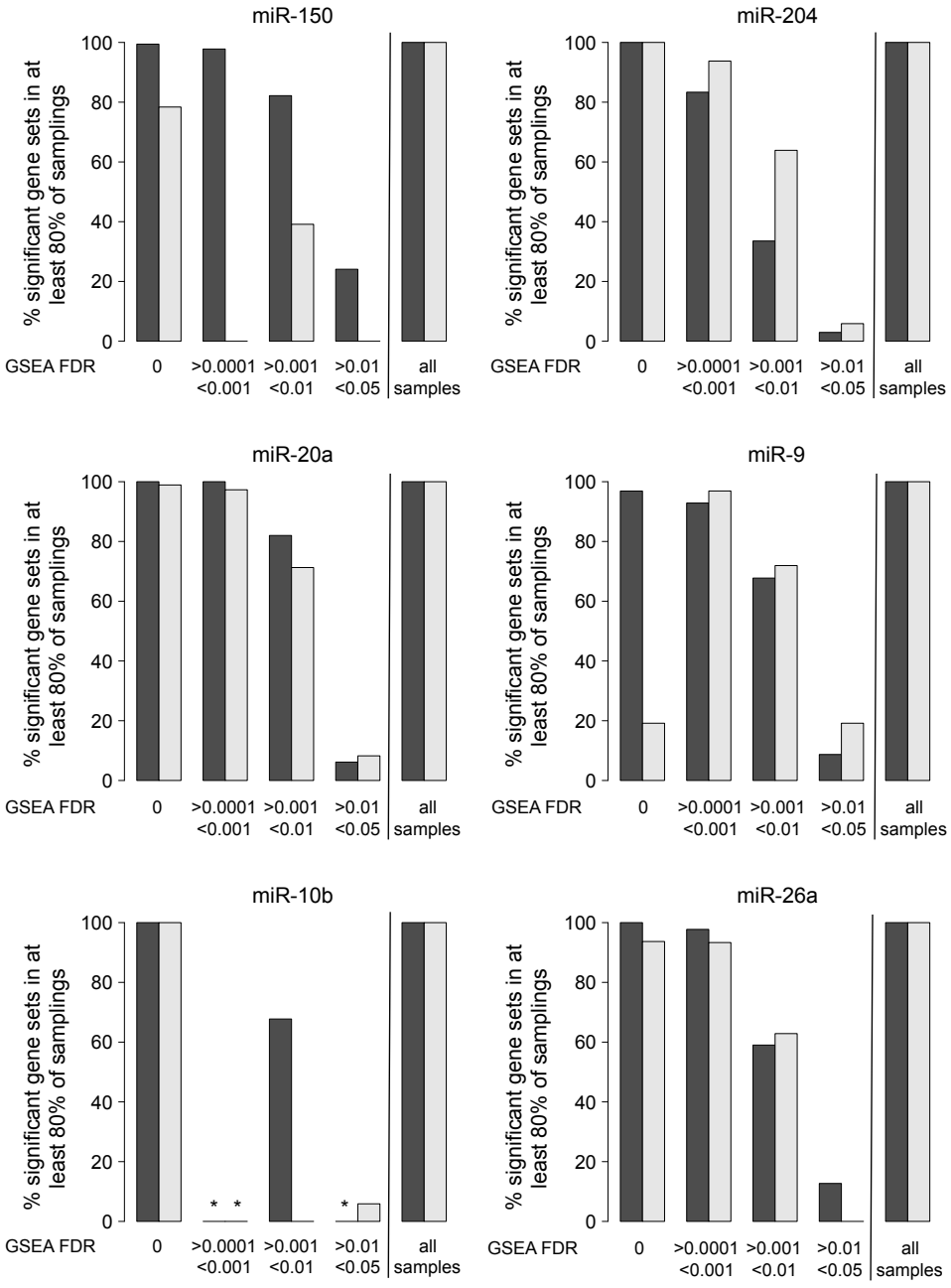
gene negatively correlated to miRNA

gene positively correlated to miRNA

Supplemental Figure 2 impact of sample size on predicted miRNA functions

The function of a miRNA depends on the cellular environment that dictates which of the putative target genes are expressed. Tissue or disease specific miRNA functions have been previously described, amongst others for the miR-17-92 cluster that can function either as a tumour suppressor or an oncomir. To evaluate whether the predicted miRNA functions can differ between datasets containing different tissue types, we compared miRNA predictions between the neuroblastoma and myeloma dataset. Only the most significant predictions (GSEA FDR=0) were used in the analysis as we found these to be independent of sample size. Rather surprisingly, the majority of predicted miRNA functions were specific to only one of the datasets suggesting that predicted miRNA functions cannot always be generalized and should be interpreted within a tissue or cell-specific context. These results clearly support previous observations that miRNAs often function in a tissue specific manner. Further experiments are needed to validate the extent of these findings.

Each graph in Supplemental Figure 3 represents the overlap in predicted miRNA functions when repeatedly (n = 50) sampling 60 samples from the neuroblastoma dataset. Significant gene sets, selected using different GSEA FDR cut-offs, that are identified when using the entire neuroblastoma dataset were compared to significant gene sets identified when using only 60 samples. Gene sets that were identified as significant in at least 40 out of 50 samplings (power 80%) were considered to be common between both datasets and thus independent of sample size. The Y-axis shows the percentage of gene sets that were identified in at least 40/50 samplings relative to the total number of gene sets that can be identified when using all samples (set to 100%). Positive miRNA – gene set correlations are indicated in dark grey, negative miRNA – gene set correlations in light grey. Most gene sets with a GSEA FDR = 0 are also identified when using only 60 samples.


miR-150
miR-204
miR-20a
miR-9
miR-10b
miR-26a
% significant gene sets in at least 80% of samplings
100
80
60
40
20
0
GSEA FDR
0
>0.0001 <0.001
>0.001 <0.01
>0.01 <0.05
all samples
*

Supplemental Experimental Procedures

miRNA prediction databases

The following miRNA target prediction databases were used: MIRDB release 14 (Wang, 2008; Wang and El Naqa, 2008), Targetscan 5.1 (Friedman *et al*, 2009), MICROCOSM v5 (Griffiths-Jones *et al*, 2008), DIANA 3.0 (Maragkakis *et al*, 2009a; Maragkakis *et al*, 2009b), RNA22 (august 2007) (Miranda *et al*, 2006) and PITA v6 (Kertesz *et al*, 2007).

Impact of sample size on predicted functions

In order to compare miRNA predictions between the myeloma and neuroblastoma dataset we first evaluated the impact of the sample size on miRNA annotations, as the number of samples between both datasets differs (the neuroblastoma dataset consists of 99 tumour samples while the myeloma dataset contains 60 tumour samples). We repeatedly (n = 50) selected 60 random samples within the neuroblastoma dataset to evaluate the impact on the inferred miRNA annotations. Individual miRNA – mRNA correlations were recalculated for a selection of 6 miRNAs and GSEA was performed to determine miRNA functions as described earlier. For each of these miRNAs, significant gene sets that were identified when using the entire dataset were compared to those obtained when selecting only 60 samples. The most significant miRNA - gene set associations (GSEA FDR=0) remained significant when reducing the sample size to 60 whereas for less significant miRNA - gene set associations this was not always the case (Supplemental Figure 3).

MYCN 3'UTR luciferase assay

74 bp oligonucleotides spanning the predicted 3'UTR miRNA binding site and flanked by XhoI and NotI restriction sites were annealed and cloned into psicheck2 (Promega). The following oligonucleotides were used:

MYCN_miR-29a_F
TCGAGACAACAGAAAGTCATTCCTTCTTTTTAAAATGGTGCTTAAGTTCCA
GCAGATGCCACATAAGGGGTTGC
MYCN_miR-29a _R
GGCCGCAACCCCTTATGTGGCATCTGCTGGAACTTAAGCACCATTTTAAA
AAGAAGGAATGACTTTCTGTTGTC
MYCN_miR-29a_mut _F

TCGAGACAACAGAAAGTCATTCCTTCTTTTTAAAAGGTGTCTTAAGTTCCA
GCAGATGCCACATAAGGGGTTGC

MYCN_miR-29a _mut_R
GGCCGCAACCCCTTATGTGGCATCTGCTGGAACTTAAGACACCTTTTAAA
AAGAAGGAATGACTTTCTGTTGTC

Supplemental References

Fontana L, Fiori ME, Albini S, Cifaldi L, Giovinazzi S, Forloni M, Boldrini R, Donfrancesco A, Federici V, Giacomini P, Peschle C, Fruci D (2008) Antagomir-17-5p abolishes the growth of therapy-resistant neuroblastoma through p21 and BIM. *PLoS One* **3:** e2236.

Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19:** 92-105.

Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36:** D154-158.

Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E (2007) The role of site accessibility in microRNA target recognition. *Nat Genet* **39:** 1278-1284.

Ma L, Young J, Prabhala H, Pan E, Mestdagh P, Muth D, Teruya-Feldstein J, Reinhardt F, Onder TT, Valastyan S, Westermann F, Speleman F, Vandesompele J, Weinberg RA (2010) miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat Cell Biol* **12:** 247-256.

Maragkakis M, Alexiou P, Papadopoulos GL, Reczko M, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Simossis VA, Sethupathy P, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG (2009a) Accurate microRNA target prediction correlates with protein repression levels. *BMC Bioinformatics* **10:** 295.

Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG (2009b) DIANA-microT web server: elucidating microRNA functions through target prediction. *Nucleic Acids Res* **37:** W273-276.

Mestdagh P, Bostrom AK, Impens F, Fredlund E, Van Peer G, De Antonellis P, von Stedingk K, Ghesquiere B, Schulte S, Dews M, Thomas-Tikhonenko A, Schulte JH, Zollo M, Schramm A, Gevaert K, Axelson H, Speleman F, Vandesompele J (2010a) The miR-17-92 microRNA cluster regulates multiple components of the TGF-beta pathway in neuroblastoma. *Mol Cell* **40:** 762-773.

Mestdagh P, Fredlund E, Pattyn F, Schulte JH, Muth D, Vermeulen J, Kumps C, Schlierf S, De Preter K, Van Roy N, Noguera R, Laureys G, Schramm A, Eggert A, Westermann F, Speleman F, Vandesompele J (2010b) MYCN/c-MYC-induced microRNAs repress coding gene networks associated with poor outcome in MYCN/c-MYC-activated tumors. *Oncogene* **29:** 1394-1404.

Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* **126:** 1203-1217.

Wang X (2008) miRDB: a microRNA target prediction and functional annotation database with a wiki interface. *RNA* **14:** 1012-1017.

Wang X, El Naqa IM (2008) Prediction of both conserved and nonconserved microRNA targets in animals. *Bioinformatics* **24:** 325-332.

Yamakuchi M, Lotterman CD, Bao C, Hruban RH, Karim B, Mendell JT, Huso D, Lowenstein CJ (2010) P53-induced microRNA-107 inhibits HIF-1 and tumor angiogenesis. *Proc Natl Acad Sci U S A* **107:** 6334-6339.