Volinia et al.

Supplemental Material

Reprogramming of the miRNA networks in cancer and leukemia

Datasets

The expression data analyzed in this study can be accessed at the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) under accession nos. GSE8126 and GSE7055 (Ambs et al. 2008), GSE17155 (Fassan et al. 2009), GSE3467 (He et al. 2005), GSE7828 (Schetter et al. 2008), GSE6857 (Budhu et al. 2008), GSE16654 (Chin et al. 2009), and GSE14936 (Seike et al. 2009), and at ArrayExpress (www.ebi.ac.uk/microarray-as/ae) under accession nos. E-TABM-866 (Pineau et al.), E-TABM-664 (Bloomston et al. 2007), E-TABM-762 and E-TABM-763 (Visone et al. 2009), E-TABM-508 (Pichiorri et al. 2008) , E-TABM-429 (Garzon et al. 2008a), E-TABM-434 (Petrocca et al. 2008), E-TABM-405 (Garzon et al. 2008b), E-TABM-343 (Iorio et al. 2007), E-TABM-41 and E-TABM-42 (Calin et al. 2005), E-TABM-48 (Roldo et al. 2006), E-TABM-22 (Yanaihara et al. 2006), E-TABM-23 (Iorio et al. 2005), E-TABM-46, E-TABM-47, E-TABM-49, and E-TABM-50 (Volinia et al. 2006), and E-MEXP-1796 (Godlewski et al. 2008), E-TABM-37 (Ciafre et al. 2005), E-TABM-341 (Ueda et al.). The additional samples we used here were unpublished or previously reported in context-specific papers (Baffa et al. 2009; Garzon et al. 2006; Garzon et al. 2007; Zhang et al. 2008). The respective accession numbers are: E-TABM-969 and E-TABM-970 for normal tissues, E-TABM-971 for breast cancer, E-TABM-972 and E-TABM-974 for acute myeloid leukemia, E-TABM-973 for chronic lymphocytic leukemia and E-TABM-975 for ovarian cancer.

Supplemental Results and Discussion

The miRNA specificity and network in normal tissues and ES differentiation. Embryonic cell types, such as embryonic bodies, trophoblasts, endoderm, or other stem cells including induced pluripotent stem cells (iPS) also harbored high levels of hsa-miR-302. Since hsa-miR-302 expression only partially decreased throughout these stages of ES differentiation, its IC in embryonic tissues was low (0.5). The miRNAs which had highest tissue specificity in embryos were: hsa-miR-211 in 14 day embryoid bodies (EBs), hsa-miR-10b, hsa-miR-218, hsa-miR-122, and hsa-miR-148a in spontaneous differentiated monolayers, hsa-miR-138 and hsa-miR-338-3p in 7 day EBs, and hsa-miR-99a in trophoblast (Supplemental Figure 2 and Supplemental Table III). The information content (IC) of hsa-miR-302 genes is in our datasets even higher than that found by Landgraf, who considered a lower number of tissues/organs/systems. After hsa-miR-302, the most specific miRNAs are hsa-miR-338-5p, hsa-miR-323-3p, hsa-miR-335, with highest expression in epidermis/nervous system, nervous system and breast respectively (Supplemental Figure 1 and Supplemental Table I). hsa-miR-122 is the miRNA with the highest IC difference between ours and the sequencing dataset. hsa-miR-371-5p is here well represented in embryonic cell types, since we have high hsa-miR-371-5p expression in embryoid bodies, not assayed in the clones' dataset. hsa-miR-129-3p is specific for nervous system in both datasets. We detected hsa-miR-142-5p in hematopoietic system and connective tissues. hsa-miR-9, hsa-miR-128 and hsa-miR-138 were nervous system specific in both datasets. We report twice as many different tissue groups that in the clones' dataset, thus we expected less tissue specific miRNAs and lower information contents. Our hypothesis is confirmed by the IC table (Supplemental Table I). Nevertheless, the IC distributions are very similar: 25 miRNAs with IC higher than 1.35 for the microarrays, in comparison to 29 for the clones. The only notable exception to the lower IC is hsa-miR-302a/b/c. This variation might be

due to the fact that we assayed different cell types of embryonic origin, in comparison to only ES cells for the clones.

Supplemental Methods

Tissue specificity. First, all samples were classified according to their organ-, tissue- and cell-type; then the normal samples were grouped in specific systems and the disease samples in specific pathological states. To assess the specificity of miRNA expression across groups, we needed to estimate what fraction of the total, for a given miRNA belonged to each single group. Therefore, we used the procedure described in the first miRNA expression atlas (Landgraf et al. 2007), with the only exception that we called *Em,t* the value of miRNA *m* in the group *t* referred to as "mean expression value (subtracted of the background value, 100). From here onwards, we essentially proceeded as the reference. The specificity score varies between 0, when the expression level of the miRNA *m* is the same across all tissues, and log2 of the number of tissue types, when only one tissue expresses the miRNA. To minimize artifacts from miRNAs or tissues with very small expression levels, we considered only miRNAs with a total expression value above 10 times the number of normal tissues and above 100 times the number of cancer types; with a minimal expression value (after background subtraction) of 100. For the calculation of overall specificity, we thus included 130 and 133 different mature miRNAs for normal tissues and cancer, respectively. Therefore, it is possible that we missed some specifically expressed miRNAs that in our data had either very low expression or were specific to tissues/diseases that we did not sample sufficiently.

Array CGH. Each cancer sample was compared to a healthy control on a two channel oligonucleotide-based platform. Each gene was evaluated in each sample by using the

normalized log2 ratio (cancer over control). Different probes related to the same gene were averaged (gene symbols were used as keys). Data were normalized according to the providers. As a pre-processing step we retained only those genes with high variability (standard deviation> 0.2). For each gene we computed the 5^{th} and 95^{th} percentiles (only for genes measured in at least 300 samples). A gene harboring recurrent deletions in tumors would result in a low 5^{th} percentile log2 ratio (negative), while one with amplifications would display a high 95^{th} percentile (positive).

To identify the miRNAs with structural alterations in cancer we followed the following procedure (illustrated here for amplifications): we selected all miRNA families where at least 1 member was significantly amplified ($p<0.05$). We defined the family p-value for amplification, as the product of the amplification p-values for each family member (including also the non significant miRNAs), with the following exceptions: i) replicated identical mature miRNAs were considered only when mapping on different loci (i.e. represented by different host genes or flanking genes); ii) physically clustered family members were scored only once. The distinct human genes we assayed in the aCGH dataset were 19,654 in total. We studied 530 distinct miRNA precursors in 308 chromosomal loci, corresponding to 471 different mature miRNAs and 356 distinct miRNA families. The average distance between the miRNAs and their flanking genes was of 188 Kb for the 5' and of 240 Kb for the 3' gene. The number of miRNA loci deleted/amplified in cancer was not significantly different from expectation, when compared to the whole coding genome (in both cases $p>>0.05$).

Bootstrap analysis (random swap between cancer and control channels) was used to simulate gene specific 5^{th} and 95^{th} percentiles. Gene-specific p-values for deletions were calculated as the percentage of resampled 5^{th} percentiles which exceeded the original 5^{th} percentile. We had to take in consideration two phenomena, associated to aCGH, but not linked to cancer: sex

chromosomes and polymorphic copy number variations (CNV). Since the control sample was more frequently from male, while roughly half of the tumors were of female origin, the Y-chromosome genes were incorrectly expected to appear as deleted. Conversely, we expected the X chromosome genes, except for those belonging to the pseudo-autonomic region, to incorrectly appear as amplified. Genes located in the sex chromosomes were indeed behaving exactly as expected (data not shown). Polymorphic CNVs could also display large fold-changes, resulting in high 95^{th} or low 5^{th} percentiles. But, we expected that such CNVs, not associated to cancer, would not display significant p-values. Indeed, most polymorphic sites for copy number variations (CNVs) did not filtered through the aCGH assay, since the different alleles were balanced in the cancer and control groups. Only a small percentage of miRNA coincided with polymorphic CNVs and that fraction was not enriched in the cancer subset (data not shown).

Supplemental Table I. Expression levels of tissue specific microRNAs in normal tissues. Sorted by information content (IC)

Name	Tissue Specificity	Expression Value	Information Content	Chromosomal Location
hsa-miR-302a	EMBRYO	4380	4.00	4:113569339-113569407
hsa-miR-302c	EMBRYO	1709	3.99	4: 113569519-113569586
hsa-miR-302b	EMBRYO	4929	3.73	4: 113569641-113569713
hsa-miR-338-5p	EPIDERMIS / NERVOUS SYSTEM	573 / 451	2.51	17: 79099683-79099749
hsa-miR-323-3p	NERVOUS SYSTEM	167	2.47	14: 101492069-101492154
hsa-miR-335	BREAST	322	2.37	7: 130135952-130136045
hsa-miR-124	NERVOUS SYSTEM	533	2.33	8:9,798,308-9,798,392 8:65,454,260-65,454,368
hsa-miR-139-5p	NERVOUS SYSTEM	245	2.27	11: 72326107-72326174
hsa-miR-371-5p	EMBRYO	443	1.93	19: 54290929-54290995
hsa-miR-325	EPIDERMIS	129	1.91	X: 76225829-76225926
hsa-miR-133a	CARDIOVASCULAR SYSTEM / SKELETAL MUSCLE	464/ 661	1.88	18:17,659,657-17,659,744 20:60,572,564-60,572,665
hsa-miR-184	RESPIRATORY SYSTEM	259	1.86	15: 79502130-79502213
hsa-miR-133b	CARDIOVASCULAR SYSTEM / SKELETAL MUSCLE	394 / 371	1.83	6: 52013721-52013839
hsa-miR-211	NERVOUS SYSTEM	180	1.78	15: 31357235-31357344
hsa-miR-9	NERVOUS SYSTEM	1346	1.73	1:154656757-154656845 5:87998427-87998513
hsa-miR-370	NERVOUS SYSTEM	146	1.72	14: 101377476-101377550
hsa-miR-376c	EPIDERMIS	121	1.61	14: 101506027-101506092
hsa-miR-132	NERVOUS SYSTEM	552	1.59	17: 1953202-1953302
hsa-miR-1	SKELETAL MUSCLE / CARDIOVASCULAR SYSTEM	500 / 350	1.56	20:60,561,958-60,562,028 18:17,662,963-17,663,047
hsa-miR-215	KIDNEY	2630	1.55	1: 220291195-220291304
hsa-miR-339-5p	LIVER	107	1.54	7: 1062569-1062662
hsa-miR-330-3p	LIVER	125	1.53	19: 46142252-46142345
hsa-miR-342-5p	EPIDERMIS	326	1.49	14: 100575992-100576090
hsa-miR-199b-5p	RESPIRATORY SYSTEM	144	1.39	9: 131007000-131007109
hsa-miR-129-3p	NERVOUS SYSTEM	266	1.36	11: 43602944-43603033

Supplemental Table II. Normal Tissues grouped by systems

Normal Organism Part	Normal System	N of Samples
ADIPOSE TISSUE	Adipose	5
BREAST	Breast	4
HEART	Cardiovascular System	6
DERMAL FIBROBLAST	Connective Tissue	11
EMBRYONIC STEM CELLS / DEFINITIVE ENDODERM / SPONTANEOUS DIFFERENTIATED MONOLAYERS / EMBRYONIC BODIES / INDUCED PLURIPOTENT STEM CELLS/	Embryo	68
THYROID / PANCREAS / PITUITARY /	Endocrine Gland	107
SKIN	Epidermis	20
ESOPHAGUS / COLON / STOMACH / RECTUM / SIGMA / SMALL INTESTINE	Gastrointestinal System	394
BLOOD / BONE MARROW / LYMPHONODE / THYMUS / TONSIL / SPLEEN	Hematopoietic System	225
KIDNEY	Kidney	19
LIVER	Liver	38
OCCIPITAL CORTEX / TEMPORAL CORTEX / PRIMARY VISUAL CORTEX / THALAMUS / BRAIN / MEDULLA / CEREBELLUM / PREFRONTAL CORTEX / AUDITORY CORTEX / AMYGDALA/ HYPOTHALAMUS / MOTOR CORTEX / CAUDATE NUCLEUS / PARIETAL CORTEX	Nervous System	58
PLACENTA	Placenta	6
FALLOPIAN TUBE / CERVIX / OVARY / PROSTATE / TESTIS	Reproductive System	59
LUNG / TRACHEA	Respiratory System	75
SKELETAL MUSCLE	Skeletal Muscle	6
BLADDER	Urinary System	6

Supplemental Table III. Expression levels of tissue specific microRNAs during differentiation of embryonic stem cells (ES) . Sorted by information content (IC)

Name	Embryonic tissue Specificity	Expression Value	IC	Chromosomal Location
hsa-miR-211	EB 14 day	165	2.65	15: 31357235-31357344
hsa-miR-10b	Monolayer	909	2.20	2: 177015031-177015140
hsa-miR-138	EB 7 day	181	1.88	3:44,130,708-44,130,806 16:55,449,931-55,450,014
hsa-miR-218	Monolayer	665	1.65	4:20,138,996-20,139,105 5:168,127,729-168,127,838
hsa-miR-122	Monolayer	3660	1.57	18: 56118306-56118390
hsa-miR-99a	Trophoblast	197	1.54	21: 17911409-17911489
hsa-miR-215	Monolayer /Definitive Endoderm	1673 / 325	1.42	1: 220291195-220291304
hsa-miR-338-3p	EB 7 day	432	1.26	17: 79099683-79099749
let-7e	iPS / Embryonic Stem Cells	284 / 181	1.23	19: 52196039-52196117
hsa-miR-371-5p	EB 14 day / EB 7 day / ES	1424 / 6274 / 848	1.20	19: 54290929-54290995
hsa-miR-148a	Monolayer	158	1.18	7: 25989539-25989606
hsa-miR-181c	Trophoblast / Monolayer	1500 / 927	1.07	19: 13985513-13985622
hsa-miR-203	EB 14 day / Monolayer / Trophoblast	974 / 750 / 934	1.05	14: 104583742-104583851
hsa-miR-181a	Trophoblast / Monolayer	2943 / 1106	0.99	1:197,094,796-197,094,905 9:126,494,542-126,494,651
hsa-miR-181b	Trophoblast / Monolayer	8290 / 3397	0.93	1:197,094,625-197,094,734 9:126,495,810-126,495,898
hsa-miR-192	Monolayer / Definitive Endoderm	1804 /474	0.84	11: 64658609-64658718
hsa-miR-150	EB 7 day / iPS	501 / 347	0.83	19: 50004042-50004125
hsa-miR-139-3p	iPS / ES	477 / 250	0.82	11: 72326107-72326174
hsa-miR-27b	Trophoblast / Monolayer	554 / 317	0.77	9: 97847727-97847823
hsa-miR-10a	Monolayer	2547	0.75	17: 46657200-46657309

Legend: Embryonic Bodies: EB; Spontaneous Differentiated Monolayer: Monolayer; Induced Pluripotent Stem Cells: iPS; Embryonic Stem Cells: ES.

Supplemental Table IV. Up-regulated microRNAs in 31 types of solid cancers (2532 cancers samples vs. 806 corresponding normal samples)

Parametric p-value	FDR	intensities in Solid Cancers	intensities in Normal Tissues	Fold change-	microRNA	Chromosomal location
< $1x10^{-07}$	< $1x10^{-07}$	967	617.9	**1.57**	**hsa-miR-21**	17q23.1
< $1x10^{-07}$	< $1x10^{-07}$	1378.2	917.7	**1.5**	**hsa-miR-25**	7q22.1
< $1x10^{-07}$	< $1x10^{-07}$	902.7	626.3	**1.44**	**hsa-miR-20a**	13q31.3
< $1x10^{-07}$	< $1x10^{-07}$	925.7	646.9	**1.43**	**hsa-miR-17**	13q31.3
< $1x10^{-07}$	< $1x10^{-07}$	652.3	469	**1.39**	**hsa-miR-106a**	Xq26.2
< $1x10^{-07}$	< $1x10^{-07}$	410	297.8	**1.38**	**hsa-miR-106b**	7q22.1
$1.30x10^{-06}$	$1.08x10^{-05}$	918.8	697.9	**1.32**	**hsa-miR-146a**	5q34
< $1x10^{-07}$	< $1x10^{-07}$	11893.3	9370.9	**1.27**	**hsa-miR-92a**	13q31.3,Xq26.2
$1.60x10^{-06}$	$1.29x10^{-05}$	2354.9	1919.4	**1.23**	**hsa-miR-103**	5q35.1,20p13
< $1x10^{-07}$	< $1x10^{-07}$	289.4	237.8	**1.22**	**hsa-miR-130b**	22q11.21
$6.81x10^{-05}$	0.000387	750.4	615.7	**1.22**	**hsa-miR-30c**	1p34.2,6q13
< $1x10^{-07}$	< $1x10^{-07}$	452.7	372	**1.22**	**hsa-miR-93**	7q22.1
$3.90x10^{-06}$	$2.74x10^{-05}$	2116.4	1743.6	**1.21**	**hsa-miR-107**	10q23.31
< $1x10^{-07}$	< $1x10^{-07}$	297.7	248	**1.2**	**hsa-miR-30e**	1p34.2
< $1x10^{-07}$	< $1x10^{-07}$	227	191.9	**1.18**	**hsa-miR-15a**	13q14.2
0.000433	0.002	1209.2	1031.6	**1.17**	**hsa-miR-181b**	1q32.1,9q33.3
< $1x10^{-07}$	< $1x10^{-07}$	191.3	164.8	**1.16**	**hsa-miR-15b**	3q25.33
0.000605	0.002648	508.2	438.5	**1.16**	**hsa-miR-181a**	1q32.1,9q33.3
0.000442	0.002002	333.8	287	**1.16**	**hsa-miR-32**	9q31.3
$5.10x10^{-06}$	$3.40x10^{-05}$	160	147.3	**1.09**	**hsa-miR-345**	14q32.2
0.000494	0.0022	208.1	193.5	**1.08**	**hsa-miR-34a**	1p36.22
0.000864	0.00372	154.1	146.2	**1.05**	**hsa-miR-374a**	Xq13.2

*Intensities (expression levels) are reported as geometric means after quantiles normalization.

Supplemental Table V. Down-regulated microRNAs in 31 types of solid cancers (2532 cancer samples vs. 806 corresponding normal samples)

Parametric p-value	FDR	Intensities in Solid Cancers	Intensities in Normal Tissues	Fold Change	microRNA	Chromosomal location
$< 1x10^{-07}$	$< 1x10^{-07}$	275.7	403	**0.68**	**hsa-miR-203**	14q32.33
$< 1x10^{-07}$	$< 1x10^{-07}$	714.9	1015.1	**0.7**	**hsa-miR-145**	5q32
0.000003	$2.23x10^{-05}$	273.7	349.6	**0.78**	**hsa-miR-205**	1q32.2
$2.00x10^{-07}$	$1.80x10^{-06}$	382.5	471.2	**0.81**	**hsa-miR-206**	6p12.2
$< 1x10^{-07}$	$< 1x10^{-07}$	204	247.9	**0.82**	**hsa-miR-33b**	17p11.2
$3.70x10^{-06}$	$2.67x10^{-05}$	419.9	507.8	**0.83**	**hsa-miR-193a**	17q11.2
0.000042	0.000249	175.1	208.5	**0.84**	**hsa-miR-204**	9q21.12
$< 1x10^{-07}$	$< 1x10^{-07}$	173.9	204.2	**0.85**	**hsa-miR-143**	5q32
$1.00x10^{-07}$	0.000001	220.5	260.2	**0.85**	**hsa-miR-326**	11q13.4
$< 1x10^{-07}$	$< 1x10^{-07}$	150.2	176.9	**0.85**	**hsa-miR-338**	17q25.3
$< 1x10^{-07}$	$< 1x10^{-07}$	164.8	193.6	**0.85**	**hsa-miR-9**	1q22,5q14.3, 15q26.1
0.000355	0.001789	190.6	220.4	**0.86**	**hsa-miR-95**	4p16.1
$7.40x10^{-06}$	$4.82x10^{-05}$	225.6	259.4	**0.87**	**hsa-miR-138**	16q13,3p21.33
0.000435	0.002	160.4	183.8	**0.87**	**hsa-miR-183**	7q32.2
$< 1x10^{-07}$	$< 1x10^{-07}$	150.4	172.9	**0.87**	**hsa-miR-202**	10q26.3
$8.24x10^{-05}$	0.000458	529.1	603.1	**0.88**	**hsa-miR-128a**	2q21.3,3p22.3
0.000921	0.003905	617	694.2	**0.89**	**hsa-miR-214**	1q24.3
$< 1x10^{-07}$	$< 1x10^{-07}$	139	153.9	**0.9**	**hsa-miR-132**	17p13.3
$< 1x10^{-07}$	$< 1x10^{-07}$	140.7	156.7	**0.9**	**hsa-miR-299**	14q32.31
$< 1x10^{-07}$	$< 1x10^{-07}$	140.7	154.8	**0.91**	**hsa-miR-129**	7q32.1,11p11.2
$2.00x10^{-07}$	$1.80x10^{-06}$	139.6	152.7	**0.91**	**hsa-miR-133a**	18q11.2,20q13.33
$< 1x10^{-07}$	$< 1x10^{-07}$	129.8	141.1	**0.92**	**hsa-miR-139**	11q13.4
$8.80x10^{-06}$	$5.46x10^{-05}$	153.4	165	**0.93**	**hsa-miR-339**	7p22.3
$4.50x10^{-06}$	$3.08x10^{-05}$	140	149.6	**0.94**	**hsa-miR-1**	20q13.33,18q11.2
$9.68x10^{-05}$	0.000528	140.4	149.6	**0.94**	**hsa-miR-133b**	6p12.2
$< 1x10^{-07}$	$< 1x10^{-07}$	129.7	137.8	**0.94**	**hsa-miR-323**	14q32.31
0.000394	0.001946	144.7	152.7	**0.95**	**hsa-miR-218**	4p15.31,5q34
0.000193	0.00101	130.2	133	**0.98**	**hsa-miR-335**	7q32.2

*Intensities (expression levels) are reported as geometric means after quantiles normalization.

Supplemental Table VI. Expression levels of tissue specific microRNAs in cancer and leukemia

Name	Cancer Specificity	Expression	IC	Chromosomal location
hsa-miR-369-3p	OVARIAN ADENOCARCINOMA	2737	3.95	14:101531935-101532004
hsa-miR-325	PANCREAS NON-FUNCTIONING ENDOCRINE CANCER / INSULINOMA	1625 / 365	3.67	X: 76225829-76225926
hsa-miR-208	CC / BCC / EPIDERMIS SCC	256/ 244 / 418	3.36	14:22,927,645-22,927,715 14:22,957,036-22,957,112
hsa-miR-190	TPC / HEPATOBLASTOMA	211 / 220	3.04	15: 63116156-63116240
hsa-miR-184	AMoL / HEPATOBLASTOMA /TPC / MM	448 / 324 / 214 / 171	3.01	15: 79502130-79502213
hsa-miR-338-5p	EPIDERMIS SCC / BCC / CC	1928 / 808 / 472	2.70	17: 79099683-79099749
hsa-miR-105	HGSIL	234	2.57	X:151,311,347-151,313,620
hsa-miR-376c	LYMPHOMA CUTANEOUS / EPIDERMIS SCC / MM	1134 / 626 / 1223 / 525	2.50	14: 101506027-101506092
hsa-miR-302b	EPIDERMIS SCC / LYMPHOMA CUTANEOUS	336 / 110	2.48	4: 113569641-113569713
hsa-miR-299-5p	HEPATOBLASTOMA / AMoL / TPC	287 / 122 / 148	2.03	14: 101490131-101490193
hsa-miR-19a	LYMPHOMA CUTANEOUS / CML	388 /367	1.83	13: 92003145-92003226
hsa-miR-200c	AMoL / HEPATOBLASTOMA	4460 / 4138	1.99	12: 7072862-7072929
hsa-miR-143	AMoL / HEPATOBLASTOMA	1043 / 1098	1.94	5: 148808481-148808586
hsa-miR-211	EPIDERMIS SCC / NSCLC	129 / 89	1.86	15: 31357235-31357344
hsa-miR-374a	BL SPORADIC / CML / APL	619 / 169 / 212	1.69	X: 73507121-73507192
hsa-miR-339-5p	BL ENDEMIC / BL SPORADIC	616 / 738	1.62	7: 1062569-1062662
hsa-miR-148a	LYMPHOMA CUTANEOUS	1161	1.64	7: 25989539-25989606
hsa-miR-370	BREAST, DUCTAL, LOBULAR, MALE CARCINOMA	157, 227, 239	1.69	14: 101377476-101377550
hsa-miR-199a-5p	HEPATOBLASTOMA	1122	1.52	19:10,789,102-10,789,172 1:170,380,298-170,380,407
hsa-miR-215	KIDNEY CARCINOMA / HEPATOBLASTOMA	7247 /2461	1.45	1: 220291195-220291304
hsa-miR-19b	CML	232	1.33	13:90,801,447-90,801,533 X:133,131,367-133,131,462
hsa-miR-148b	LYMPHOMA CUTANEOUS	657	1.39	12: 54731000-54731098
hsa-miR-18a	AMoL / FOLLICULAR ADENOMA	166 / 76	1.38	13: 92003005-92003075
hsa-miR-202	AMoL/ HEPATOBLASTOMA	362 / 287	1.32	10: 135061015-135061124
hsa-miR-205	AMoL / HEPATOBLASTOMA / LUNG SCC	2029 / 2275 / 2665	1.31	1: 209605478-209605587

Legend: Burkitt's Lymphoma: BL; Thyroid Papillary Cancer: TPC; Multiple Myeloma: MM; Squamous Cell Carcinoma: SCC; Basal Cell Carcinoma: BCC; Chronic Myelogenous Leukemia: CML; Acute Promyelocytic Leukemia: APL; High Grade Squamous Intraepithelial Lesion: HGSIL; Non-Small Cell Lung Cancer: NSCLC; Acute Monocytic Leukemia: AmoL; Cervix Carcinoma: CC.

Supplemental Table VII. Array CGH datasets

GEO datasets	**# of samples**	**Cancer type**	**Platform**
GSE4659	32	AML	GPL2873
GSE6472	6	Nasopharyngeal carcinoma	GPL2879
GSE7077	4	Osteosarcoma	GPL2879
GSE7344	16	Glioma	GPL2873
GSE7482	25	ACC	GPL2879
GSE7615	298	Pancreas, glioblastoma, T-ALL, melanoma, colon	GPL2879 GPL4091
GSE7822	14	Melanoma	GPL2879
GSE8398	25	Ewing Sarcoma	GPL2879
GSE8804	13	Myelodysplasia	GPL2879
GSE8918	87	Follicular Lymphoma, CLL, Mantle Cell Lymphoma, Nodal Marginal Zone Lymphoma, Lymphoplasmacytic Lymphoma, Splenic Marginal Zone Lymphoma , MALT	GPL2879
GSE9015	7	Breast	GPL4091
GSE9654	10	Osteosarcoma	GPL2879
SMD	207	Lung, Pancreas, Breast, Melanoma, Fibrous Histiocytoma	Stanford

- GPL2873 Agilent- Human Genome CGH Microarray 44A G4410A
- GPL2879 Agilent- Human Genome CGH Microarray 44B G4410B
- GPL4091 Agilent- Human Genome CGH Microarray 244A G4411B

Supplemental Table VIII. miRNA families amplified in cancer

miRNA family	p-value	FDR	Amplified (p<0.05)	Family members
hsa-miR-550	$1.25x10^{-09}$	0.006	*hsa-miR-550-2;hsa-miR-550-1*	2
hsa-miR-9	$7.00x10^{-11}$	0.012	*hsa-miR-9-1;hsa-miR-9-2; hsa-miR-9-3*	3
hsa-miR-133	$2.50x10^{-09}$	0.060	*hsa-miR-133a-2;hsa-miR-133b*	3
hsa-miR-153	$5.00x10^{-07}$	0.062	*hsa-miR-153-1;hsa-miR-153-2*	2
hsa-miR-548d	$1.40x10^{-06}$	0.073	*hsa-miR-548d-1;hsa-miR-548d-2*	2
hsa-miR-1/206	$7.55x10^{-10}$	0.114	*hsa-miR-1-1;hsa-miR-206*	4
hsa-miR-146	$1.97x10^{-05}$	0.116	*hsa-miR-146a*	2
hsa-miR-148/152	$2.00x10^{-07}$	0.133	*hsa-miR-152;hsa-miR-148a*	3
hsa-miR-196	$5.54x10^{-07}$	0.150	*hsa-miR-196b;hsa-miR-196a-2*	3
hsa-miR-17/20/93/ 106/519	$3.70x10^{-06}$	0.194	*hsa-miR-106b*	3
hsa-miR-30-5p	$1.06x10^{-05}$	0.226	*hsa-miR-30d*	3
hsa-miR-192/215	$4.95x10^{-05}$	0.239	*hsa-miR-192*	2
hsa-miR-194	$4.95x10^{-05}$	0.239	*hsa-miR-194-2*	2
hsa-miR-551	$4.97x10^{-05}$	0.239	*hsa-miR-551b*	2

* FDR < 0.3 (i.e. less than 30% of the identified miRNA families were expected to be false positives)

Supplemental Table IX. miRNA families deleted in cancer

miRNA family	p-value	FDR	Deleted (p<0.05)	Family members
hsa-miR-204/211	6.25×10^{-10}	0.003	*hsa-miR-211;hsa-miR-204*	2
hsa-miR-103/107	2.45×10^{-09}	0.064	*hsa-miR-107;hsa-miR-103-2*	3
hsa-miR-380-5p	2.85×10^{-06}	0.109	*hsa-miR-563*	2
hsa-miR-128	3.25×10^{-06}	0.113	*hsa-miR-128a*	2
hsa-miR-141/200a	1.52×10^{-05}	0.170	*hsa-miR-200a*	2
hsa-miR-200bc/429	1.52×10^{-05}	0.170	*hsa-miR-429*	2
hsa-miR-125/351	3.50×10^{-07}	0.189	*hsa-miR-125b-1;hsa-miR-125a*	3
hsa-miR-218	2.41×10^{-05}	0.193	*hsa-miR-218-1*	2
hsa-miR-101	2.70×10^{-05}	0.230	*hsa-miR-101-1;hsa-miR-101-2*	2
hsa-miR-99/100	4.38×10^{-05}	0.241	*hsa-miR-99b*	2
hsa-miR-23	8.80×10^{-05}	0.290	*hsa-miR-23b;hsa-miR-23a*	2
hsa-miR-24	8.80×10^{-05}	0.290	*hsa-miR-24-1;hsa-miR-24-2*	2
hsa-miR-27	8.80×10^{-05}	0.290	*hsa-miR-27b;hsa-miR-27a*	2

* FDR < 0.3 (i.e. less than 30% of the identified miRNA families were expected to be false positives)

Supplemental Table X. Deregulated miRNAs in leukemia from Mir155 transgenic mice.

Parametric p-value	FDR	Intensity in hsa-miR-155 Leukemia	Intensity in Wt	Fold Change	miRNA
$1.00x10^{-06}$	$2.31x10^{-05}$	832.1	10.08	**82.57**	**hsa-miR-224**
$4.20x10^{-06}$	$5.67x10^{-05}$	3426.09	132.36	**25.89**	**hsa-miR-9**
$1.00x10^{-07}$	$7.20x10^{-06}$	21760.18	1005.44	**21.64**	**hsa-miR-217**
$5.70x10^{-06}$	$7.39x10^{-05}$	3387.97	219.75	**15.42**	**hsa-miR-131**
$1.00x10^{-07}$	$7.20x10^{-06}$	5334.39	375.03	**14.22**	**hsa-miR-128b**
$1.00x10^{-07}$	$7.20x10^{-06}$	31460.89	2922.63	**10.76**	**hsa-miR-181b**
$2.00x10^{-07}$	$7.20x10^{-06}$	17040.98	1615.73	**10.55**	**hsa-miR-181c**
$2.00x10^{-07}$	$7.20x10^{-06}$	18356.22	1976.64	**9.29**	**hsa-miR-128a**
$1.28x10^{-05}$	0.000148	759.87	2695.34	**0.28**	**hsa-miR-140**
0.000659	0.003556	228.42	874.66	**0.26**	**hsa-miR-135a**
$3.69x10^{-05}$	0.000332	8491.07	34327.47	**0.25**	**hsa-miR-26a**
0.000128	0.000984	528.06	2433.76	**0.22**	**hsa-miR-10a**
$3.50x10^{-06}$	$5.55x10^{-05}$	590.74	2779.66	**0.21**	**hsa-miR-425**
0.000193	0.001351	514.05	2467.2	**0.21**	**hsa-miR-340**
0.000474	0.002742	81.06	403.52	**0.2**	**hsa-miR-24**
$2.00x10^{-06}$	$3.78x10^{-05}$	656.68	3304.56	**0.2**	**hsa-miR-140**
0.000579	0.003178	30.24	196.77	**0.15**	**hsa-miR-218**
$2.55x10^{-05}$	0.000243	1006.31	6657.6	**0.15**	**hsa-miR-29b**
$6.00x10^{-07}$	$1.62x10^{-05}$	1007.33	7465.65	**0.13**	**hsa-miR-29a**
$2.54x10^{-05}$	0.000243	357.49	3183.78	**0.11**	**hsa-miR-29c**
$< 1x10^{-07}$	$< 1x10^{-07}$	2176.08	37081.81	**0.06**	**hsa-miR-146**
$2.70x10^{-06}$	$4.60x10^{-05}$	3339.45	59257.53	**0.06**	**hsa-miR-150**

Legends to Supplemental Figures

Supplemental Figure 1. miRNA specificity in 50 normal tissues grouped by system. The tissue specificity was calculated by using the information content (IC), value expressed on y-axis; each color represents a system. The hsa-miR-302 cluster is the most representative for embryonic tissues.

Supplemental Figure 2. The miRNA specificity during ES cell differentiation. miRNA specificity in 7 different types of embryonic tissues (embryonic stem cells, 7 days and 14 days embryonic bodies, trophoblasts, endoderm, induced pluripotent stem cells (iPS), spontaneously differentiating monolayers). The specificity was calculated by using the information content (IC).

Supplemental Figure 3. The node degree distribution of the normal tissues and solid cancers miRNA networks. The figure illustrates the apparent scale-freeness of the graphs for normal tissues (A) and solid cancers (B). The blue curve represents the absolute frequency of node degrees and the green curve the inverse cumulative frequency (Supplementary data excel file). The exponential decrease of both curves shows that there are a lot more poorly connected nodes than highly connected (hubs). Both miRNA graphs thus present a scale free behavior.

Supplemental Figure 4. miRNA specificity in 31 solid tumors and 20 leukemia types. The specificity was calculated by using the information content (IC), value expressed on y-axis; each color represents a cancer type.

Supplemental Figure 5. Functions repressed in cancer by up-regulated miRNAs. KEGG analysis by ClueGO (Bindea et al. 2009) of terms (p-value $<1x10^{-3}$) targeted by up-regulated miRNAs (listed in Table I and Supplemental Tables IV). Target genes selection was performed with DIANA-miRpath, microT-V4.0 (Papadopoulos et al. 2009). The union of the target mRNAs with a score above 3 was used as an input to ClueGO. Right-sided hyper-geometric test yielded

the enrichment for GO-terms. Benjamini-Hochberg correction for multiple testing controlled the p-values. GO term fusion was applied for redundancy reduction.

Supplemental Figure 6. Functions activated in cancer by down-regulated miRNAs. KEGG analysis by ClueGO (Bindea, Mlecnik, Hackl, Charoentong, Tosolini, Kirilovsky, Fridman, Pages, Trajanoski and Galon 2009) of terms (p-value $<1x10^{-3}$) targeted by down-regulated miRNAs (listed in Table I and Supplemental Tables V). Target genes selection was performed with DIANA-miRpath, microT-V4.0 (Papadopoulos, Alexiou, Maragkakis, Reczko and Hatzigeorgiou 2009). The union of the target mRNAs with a score above 3 was used as an input to ClueGO. Right-sided hyper-geometric test yielded the enrichment for GO-terms. Benjamini-Hochberg correction for multiple testing controlled the p-values. GO term fusion was applied for redundancy reduction.

Supplemental Figure 7. Functions controlled in lung adenocarcinoma by miRNAs from disjointed minor clusters. The colored groups are assigned the name of the most prominent GO group. The chart presents the specific terms and information related to the miRNA targets differentially expressed in cancer. The bars represent the number of the genes from the analyzed cluster found to be associated with the term, and the label displayed on the bars is the percentage of found genes compared to all the genes associated with the term. Term significance information is included in the chart.

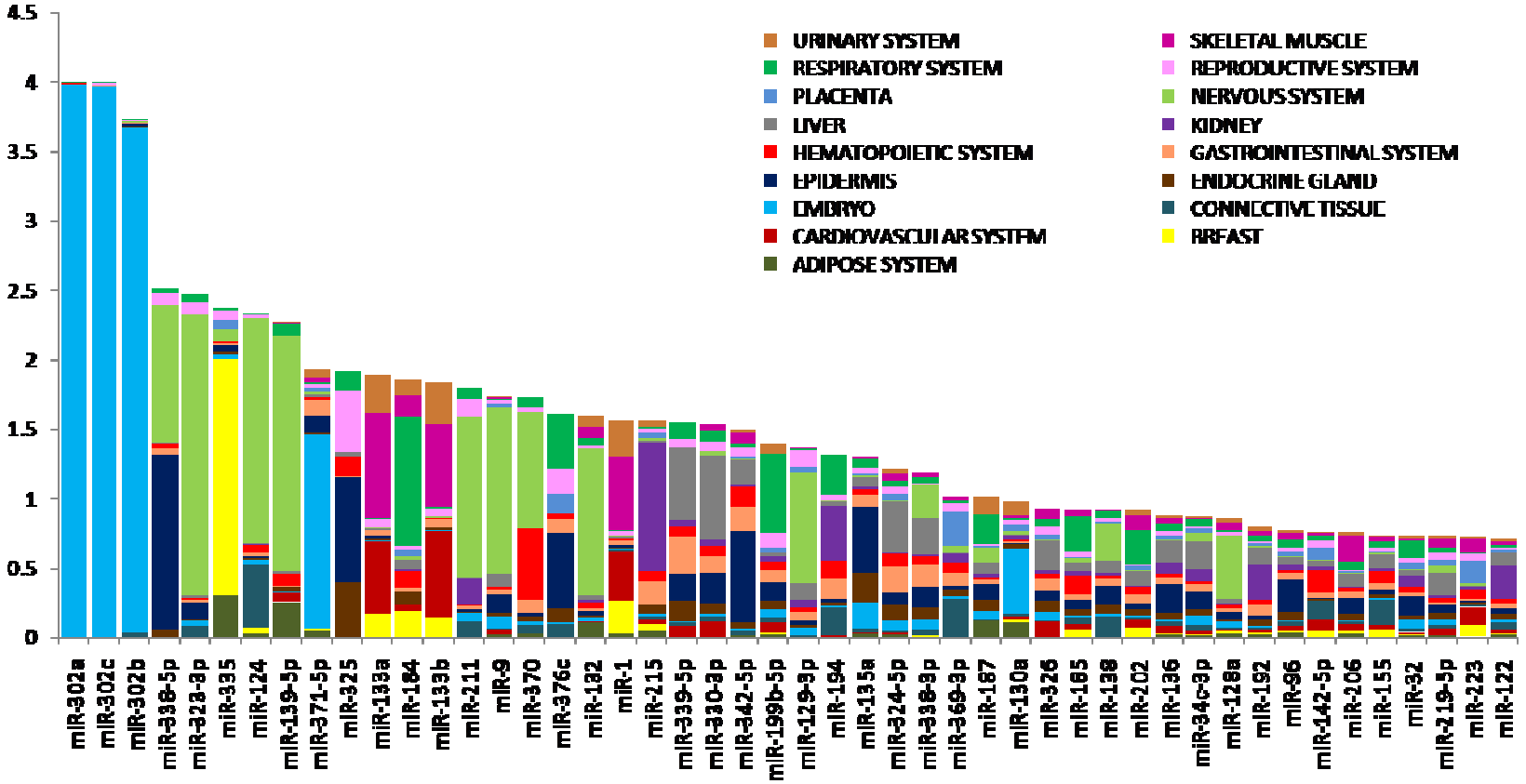
Supplemental Figure 8. MicroRNA genetic network in colon adenocarcinoma (245 samples, 10 subnets). The network was inferred for all expressed and varying miRNAs, without preselecting for differential expression. Standard Banjo parameters were adopted with a q6 discretization policy. The consensus graph depicted here was obtained from the best 100 nets (after searching through $>1x10^{10}$ networks). MCL graph-based clustering algorithm was applied to clusters extraction (miRNAs with highly related expression pattern have edges of the same color); thus, different clusters in the network are linked by different color edges. yEd graph editor

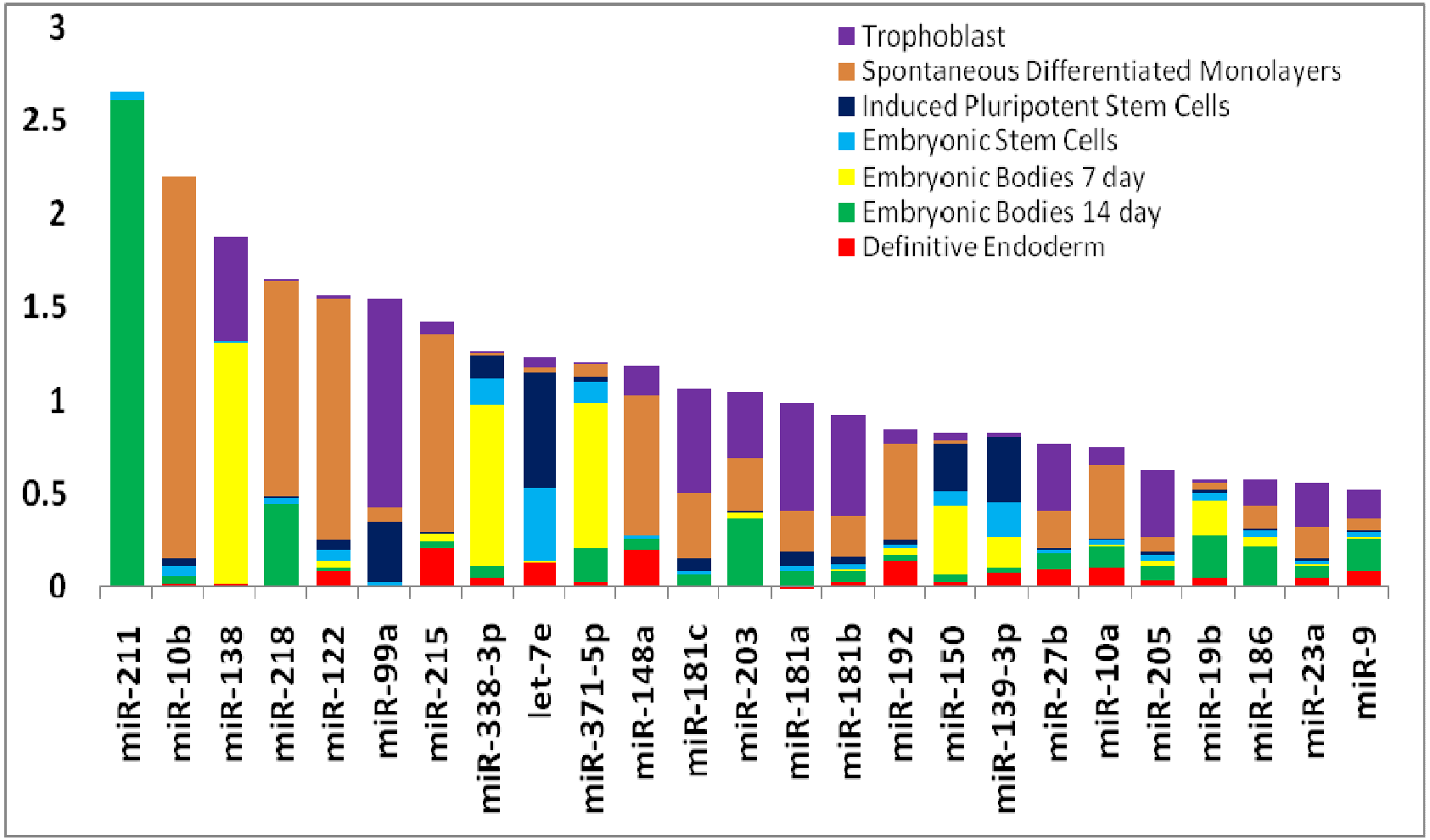
(yFiles software, Tubingen, Germany) was employed for graphs visualization. See text for further discussion.

Supplemental Figure 9. MicroRNA genetic network in breast cancer (159 samples, 14 subnets). Same procedure described in legend of Supplemental Figure 6.

Supplemental Figure 10. MicroRNA genetic network in prostate cancer (170 samples and 5 subnets). Same procedure described in legend of Supplemental Figure 6.

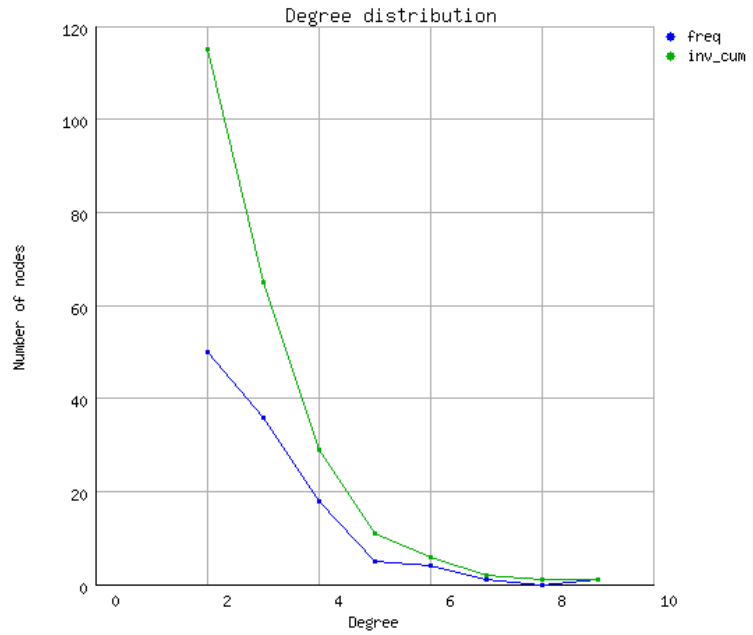
Supplemental Figure 1. miRNA specificity in 50 normal tissues grouped by system (information content).



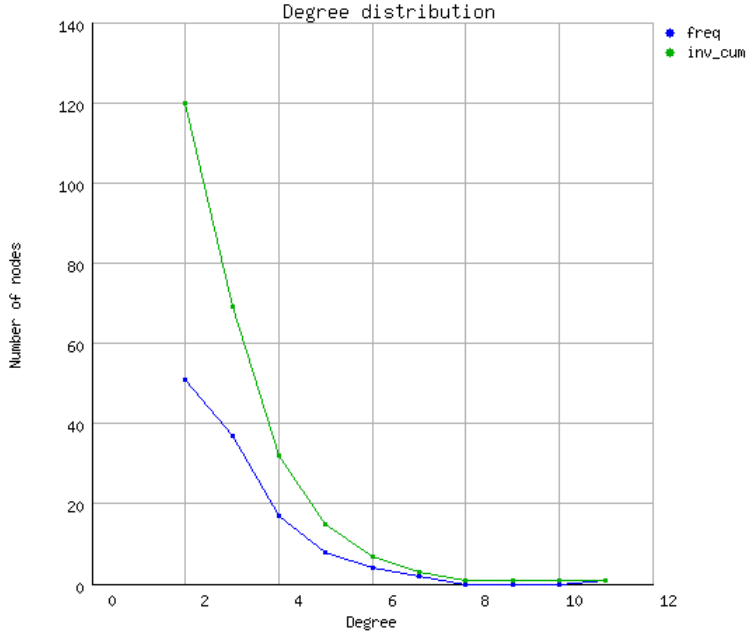


Supplemental Figure 2

Supplemental Figure 3. The node degree distribution of the normal tissues and solid cancers miRNA networks.

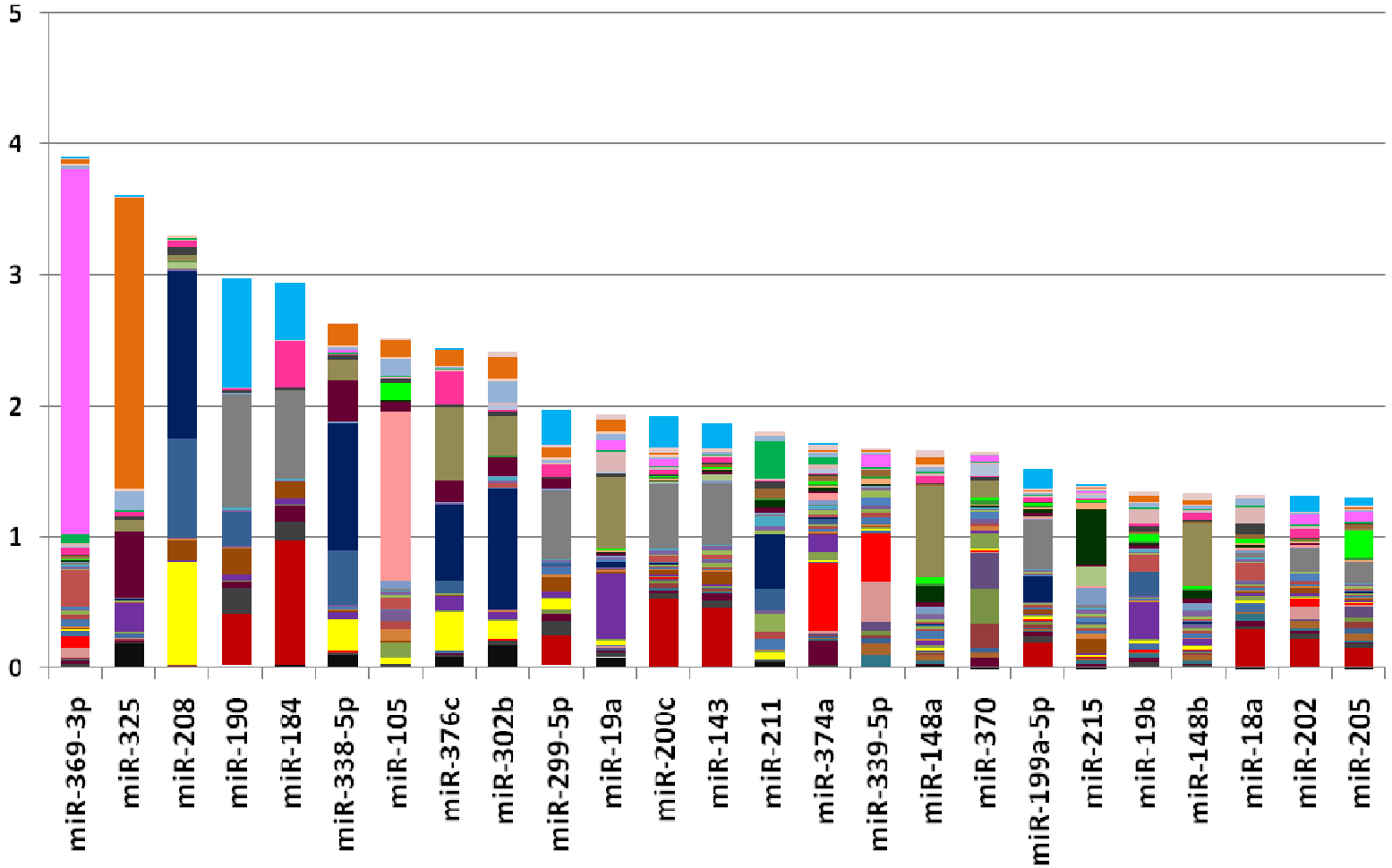


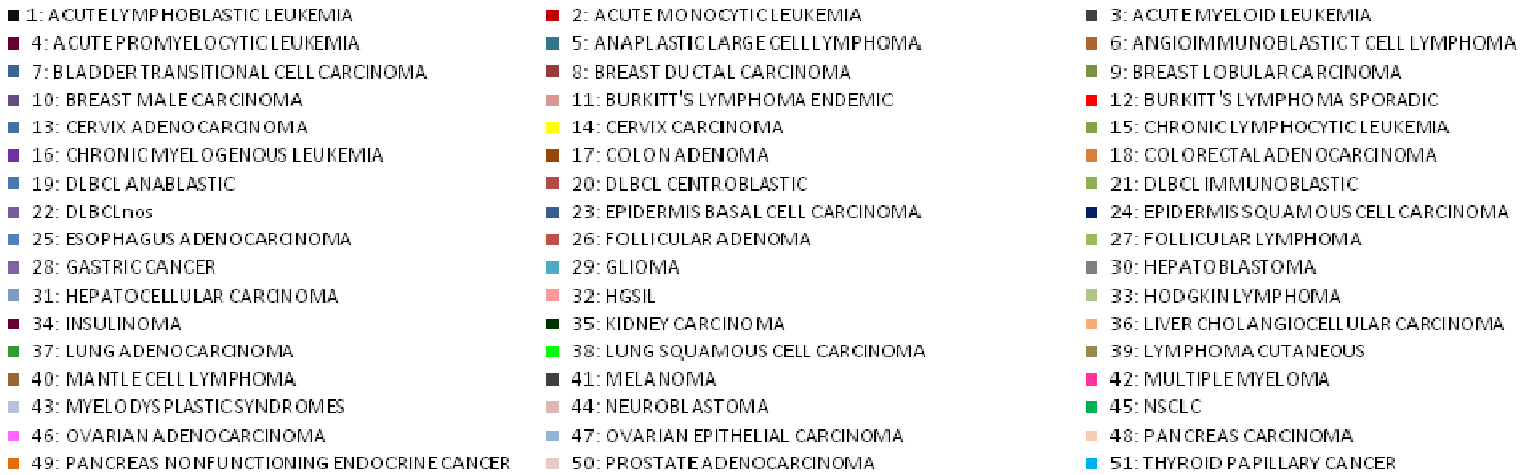
A



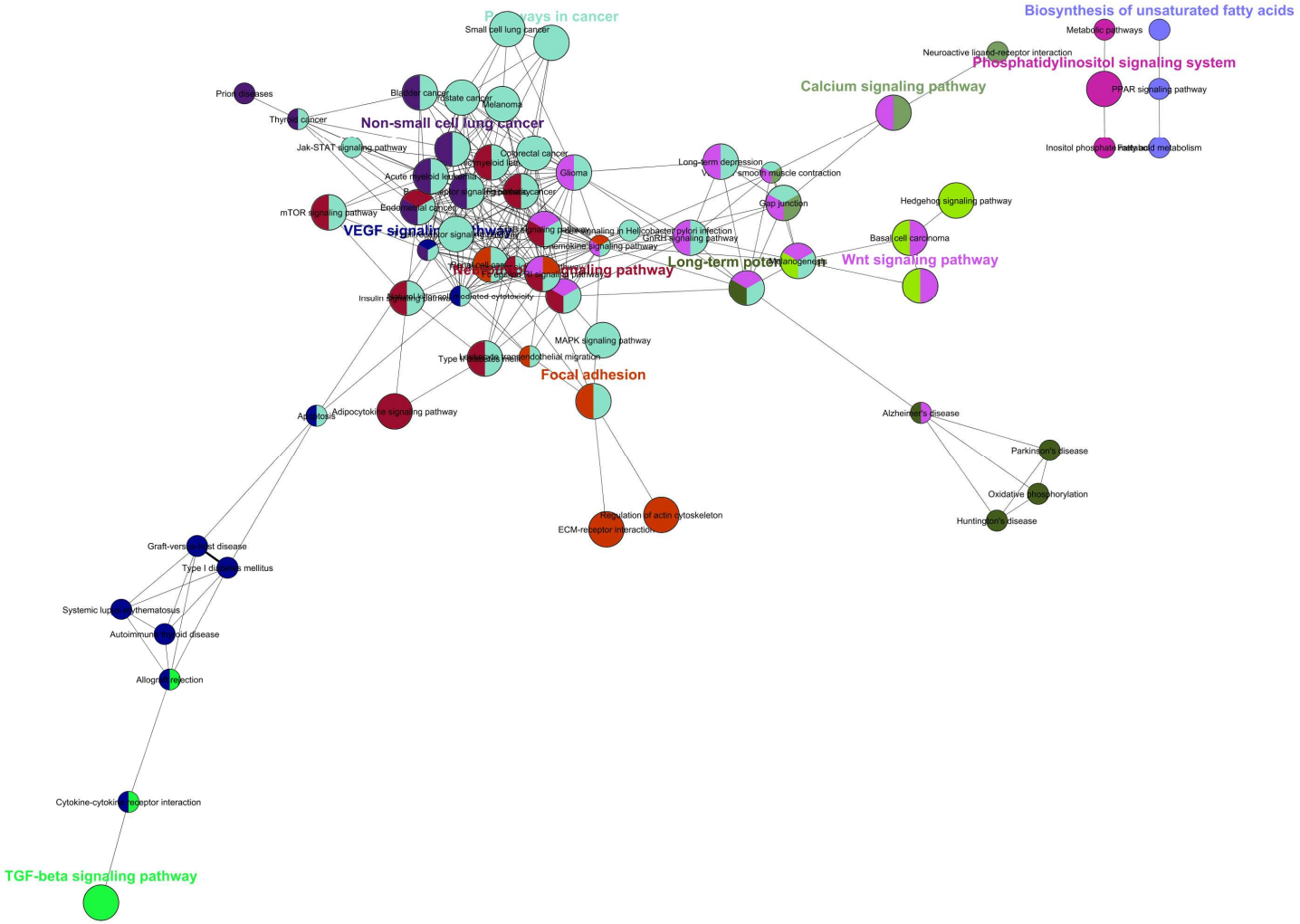
B

Supplemental Figure 4. miRNA specificity in tumors and leukemia (information content).

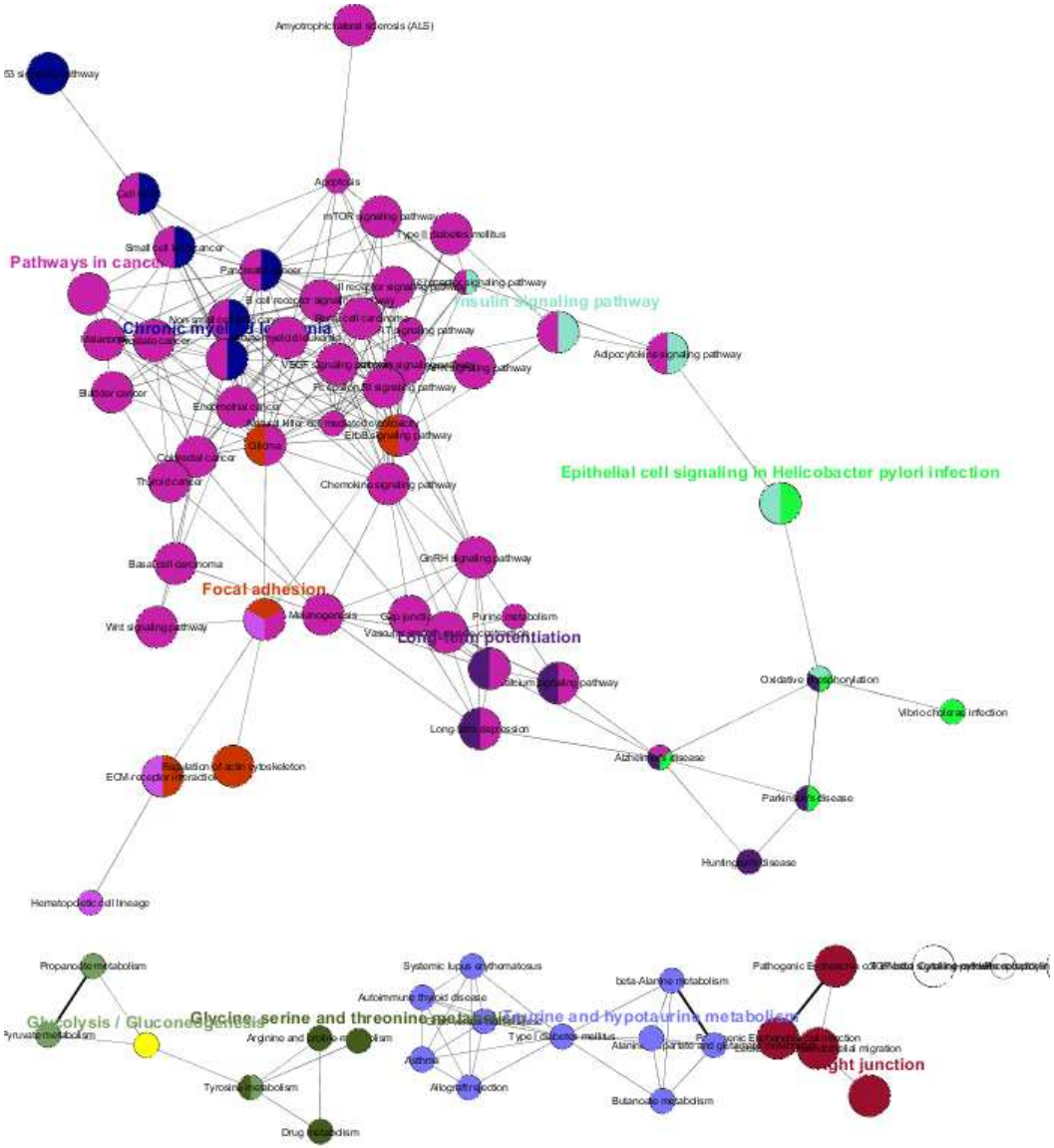



1: ACUTE LYMPHOBLASTIC LEUKEMIA
2: ACUTE MONOCYTIC LEUKEMIA
3: ACUTE MYELOID LEUKEMIA
4: ACUTE PROMYELOCYTIC LEUKEMIA
5: ANAPLASTIC LARGE CELL LYMPHOMA
6: ANGIOIMMUNOBLASTIC T CELL LYMPHOMA
7: BLADDER TRANSITIONAL CELL CARCINOMA
8: BREAST DUCTAL CARCINOMA
9: BREAST LOBULAR CARCINOMA
10: BREAST MALE CARCINOMA
11: BURKITT'S LYMPHOMA ENDEMIC
12: BURKITT'S LYMPHOMA SPORADIC
13: CERVIX ADENOCARCINOMA
14: CERVIX CARCINOMA
15: CHRONIC LYMPHOCYTIC LEUKEMIA
16: CHRONIC MYELOGENOUS LEUKEMIA
17: COLON ADENOMA
18: COLORECTAL ADENOCARCINOMA
19: DLBCL ANABLASTIC
20: DLBCL CENTROBLASTIC
21: DLBCL IMMUNOBLASTIC
22: DLBCLnos
23: EPIDERMIS BASAL CELL CARCINOMA
24: EPIDERMIS SQUAMOUS CELL CARCINOMA
25: ESOPHAGUS ADENOCARCINOMA
26: FOLLICULAR ADENOMA
27: FOLLICULAR LYMPHOMA
28: GASTRIC CANCER
29: GLIOMA
30: HEPATOBLASTOMA
31: HEPATOCELLULAR CARCINOMA
32: HGSIL
33: HODGKIN LYMPHOMA
34: INSULINOMA
35: KIDNEY CARCINOMA
36: LIVER CHOLANGIOCELLULAR CARCINOMA
37: LUNG ADENOCARCINOMA
38: LUNG SQUAMOUS CELL CARCINOMA
39: LYMPHOMA CUTANEOUS
40: MANTLE CELL LYMPHOMA
41: MELANOMA
42: MULTIPLE MYELOMA
43: MYELODYSPLASTIC SYNDROMES
44: NEUROBLASTOMA
45: NSCLC
46: OVARIAN ADENOCARCINOMA
47: OVARIAN EPITHELIAL CARCINOMA
48: PANCREAS CARCINOMA
49: PANCREAS NONFUNCTIONING ENDOCRINE CANCER
50: PROSTATE ADENOCARCINOMA
51: THYROID PAPILLARY CANCER

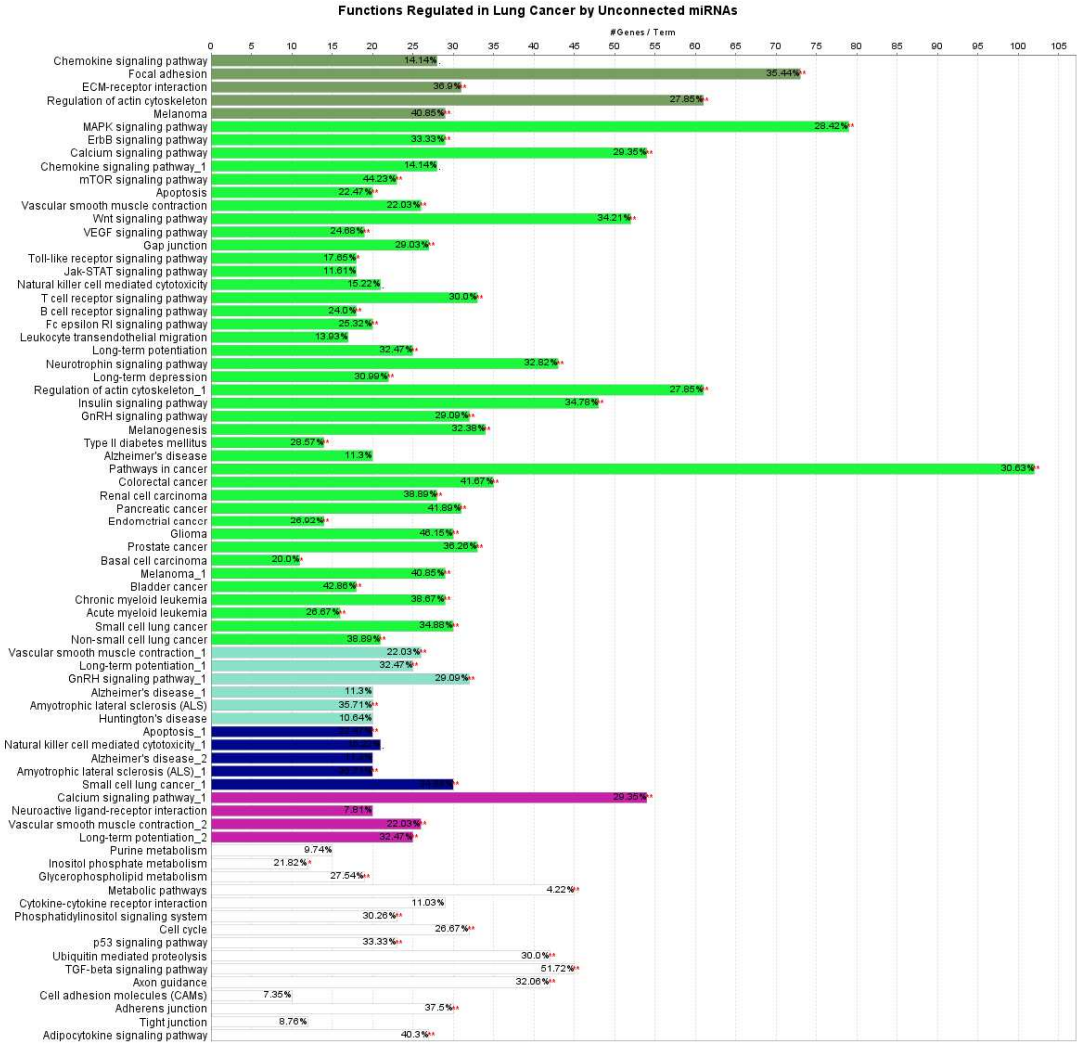
Supplemental Figure 5. Functions repressed in cancer by up-regulated miRNAs.



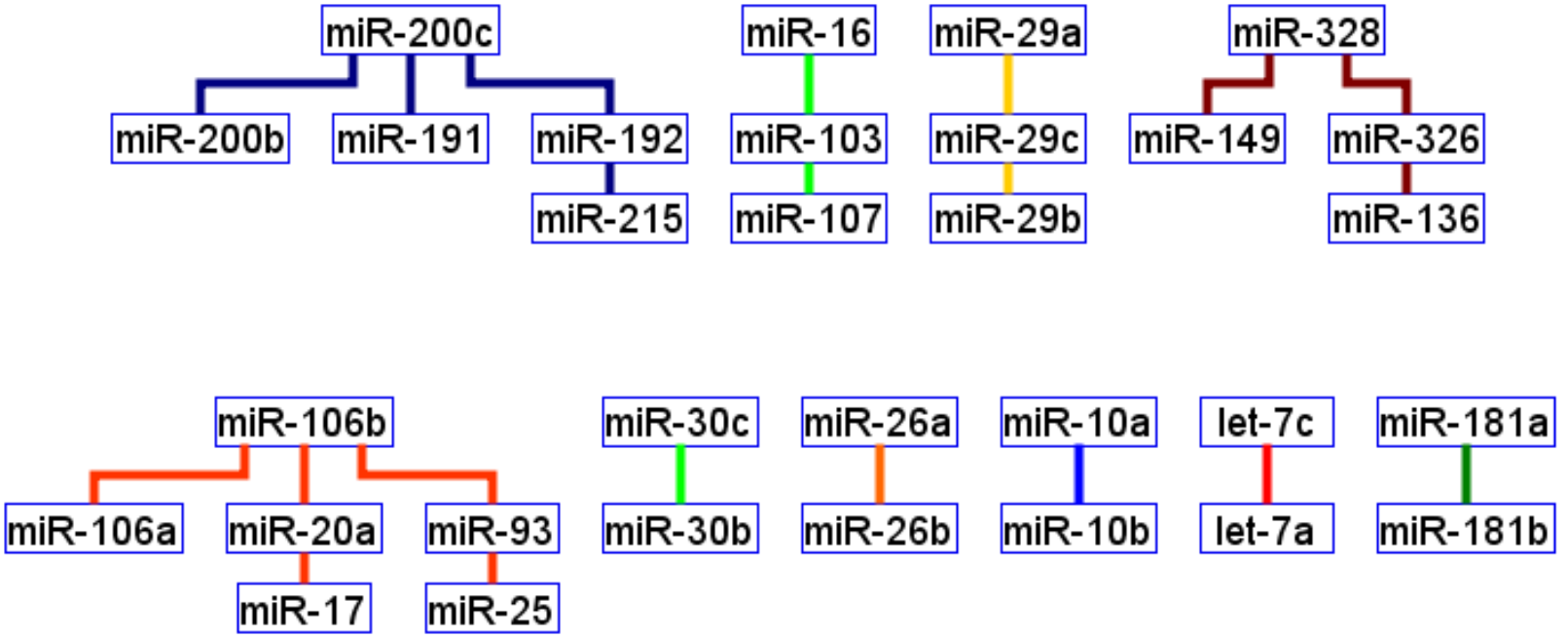
Supplemental Figure 6. Functions activated in cancer by down-regulated miRNAs.



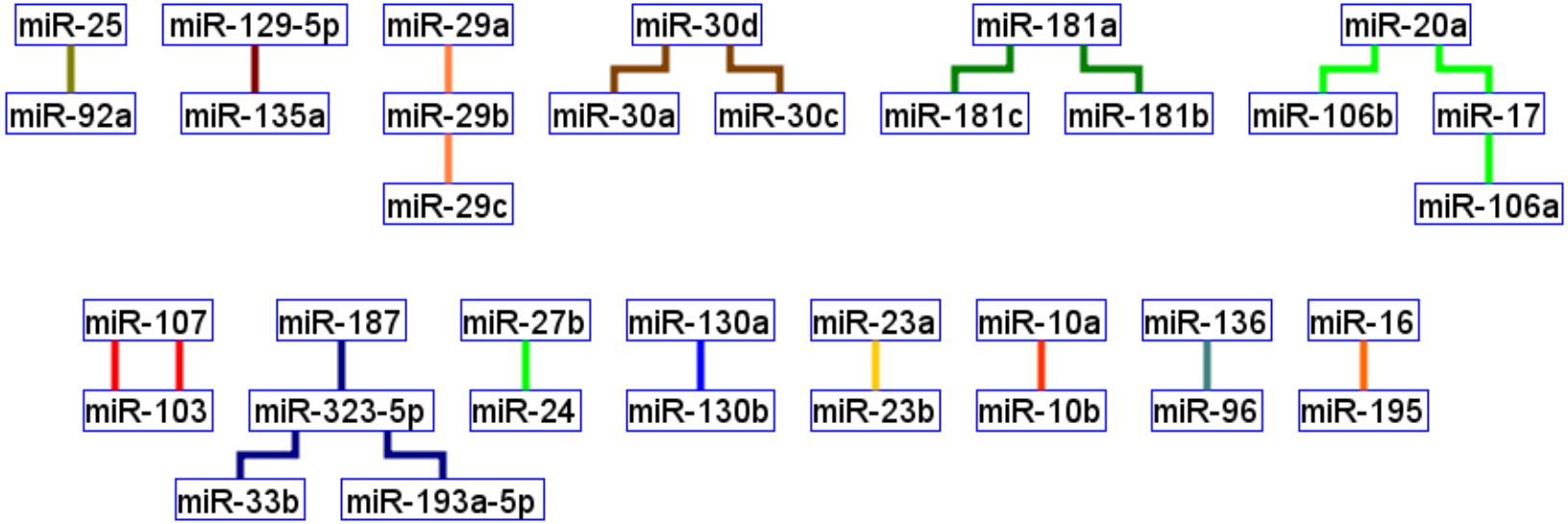
Supplemental Figure 7. Functions controlled in lung adenocarcinoma by miRNAs from disjointed minor clusters.



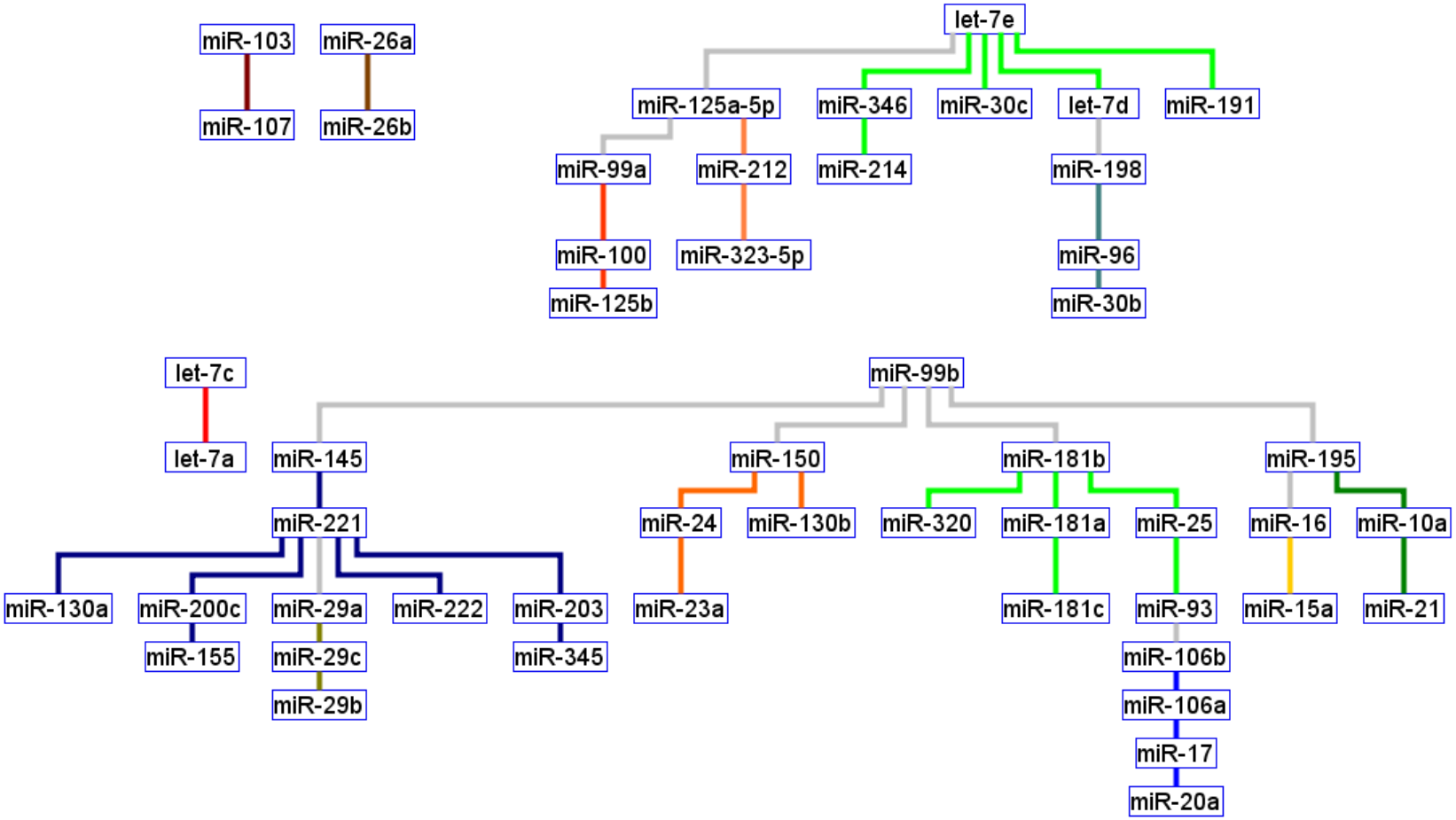
Supplemental Figure 8. MicroRNA genetic network in colon adenocarcinoma.



Supplemental Figure 9. MicroRNA genetic network in breast cancer.



Supplemental Figure 10. MicroRNA genetic network in prostate cancer.



Supplemental References

Ambs, S., Prueitt, R.L., Yi, M., Hudson, R.S., Howe, T.M., Petrocca, F., Wallace, T.A., Liu, C.G., Volinia, S., Calin, G.A. et al. 2008. Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. *Cancer research* **68:** 6162-6170.

Baffa, R., Fassan, M., Volinia, S., O'Hara, B., Liu, C.G., Palazzo, J.P., Gardiman, M., Rugge, M., Gomella, L.G., Croce, C.M. et al. 2009. MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets. *The Journal of pathology* **219:** 214-221.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pages, F., Trajanoski, Z., and Galon, J. 2009. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics (Oxford, England)* **25:** 1091-1093.

Bloomston, M., Frankel, W.L., Petrocca, F., Volinia, S., Alder, H., Hagan, J.P., Liu, C.G., Bhatt, D., Taccioli, C., and Croce, C.M. 2007. MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis. *Jama* **297:** 1901-1908.

Budhu, A., Jia, H.L., Forgues, M., Liu, C.G., Goldstein, D., Lam, A., Zanetti, K.A., Ye, Q.H., Qin, L.X., Croce, C.M. et al. 2008. Identification of metastasis-related microRNAs in hepatocellular carcinoma. *Hepatology (Baltimore, Md* **47:** 897-907.

Calin, G.A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S.E., Iorio, M.V., Visone, R., Sever, N.I., Fabbri, M. et al. 2005. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *The New England journal of medicine* **353:** 1793-1801.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J. et al. 2009. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell stem cell* **5:** 111-123.

Ciafre, S.A., Galardi, S., Mangiola, A., Ferracin, M., Liu, C.G., Sabatino, G., Negrini, M., Maira, G., Croce, C.M., and Farace, M.G. 2005. Extensive modulation of a set of microRNAs in primary glioblastoma. *Biochemical and biophysical research communications* **334:** 1351-1358.

Fassan, M., Baffa, R., Palazzo, J.P., Lloyd, J., Crosariol, M., Liu, C.G., Volinia, S., Alder, H., Rugge, M., Croce, C.M. et al. 2009. MicroRNA expression profiling of male breast cancer. *Breast Cancer Res* **11:** R58.

Garzon, R., Garofalo, M., Martelli, M.P., Briesewitz, R., Wang, L., Fernandez-Cymering, C., Volinia, S., Liu, C.G., Schnittger, S., Haferlach, T. et al. 2008a. Distinctive microRNA signature of acute myeloid leukemia bearing cytoplasmic mutated nucleophosmin. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 3945-3950.

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. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 5078-5083.

Garzon, R., Pichiorri, F., Palumbo, T., Visentini, M., Aqeilan, R., Cimmino, A., Wang, H., Sun, H., Volinia, S., Alder, H. et al. 2007. MicroRNA gene expression during retinoic acid-induced differentiation of human acute promyelocytic leukemia. *Oncogene* **26:** 4148-4157.

Garzon, R., Volinia, S., Liu, C.G., Fernandez-Cymering, C., Palumbo, T., Pichiorri, F., Fabbri, M., Coombes, K., Alder, H., Nakamura, T. et al. 2008b. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia. *Blood* **111:** 3183-3189.

Godlewski, J., Nowicki, M.O., Bronisz, A., Williams, S., Otsuki, A., Nuovo, G., Raychaudhury, A., Newton, H.B., Chiocca, E.A., and Lawler, S. 2008. Targeting of the Bmi-1 oncogene/stem cell renewal factor by microRNA-128 inhibits glioma proliferation and self-renewal. *Cancer research* **68:** 9125-9130.

He, H., Jazdzewski, K., Li, W., Liyanarachchi, S., Nagy, R., Volinia, S., Calin, G.A., Liu, C.G., Franssila, K., Suster, S. et al. 2005. The role of microRNA genes in papillary thyroid carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 19075-19080.

Iorio, M.V., Ferracin, M., Liu, C.G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M. et al. 2005. MicroRNA gene expression deregulation in human breast cancer. *Cancer research* **65:** 7065-7070.

Iorio, M.V., Visone, R., Di Leva, G., Donati, V., Petrocca, F., Casalini, P., Taccioli, C., Volinia, S., Liu, C.G., Alder, H. et al. 2007. MicroRNA signatures in human ovarian cancer. *Cancer research* **67:** 8699-8707.

Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M. et al. 2007. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129:** 1401-1414.

Papadopoulos, G.L., Alexiou, P., Maragkakis, M., Reczko, M., and Hatzigeorgiou, A.G. 2009. DIANA-mirPath: Integrating human and mouse microRNAs in pathways. *Bioinformatics (Oxford, England)* **25:** 1991-1993.

Petrocca, F., Visone, R., Onelli, M.R., Shah, M.H., Nicoloso, M.S., de Martino, I., Iliopoulos, D., Pilozzi, E., Liu, C.G., Negrini, M. et al. 2008. E2F1-regulated microRNAs impair TGFbeta-dependent cell-cycle arrest and apoptosis in gastric cancer. *Cancer cell* **13:** 272-286.

Pichiorri, F., Suh, S.S., Ladetto, M., Kuehl, M., Palumbo, T., Drandi, D., Taccioli, C., Zanesi, N., Alder, H., Hagan, J.P. et al. 2008. MicroRNAs regulate critical genes associated with multiple myeloma pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 12885-12890.

Pineau, P., Volinia, S., McJunkin, K., Marchio, A., Battiston, C., Terris, B., Mazzaferro, V., Lowe, S.W., Croce, C.M., and Dejean, A. miR-221 overexpression contributes to liver tumorigenesis. *Proceedings of the National Academy of Sciences of the United States of America* **107:** 264-269.

Roldo, C., Missiaglia, E., Hagan, J.P., Falconi, M., Capelli, P., Bersani, S., Calin, G.A., Volinia, S., Liu, C.G., Scarpa, A. et al. 2006. MicroRNA expression abnormalities in pancreatic endocrine and acinar tumors are associated with distinctive pathologic features and clinical behavior. *J Clin Oncol* **24:** 4677-4684.

Schetter, A.J., Leung, S.Y., Sohn, J.J., Zanetti, K.A., Bowman, E.D., Yanaihara, N., Yuen, S.T., Chan, T.L., Kwong, D.L., Au, G.K. et al. 2008. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *Jama* **299:** 425-436.

Seike, M., Goto, A., Okano, T., Bowman, E.D., Schetter, A.J., Horikawa, I., Mathe, E.A., Jen, J., Yang, P., Sugimura, H. et al. 2009. MiR-21 is an EGFR-regulated anti-apoptotic factor in lung cancer in never-smokers. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 12085-12090.

Ueda, T., Volinia, S., Okumura, H., Shimizu, M., Taccioli, C., Rossi, S., Alder, H., Liu, C.G., Oue, N., Yasui, W. et al. Relation between microRNA expression and progression and prognosis of gastric cancer: a microRNA expression analysis. *Lancet Oncol* **11:** 136-146.

Visone, R., Rassenti, L.Z., Veronese, A., Taccioli, C., Costinean, S., Aguda, B.D., Volinia, S., Ferracin, M., Palatini, J., Balatti, V. et al. 2009. Karyotype-specific microRNA signature in chronic lymphocytic leukemia. *Blood* **114:** 3872-3879.

Volinia, S., Calin, G.A., Liu, C.G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M. et al. 2006. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 2257-2261.

Yanaihara, N., Caplen, N., Bowman, E., Seike, M., Kumamoto, K., Yi, M., Stephens, R.M., Okamoto, A., Yokota, J., Tanaka, T. et al. 2006. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer cell* **9:** 189-198.

Zhang, L., Volinia, S., Bonome, T., Calin, G.A., Greshock, J., Yang, N., Liu, C.G., Giannakakis, A., Alexiou, P., Hasegawa, K. et al. 2008. Genomic and epigenetic alterations deregulate microRNA expression in human epithelial ovarian cancer. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 7004-7009.