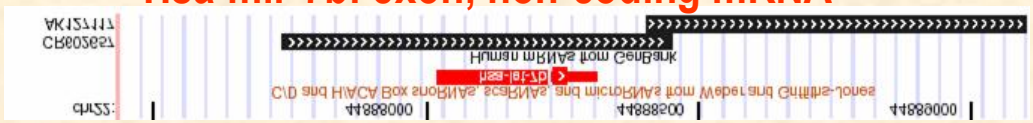
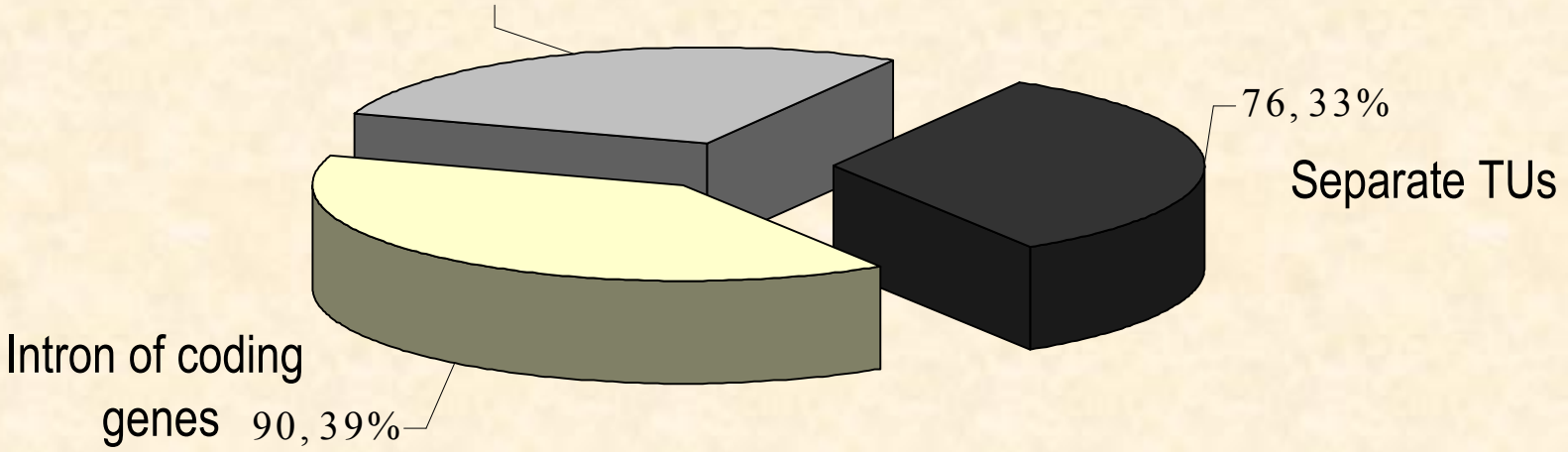
Distribution of 232 miRNAs in known transcription units (TUs)

Hsa-mir-7b: exon, non-coding mRNA



Hsa-mir-7d: intron, non-coding mRNA



Hsa-mir-126: intron, coding mRNA

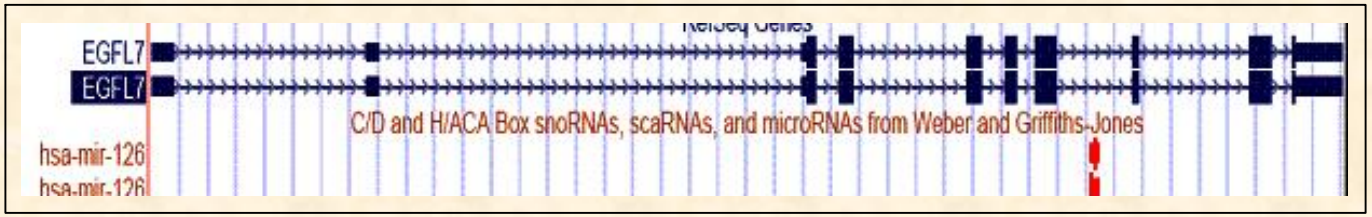
GENOME RESEARCH

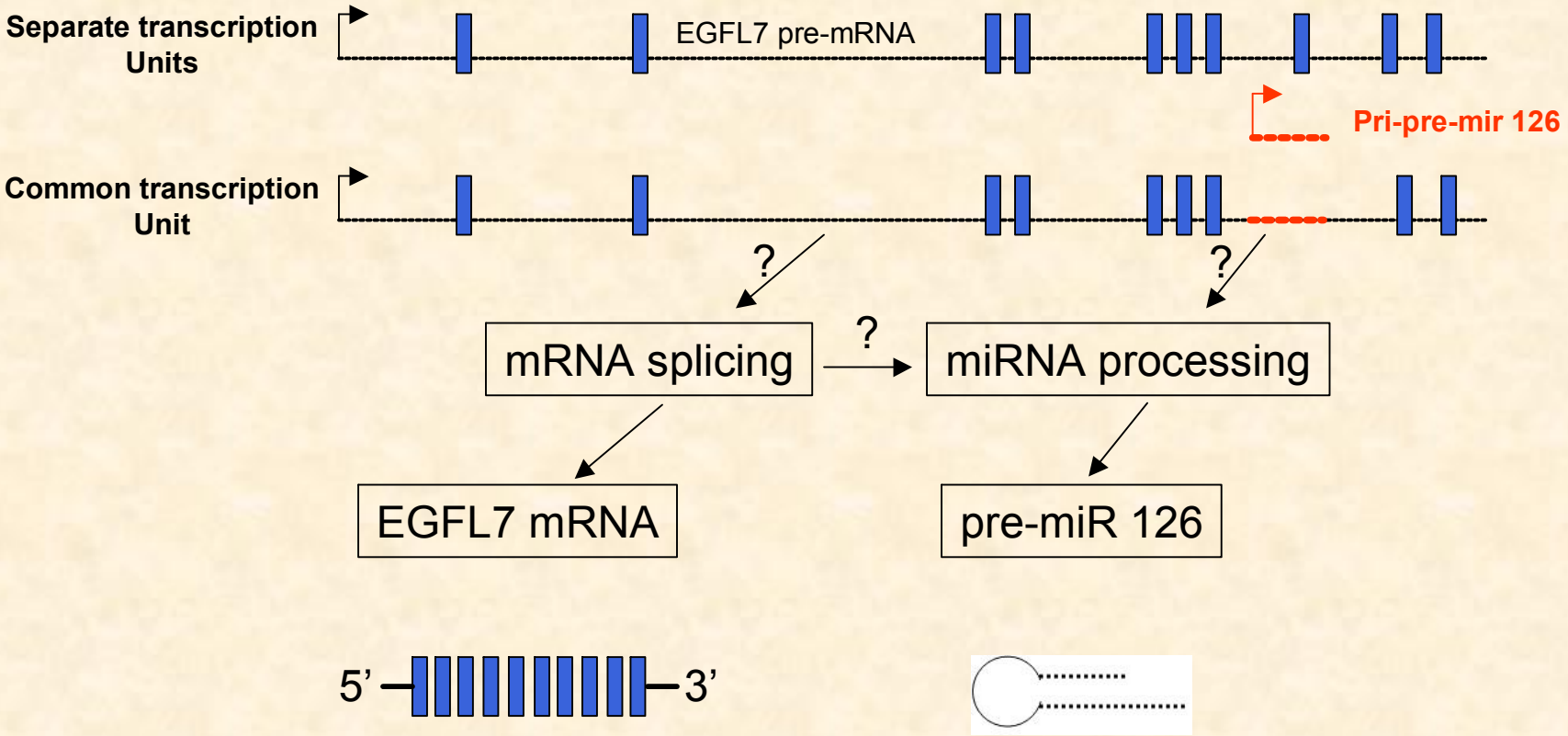
Identification of Mammalian microRNA Host Genes and Transcription Units

Antony Rodriguez, Sam Griffiths-Jones, Jennifer L. Ashurst and Allan Bradley

Genome Res. 2004 14: 1902-1910; originally published online Sep 13, 2004; doi:10.1101/gr.2722704

Possible processing steps of intronic miRNAs from a common primary host gene transcript





Host gene and intronic miRNA expression in leukemic cell lines

Background

- correlation between host genes and intronic miRNA expression has been reported for reference normal tissue

Methods and Results

- total RNA from NB4 (acute promyelocytic) and AML2 (acute myeloid) leukemia cell lines was hybridized to U133A Affymetrix gene chips as well as to a miRNA microarray
- Host genes and intronic miRNA expression was measured and compared between the two platforms for each cell line

Intronic miRNAs and host gene co-expression in reference normal tissue

TABLE 1. Correlation between expression of intronic miRNAs and host gene

Gene name	Ensembl ID[a]	MicroRNA	Host gene corr.	Upstream corr.	Downstream corr.
CROC4	125462	hsa-miR-9-1	0.999	0.083	–0.054
PDE2A	186642	hsa-miR-139	0.99	–0.293	0.493
C20orf166	174407	hsa-miR-133a-2	0.988	–0.105	0.185
		hsa-miR-1-1	0.968	–0.045	0.226
PGSF1	176840	hsa-miR-7-3	0.961	–0.394	–0.123
ABLIM2	163995	hsa-miR-95	0.96	–0.302	–0.089
LOC254559	LOC254559	hsa-miR-9-3	0.95	–0.122	–0.150
AATYK	181409	hsa-miR-338	0.921	0.906	–0.123
EGFL7	172889	hsa-miR-126	0.888	0.727	–0.246
R3HDM	48991	hsa-miR-128a	0.856	–0.037	–0.363
MCM7[b]	166508	hsa-miR-25	0.838	–0.548	–0.377

TABLE 1. Correlation between expression of intronic miRNAs and host gene

Gene name	Ensembl ID[a]	MicroRNA	Host gene corr.	Upstream corr.	Downstream corr.
TRPM3	83067	hsa-miR-204	0.796	–0.047	–0.240
TLN2	171914	hsa-miR-190	0.663	–0.269	–0.172
PANK3	120137	hsa-miR-103-1	0.638	–0.203	0.079
PTPRN	54356	hsa-miR-153-1	0.626	0.579	–0.178
CTDSP2	175215	hsa-miR-26a-2	0.609	0.492	0.265
SMC4L1	113810	hsa-miR-15b	0.509	–0.134	–0.098
		hsa-miR-16-2	0.504	–0.128	–0.021
UREB1	86758	hsa-miR-98	0.503	–0.209	–0.512
		hsa-let-7f-2	0.379	–0.242	–0.194

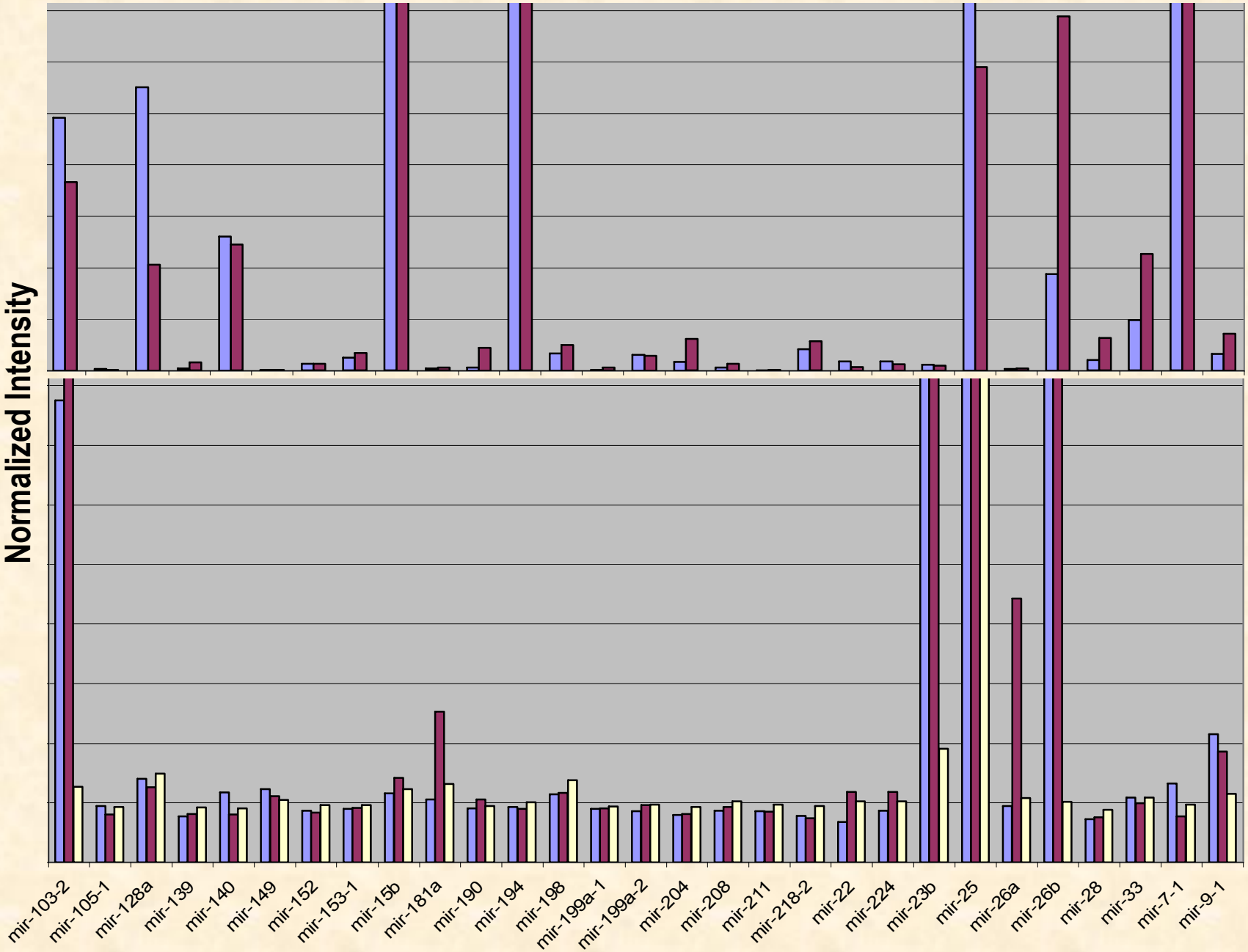
RNA

Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes

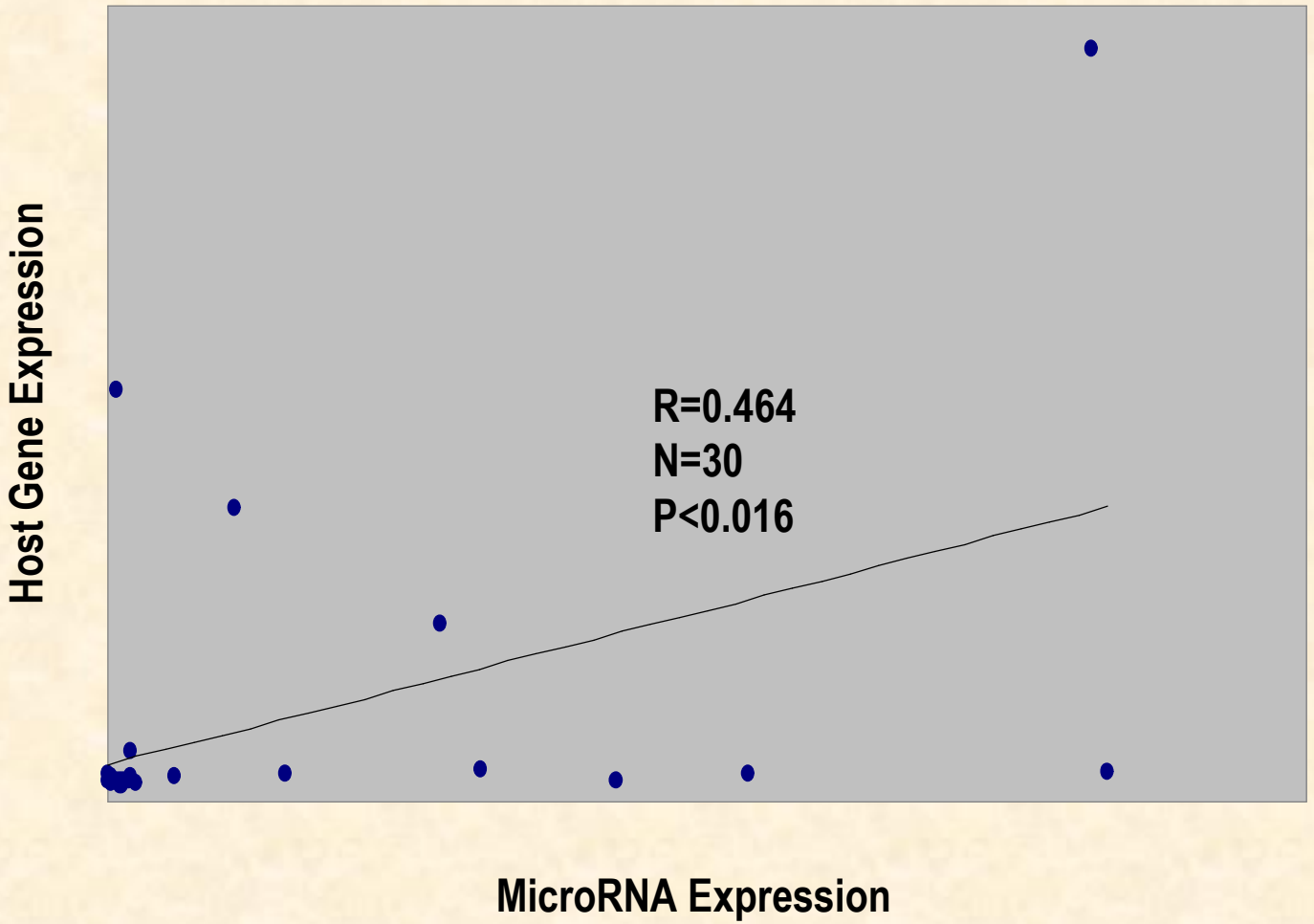
SCOTT BASKERVILLE and DAVID P. BARTEL

RNA 2005 11: 241-247
doi:10.1261/rna.7240905

Co-expression of host genes (top) and intronic miRNAs (bottom)



Correlation between host gene and intronic miRNA expression in NB4 cells



Summary of findings:co-expression of host genes and intronic miRNAs in cell lines

miRNA	Host	Chr	LOCATION	ENSEMBLE
hsa-mir-103-2	PANK2	20	3893141-3893218	ENSG00000125779
hsa-mir-128a	R3HDM1	2	136633736-136633817	ENSG00000048991
hsa-mir-15b	SMC4L1	3	161443289-161443386	ENSG00000113810
hsa-mir-25	MCM7	7	99302663-99302580	ENSG00000166508
hsa-mir-26b	CTDSP1	2	219469909-219469985	ENSG00000144579

Up-regulation of host genes and intronic miRNAs by aberrantly-expressed transcription factors

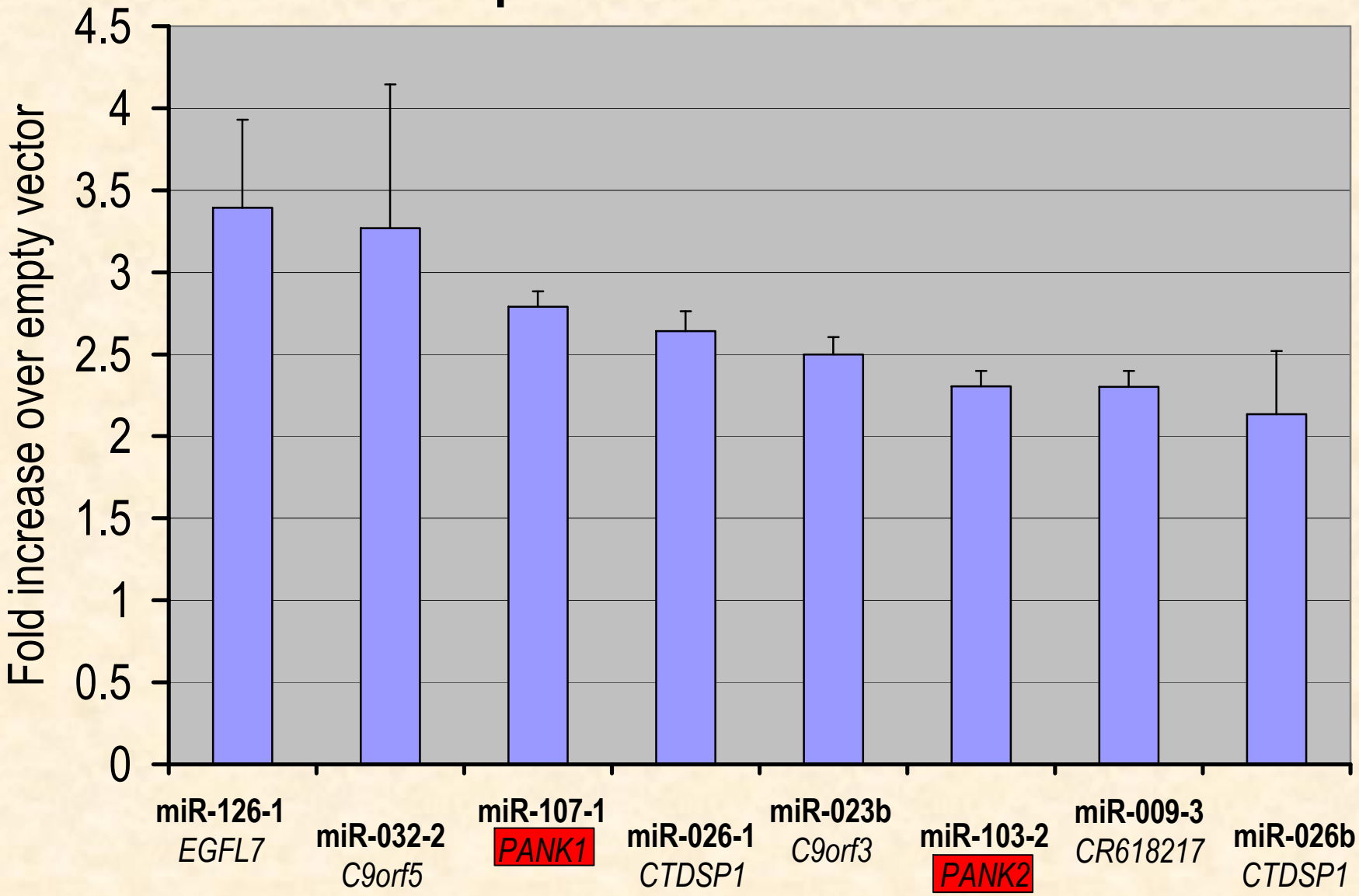
Background

- De-regulation of miRNA expression could be caused by copy number polymorphisms (CNP) or abnormal expression of transcription factors (TF) (e.g. *c-myc*)
- Lyl1 is a basic helix-loop-helix TF and suspected oncoprotein that is over-expressed in AML and T-ALL
- Lyl1 Binds to Ebox promoter DNA
- Lyl1 over-expression is associated with poor prognosis

Methods and Results

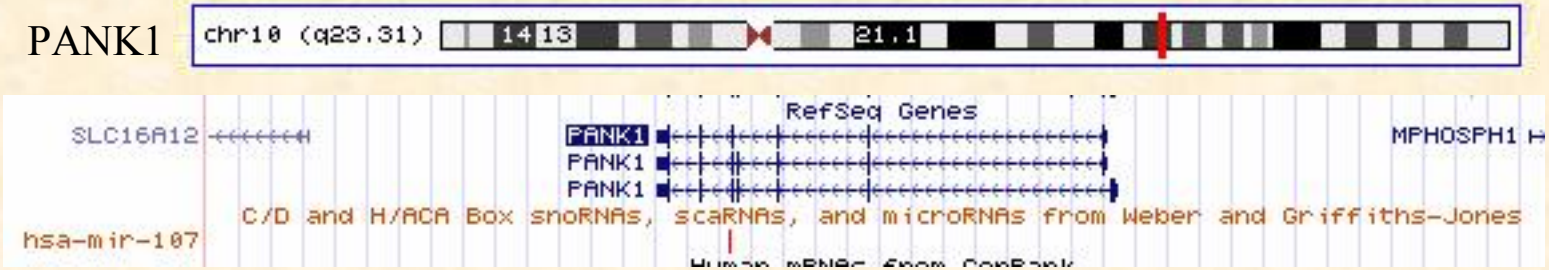
- HA-Lyl1 or empty vector expression was enforced in K562 cells
- Host gene and miRNA expression were determined by microarrays

Over-expressed Lyl1 up-regulates intronic miRNA expression in K562 cells

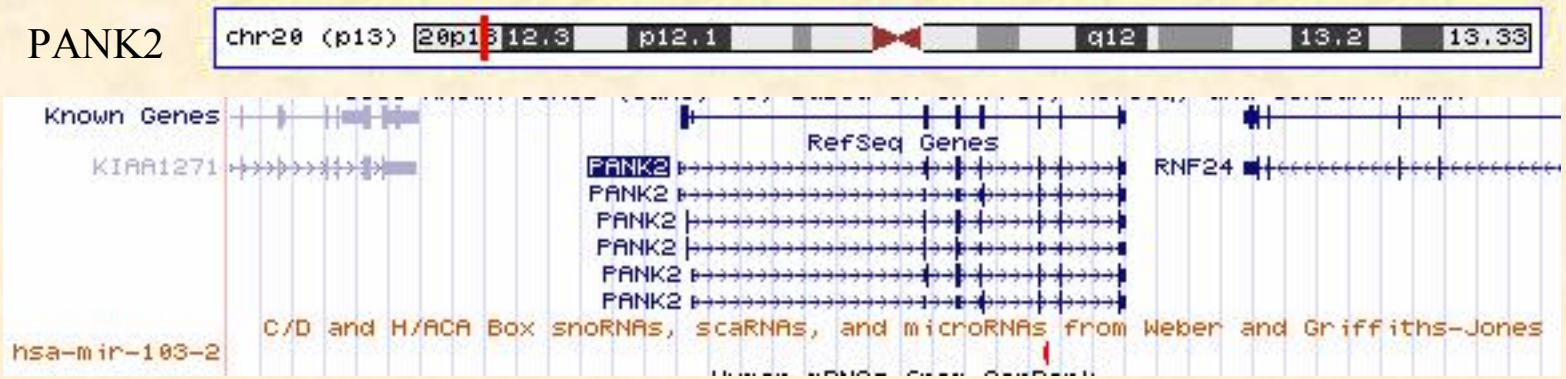


Genomic features of PANK1 and PANK2 host genes and mirs-107 and 103-2 up-regulated by Lyl1

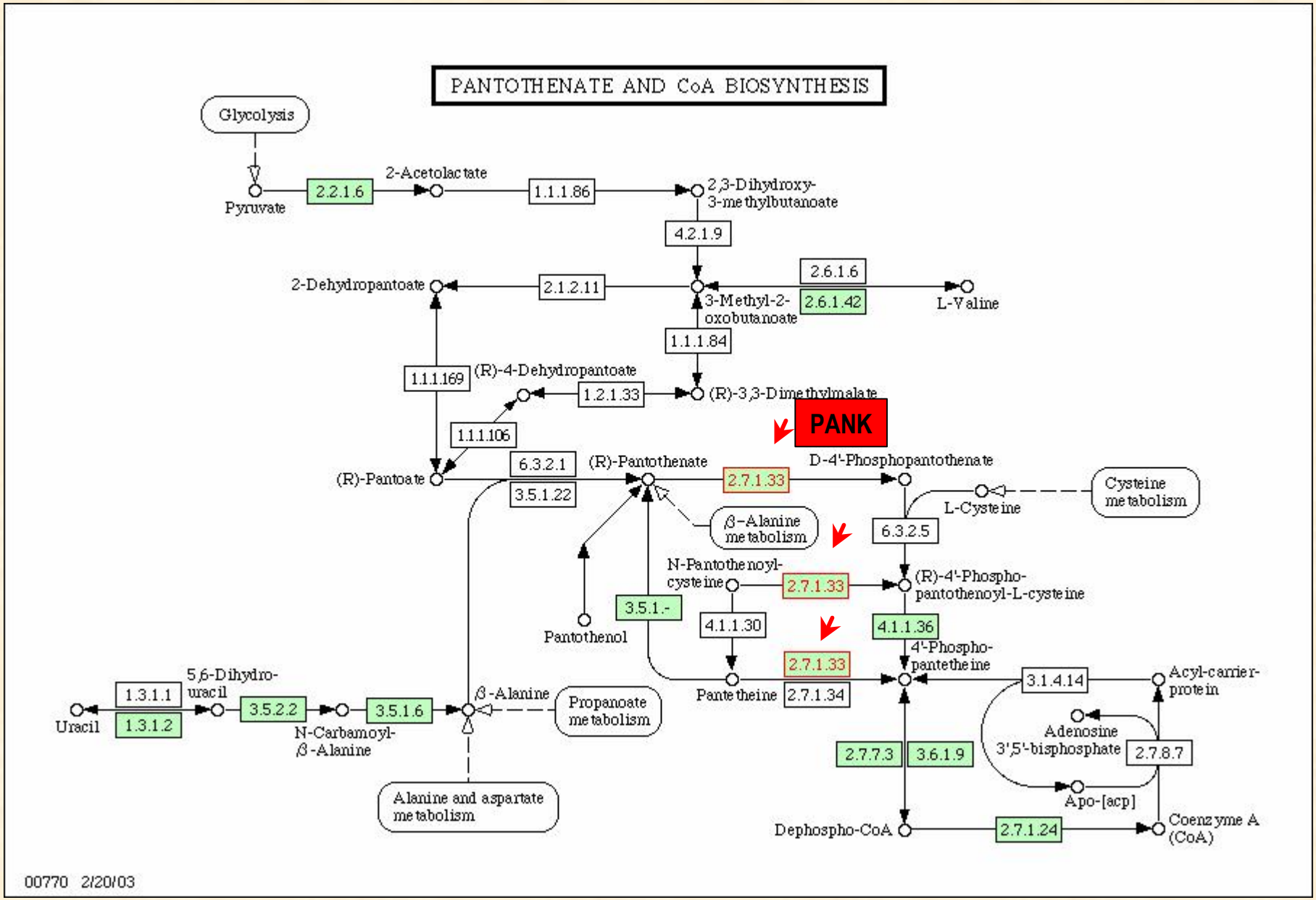
PANK1



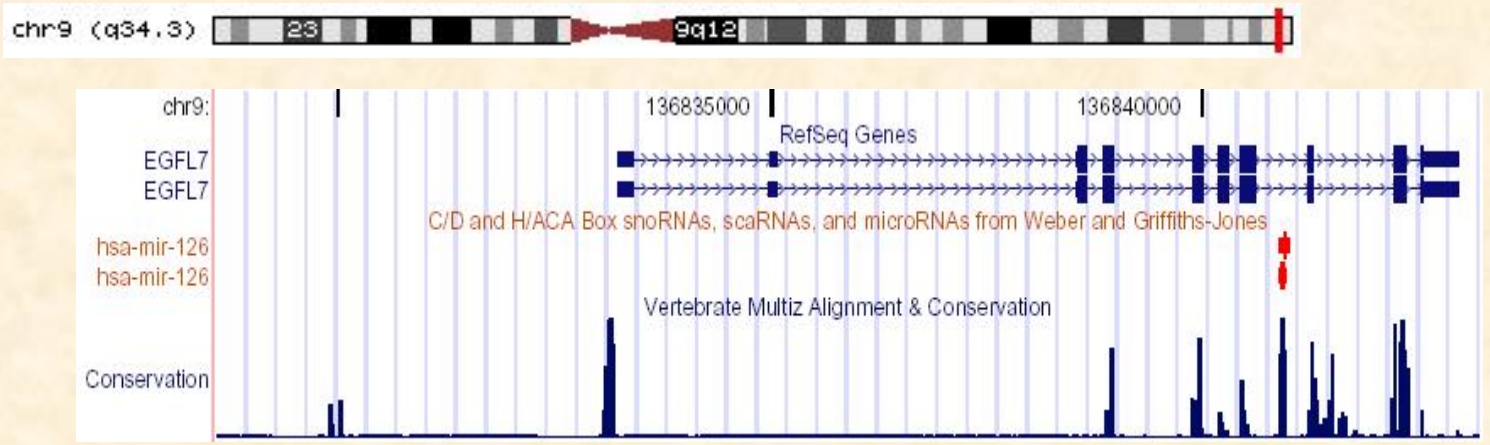
PANK2



• PANK2 deficiency caused by loss of enzymatic activity due to mutation is implicated in neurodegeneration with Brain Iron Accumulation Type1 Syndrome (Hallevorden-Spatz)

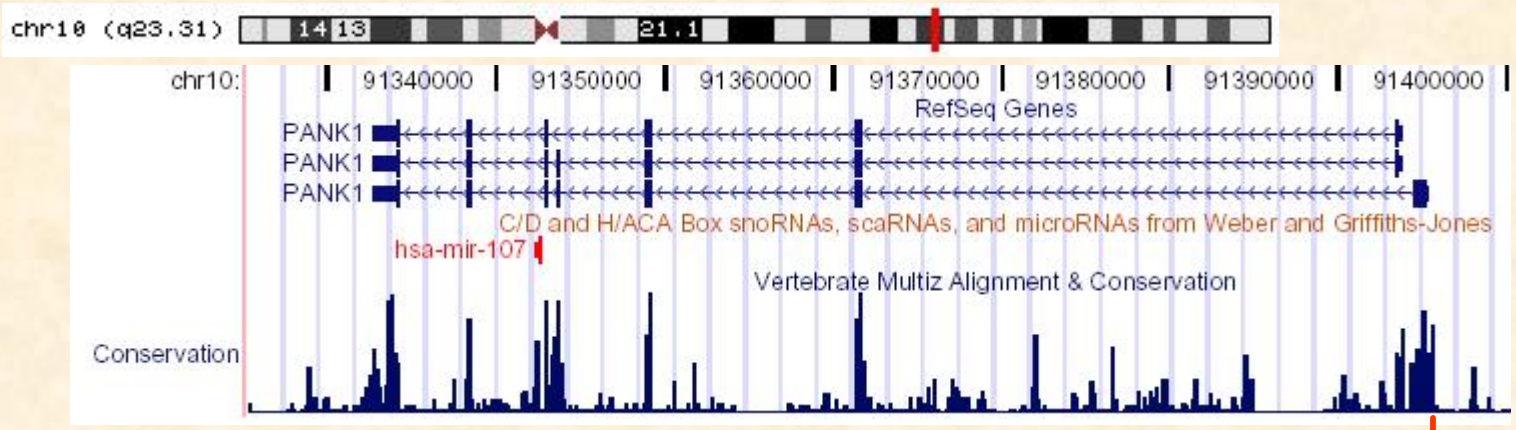

PANTOTHENATE AND CoA BIOSYNTHESIS
Glycolysis
Pyruvate
2.2.1.6
2-Acetolactate
1.1.1.86
2,3-Dihydroxy-3-methylbutanoate
4.2.1.9
2-Dehydropantoate
2.1.2.11
2.6.1.6
2.6.1.42
3-Methyl-2-oxobutanoate
L-Valine
1.1.1.84
1.1.1.169
(R)-4-Dehydropantoate
1.2.1.33
(R)-3,3-Dimethylmalate
PANK
1.1.1.106
6.3.2.1
3.5.1.22
(R)-Pantoate
(R)-Pantothenate
2.7.1.33
D-4'-Phosphopantothenate
Cysteine metabolism
L-Cysteine
β-Alanine metabolism
6.3.2.5
N-Pantothenoyl-cysteine
2.7.1.33
(R)-4'-Phospho-pantothenoyl-L-cysteine
3.5.1.-
Pantothenol
4.1.1.30
4.1.1.36
4'-Phospho-pantetheine
2.7.1.33
Pantetheine
2.7.1.34
3.1.4.14
Acyl-carrier-protein
1.3.1.1
1.3.1.2
5,6-Dihydro-uracil
3.5.2.2
3.5.1.6
β-Alanine
Propanoate metabolism
Uracil
N-Carbamoyl-β-Alanine
Adenosine 3',5'-bisphosphate
2.7.8.7
2.7.7.3
3.6.1.9
Apo-[acp]
Alanine and aspartate metabolism
Dephospho-CoA
2.7.1.24
Coenzyme A (CoA)
00770 2/20/03

Lyl1 binding sites in conserved EGFL7 and PANK1 proximal promoters of mir-126 and mir-107



Lyl1-E47 binding site

Pos	Strand	Core	Matrix	Sequence
7	(-)	0.886	0.852	cccatCACCTgctgcc



Lyl1-E47 binding site

Pos	Strand	Core	Matrix	Sequence
17	(+)	1.000	0.925	cgctgcaTCTGGccca

Functional meta-analysis of intronic miRNAs

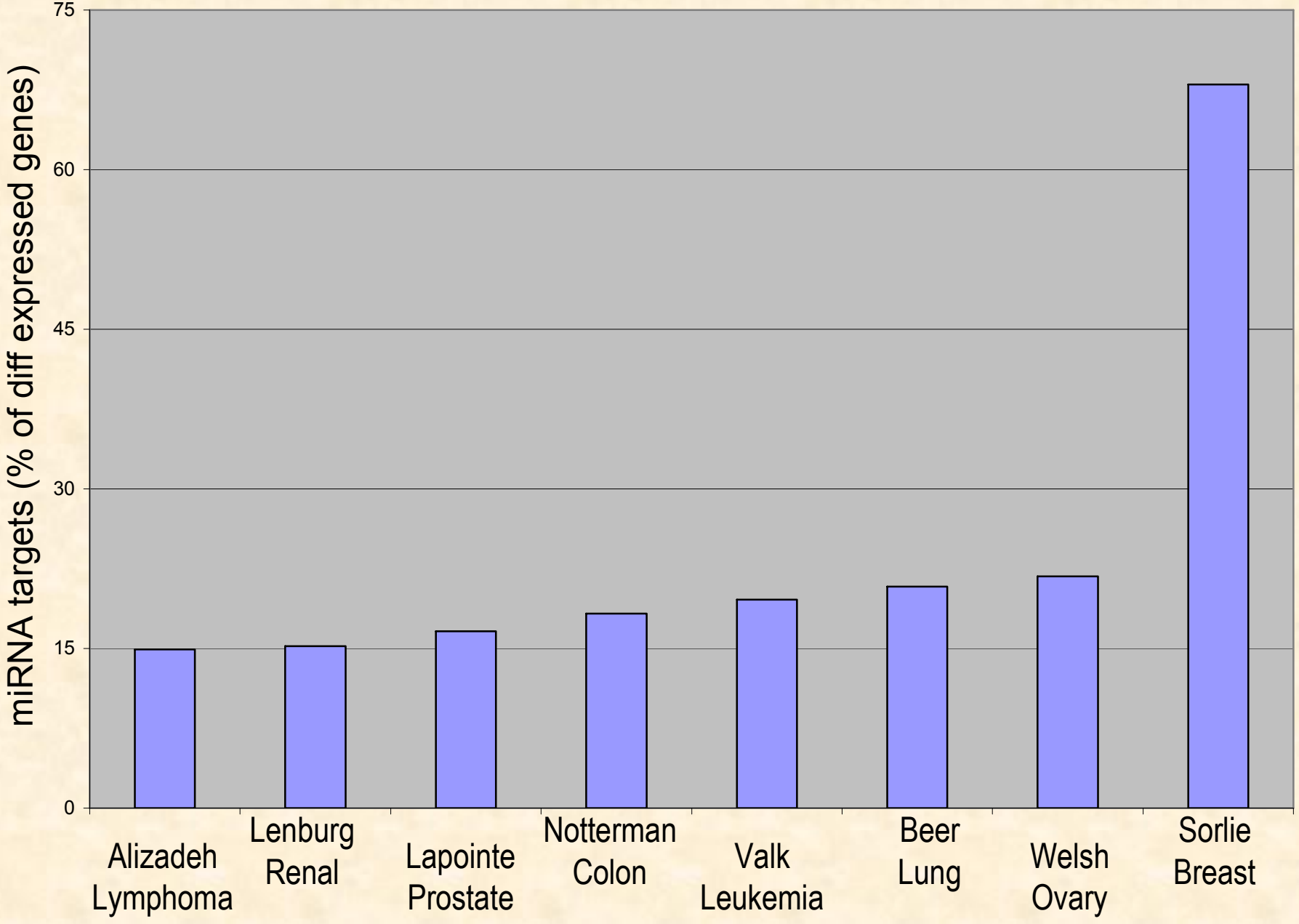
Background

- intronic miRNAs and host genes are differentially-expressed in tumors
- therefore the percentage of differentially expressed miRNA targets should also show tissue specificity
- if intronic miRNAs are co-ordinately expressed with their host genes the patterns of miRNA target gene expression should follow those of the miRNA host genes.

Methods and Results

- meta-analysis was performed using microarray data deposited at Oncomine.org
- host genes with statistically significant differential expression in tumor vs control (up and down) were identified
- differentially expressed genes in tumor sets relative to controls were filtered for 3'UTR enrichment in intronic miRNA recognition sequences
- molecular function analysis of target gene was performed using Gene Ontology at fatigo.org

Intronic miRNA targets are enriched in breast cancer



Intronic miRNA targets define tissue-specific molecular functions

1. Alizadeh_lymphoma compared to Sorlie_breast

Molecular function	Percent enrichment	Study	p value
receptor activity	8.52% / 16.11%	lymphoma breast	0.04056
transmembrane receptor activity	3.68% / 11.94%	lymphoma breast	0.00771
hydrolase activity, acting on ester bonds	8.59% / 2.24%	lymphoma breast	0.02283
phosphoric ester hydrolase activity	9.66% / 1.77%	lymphoma breast	0.00897
transmembrane receptor protein kinase activi...	0.69% / 6.19%	lymphoma breast	0.02313
phosphoric monoester hydrolase activity	11.82% / 2.47%	lymphoma breast	0.02655
transmembrane receptor protein kinase activi...	1.61% / 21.88%	lymphoma breast	0.00197
phosphoprotein phosphatase activity	20.97% / 3.12%	lymphoma breast	0.02971
transmembrane receptor protein tyrosine kina...	3.57% / 36.84%	lymphoma breast	0.00473

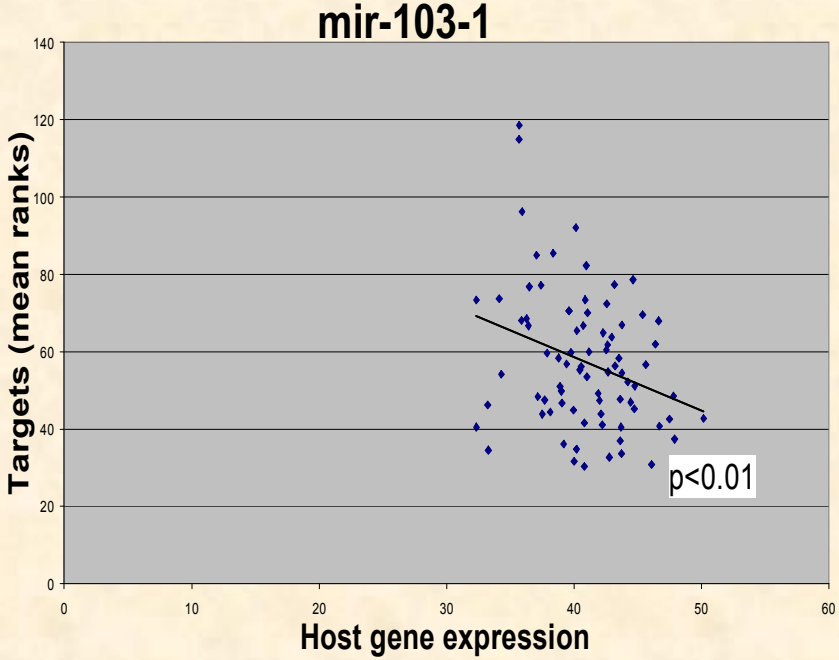
2. Valk_leukemia compared to Sorlie_breast

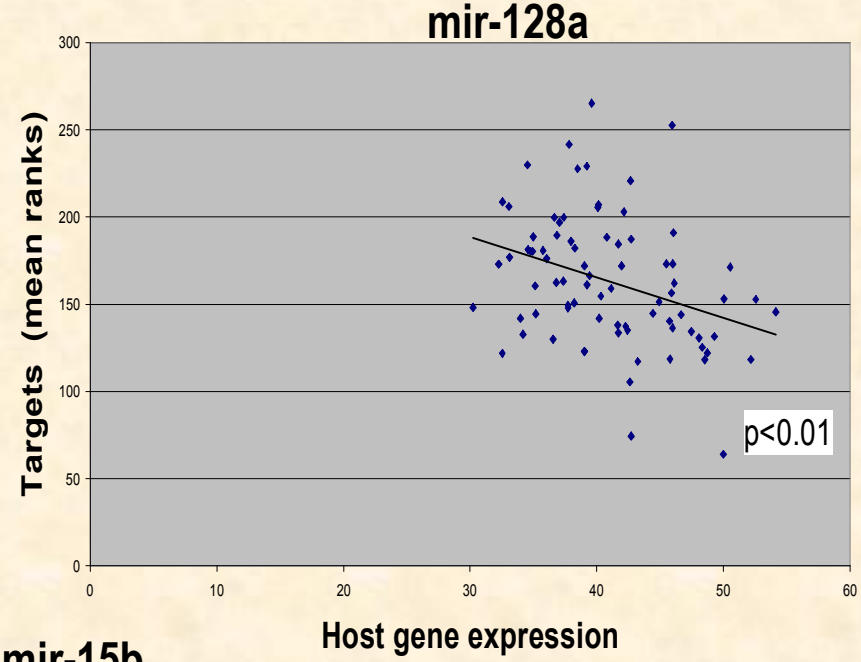
Molecular function	Percent enrichment	Study	p value
ligase activity	1.72% / 6%	leukemia / breast	0.01840
ligase activity, forming carbon-nitrogen bon...	1.92% / 6.62%	leukemia / breast	0.01895
acid-amino acid ligase activity	1.45% / 7.63%	leukemia / breast	0.00351
endopeptidase activity	0.73% / 5.08%	leukemia / breast	0.01061
ubiquitin-protein ligase activity	2.17% / 10.71%	leukemia / breast	0.00459
serine-type endopeptidase activity	0 % / 3.57%	leukemia / breast	0.03004
transmembrane receptor protein tyrosine kina...	10.77% / 33.33%	leukemia / breast	0.03557

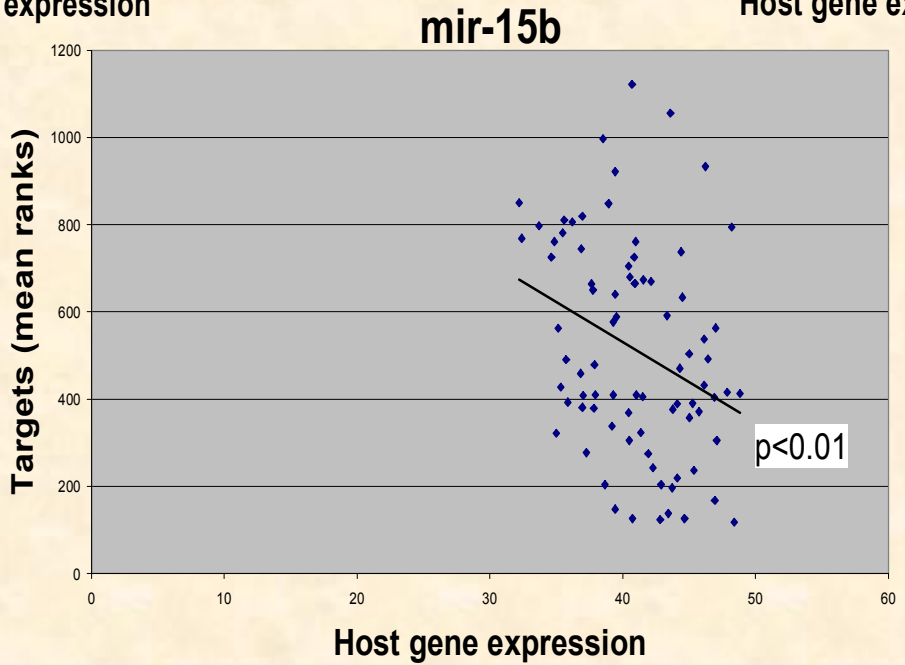
3. Valk_leukemia compared to Alizadeh_lymphoma

Molecular function	Percent enrichment	Study	p value
receptor activity	5.33%	leukemia	
	15.53%	lymphoma	0.00139
phosphatase regulator activity	2.67%	leukemia	
	0 %	lymphoma	0.00992
protein binding	51.33%	leukemia	
	40.99%	lymphoma	0.03715
receptor binding	6.67%	leukemia	
	2.48%	lymphoma	0.03743
transmembrane receptor activity	2.14%	leukemia	
	11.15%	lymphoma	0.00107
protein phosphatase regulator activity	2.86%	leukemia	
	0 %	lymphoma	0.01122
hydrolase activity, acting on ester bonds	10.71%	leukemia	
	4.88%	lymphoma	0.03841
receptor binding	7.14%	leukemia	
	2.79%	lymphoma	0.04235
protein phosphatase type 2A regulator activi...	3.15%	leukemia	
	0 %	lymphoma	0.01221
protein heterodimerization activity	3.15%	leukemia	
	0 %	lymphoma	0.01221
phosphoric ester hydrolase activity	11.81%	leukemia	
	5.16%	lymphoma	0.02310
phosphoprotein phosphatase activity	25.45%	leukemia	
	11.83%	lymphoma	0.04148

Host gene and miRNA target correlation in Valk_leukemia

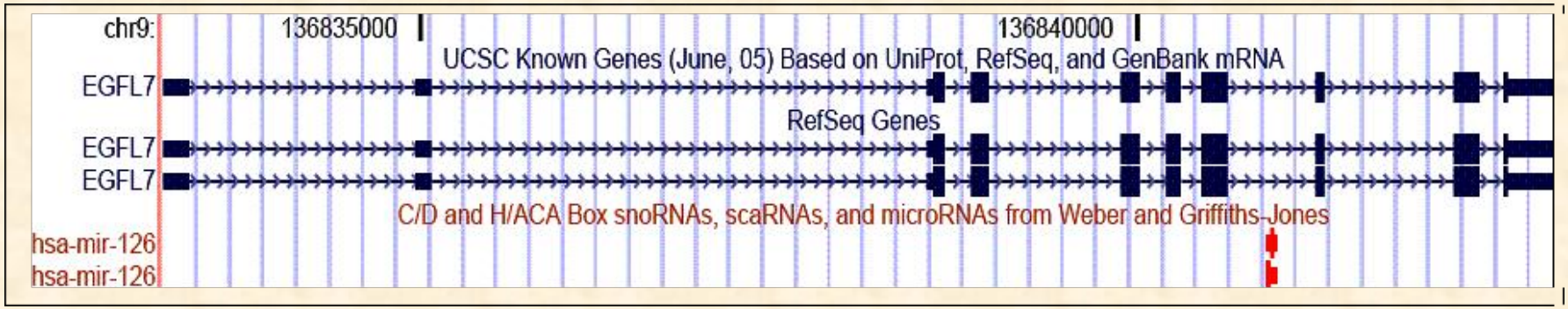






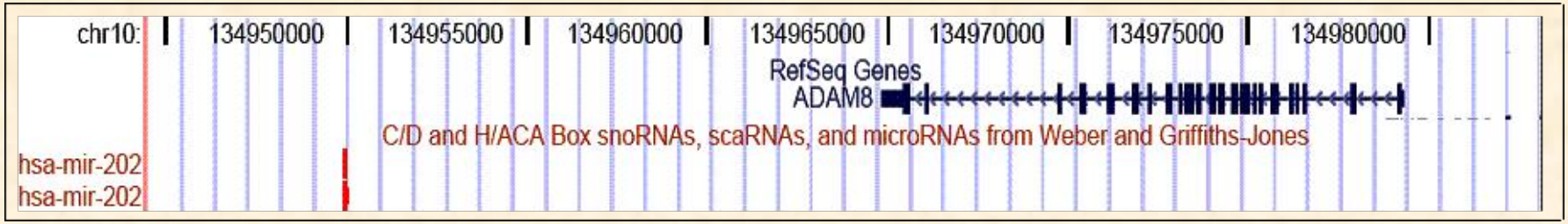
Gene set 1: Same transcription unit (34 genes)

- intronic miRNAs and host genes
- correlate host gene expression with expression of miRNA targets



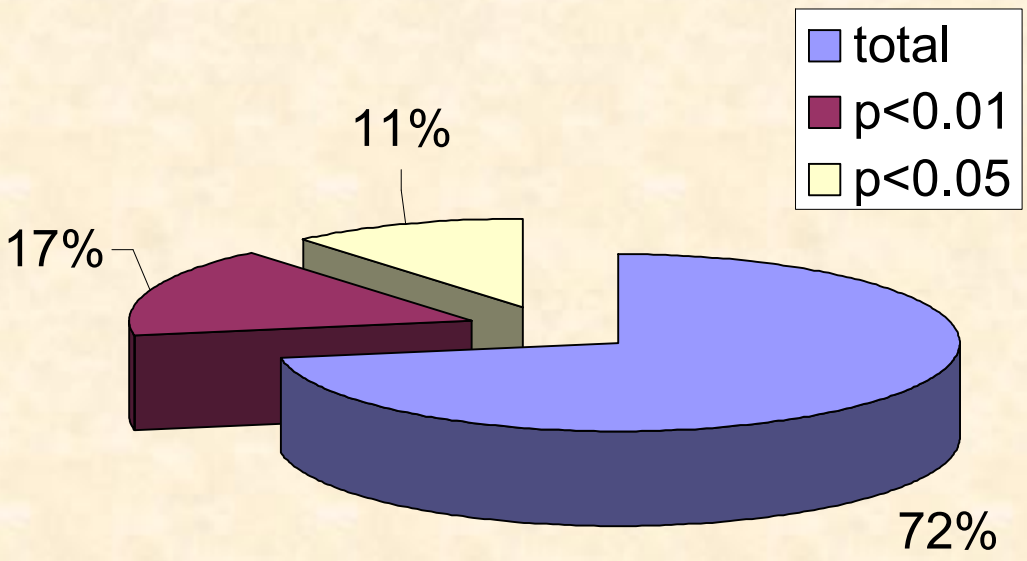
Gene set 2: Different transcription unit (35 genes)

- non-intronic miRNA and upstream genes
- correlate upstream gene expression with expression of miRNA targets

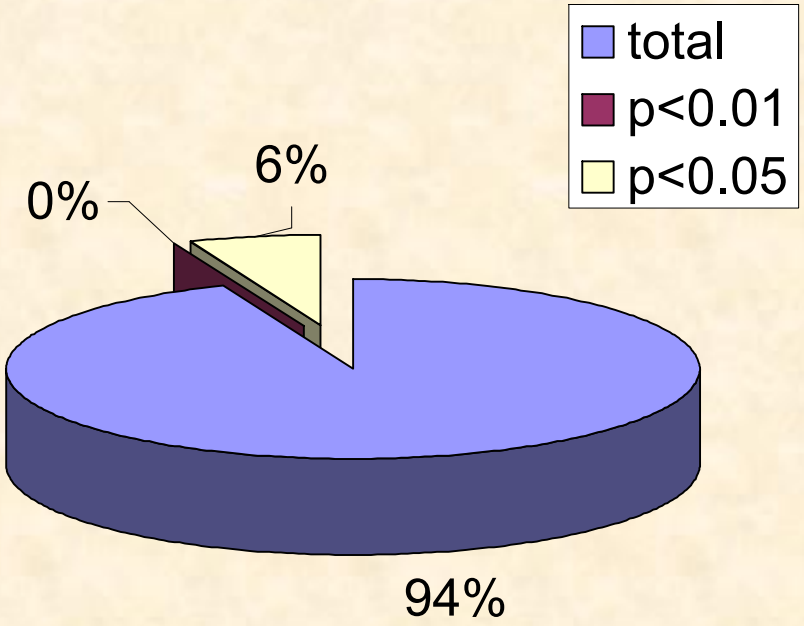


Percentage of host genes that significantly correlate with miRNA target expression in the Valk_leukemia data set

Same transcription unit



Downstream transcription unit



Cluster analysis of clinical samples using host gene expression data

Background

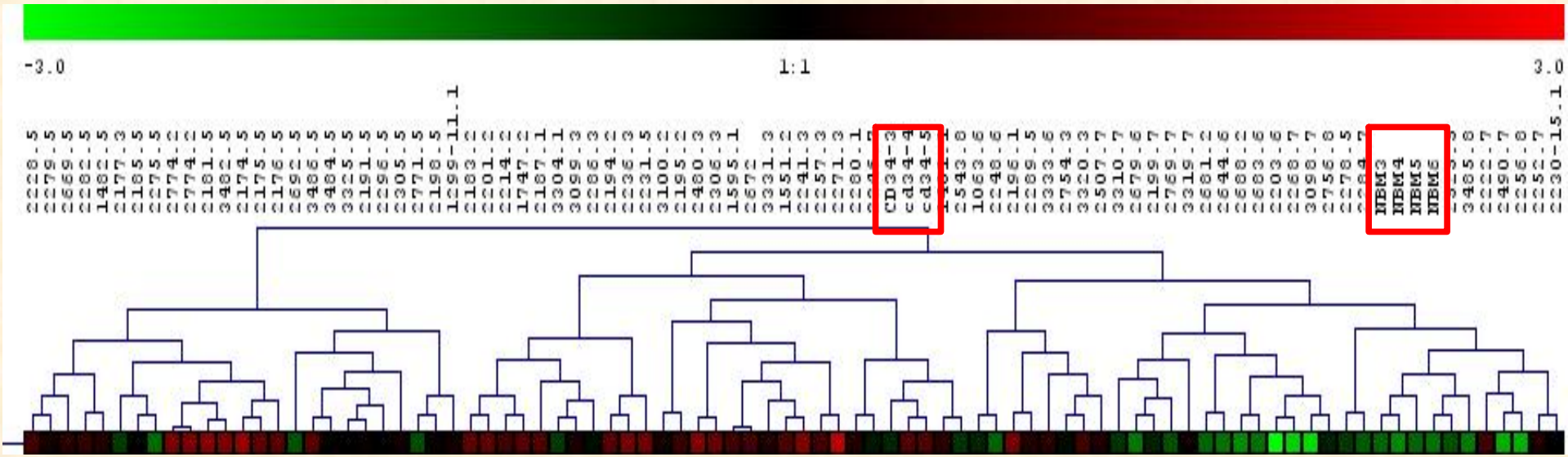
- intronic miRNAs regulate expression of ~1/3 of all miRNA targets, or between 3 and 10% of all genes
- microarray data has shown a negative correlation between miRNA expression and that of target genes

Methods and Results

- normal karyotype clinical samples were extracted from the Valk-leukemia data set
- initial cluster analysis was performed using the entire gene set
- performance of cluster analysis based on host gene expression was compared to that of a random gene subset

NBM, CD34+ and clinical samples from Valk_leukemia are clustered by hostgene expression but not by of a random gene subset

Host gene subset



Random gene subset