R Tools for microRNA Pathway Analysis

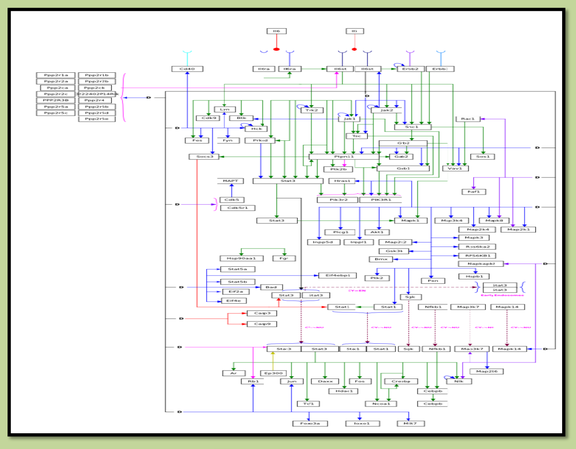
Anwesha Dutta[1,2], Pooja R.Mandaviya[1,2], Rasanpreet Kaur[1,2], Sandeep Mallya[1,2], Lars Eijssen[1], Susan Coort[1], Martijn P van Iersel[1], Jahn Saito[1], Prashantha Hebbar[2] , Chris T Evelo[1].

1. Department of Bioinformatics - BiGCaT, Maastricht University, Maastricht, the Netherlands

2. Manipal Life Sciences Center, Manipal University, Manipal, Karnataka, India

IL-6 signaling Pathway (Mus musculus)

WikiPathways



Identifier	Database
16193	Entrez Gene
16194	Entrez Gene
16194	Entrez Gene
16195	Entrez Gene
16195	Entrez Gene
14159	Entrez Gene
21682	Entrez Gene
12229	Entrez Gene
16452	Entrez Gene
19247	Entrez Gene
16451	Entrez Gene
54721	Entrez Gene
20416	Entrez Gene
14784	Entrez Gene
20848	Entrez Gene
20662	Entrez Gene
22324	Entrez Gene
26413	Entrez Gene
14388	Entrez Gene
18708	Entrez Gene
18709	Entrez Gene
20846	Entrez Gene
20846	Entrez Gene
20846	Entrez Gene

Possible gene targets

MicroRNAs are highly conserved and small non-coding RNA molecules. They inhibit protein expression through translational repression or complete degradation and gene silencing. We integrated validated and predicted miRNAs in biological pathways available on WikiPathways, using miRNA databases. We linked the validated miRNA targets to the genes in the pathways, using BridgeDb for identifier mapping. BridgeDb is a software framework that serves as middleware between relational databases, files and mapping services. The resulting pathways can be used to investigate miRNAs expression results from microarray and sequencing technologies. We developed a Bioconductor package (to be submitted) that uses the existing BridgeDb REST web services from R.

We verify the miRNA targets in the pathways by co-evaluating miRNA and mRNA microarray expression data from a mouse heart failure model. We did this using arrayanalysis.org, which currently has quality control and normalization modules. We added modules for statistical and pathway analysis. Pathway visualization was done using PathVisio which required connecting R to Java, for which a new XMLRPC interface was developed.

Conclusion: we have developed a reusable approach to integrate information about miRNA in pathways and use these for analyses.

This work is a combined effort by the exchange students of Manipal University, India in collaboration with Dept of Bioinformatics-BiGCaT, Maastricht University, Netherlands.

Find miRNAs for target genes

miRWalk · miRTarBase · miRDB · miRBase · miRecords · Tarbase

Pubmed_id	Target gene_species_scientific	Target gene_species_common	Target gene_name	Target gene_Refseq_acc	Target site_number	miRNA_sp	miRNA_m	miRNA_re	Reporter_	Reporter I	Test_meth
14510471	Mus musculus	mouse	Hes1	NM_008235		Mus musc	mmu-miR	overexpression by	siRNA trans		Western b
14510471	Mus musculus	mouse	Hes1	NM_008235		Mus musc	mmu-miR	mutation			Western b
14510471	Mus musculus	mouse	Hes1	NM_008235		Mus musc	mmu-miR	overexpression by	siRNA trans		unknown
15105502	Mus musculus	mouse	HOXB8	NM_010461	1	Mus musc	mmu-miR	196a			5'-RACE
15361871	Mus musculus	mouse	HOXB8	NM_010461	1	Mus musc	mmu-miR	overexpre	GFP	3'UTR	activity as
15538371	Mus musculus	mouse	Mtpn	NM_008098	1	Mus musc	mmu-miR	overexpression by	miRNA prec		Western b
15538371	Mus musculus	mouse	Mtpn	NM_008098	1	Mus musc	mmu-miR	underexpression by	anti-miRN		Western b
15538371	Mus musculus	mouse	Mtpn	NM_008098	1	Mus musc	mmu-miR	overexpression by	miRNA prec		RT-PCR
15685193	Mus musculus	mouse	Mapk14	NM_011951		Homo sap	hsa-miR-	124			
15685193	Mus musculus	mouse	Mapk14	NM_011951		Homo sap	hsa-miR-	124			
15685193	Mus musculus	mouse	Mapk14	NM_011951		Homo sap	hsa-miR-	1overexpression by	mature miR		Microarra
15685193	Mus musculus	mouse	Mapk14	NM_011951		Homo sap	hsa-miR-	124			
15806104	Mus musculus	mouse	Jak2	NM_008413		Mus musc	mmu-miR	overexpre	luciferase	3'UTR	activity as
15806104	Mus musculus	mouse	C1qbp	NM_007573		Mus musc	mmu-miR	375		3'UTR	
15806104	Mus musculus	mouse	Usp1	NM_146144		Mus musc	mmu-miR	375		3'UTR	
15806104	Mus musculus	mouse	Adipor2	NM_197985		Mus musc	mmu-miR	overexpression by	si	3'UTR	Western b
15806104	Mus musculus	mouse	Mtpn	NM_008098		Mus musc	mmu-miR	overexpression by	si	3'UTR	Western b
15806104	Mus musculus	mouse	Mtpn	NM_008098		Mus musc	mmu-let-	overexpression by	si	3'UTR	Western b
15806104	Mus musculus	mouse	Mtpn	NM_008098		Mus musc	mmu-miR	overexpression by	si	3'UTR	Western b
15854907	Mus musculus	mouse	Rtl1/Peg11	NM_184109	1	Mus musc	mmu-miR	434-3p			
15854907	Mus musculus	mouse	Rtl1/Peg11	NM_184109	1	Mus musc	mmu-miR	434-5p			
15854907	Mus musculus	mouse	Rtl1/Peg11	NM_184109	1	Mus musc	mmu-miR	127			
15854907	Mus musculus	mouse	Rtl1/Peg11	NM_184109	1	Mus musc	mmu-miR	433			
15854907	Mus musculus	mouse	Rtl1/Peg11	NM_184109	1	Mus musc	mmu-miR	433*			

Mapping services

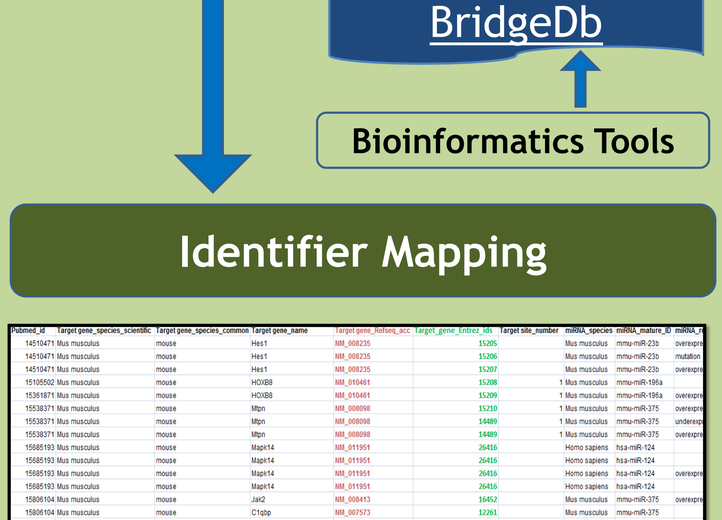
BridgeDb

Bioinformatics Tools

Identifier Mapping

Pubmed_id	Target gene_species_scientific	Target gene_species_common	Target gene_name	Target gene_Refseq_acc	Target gene_Entrez_ids	Target site_number	miRNA_species	miRNA_mature_ID	miRNA_r
14510471	Mus musculus	mouse	Hes1	NM_008235	15205		Mus musculus	mmu-miR-23b	overexpre
14510471	Mus musculus	mouse	Hes1	NM_008235	15206		Mus musculus	mmu-miR-23b	mutation
14510471	Mus musculus	mouse	Hes1	NM_008235	15207		Mus musculus	mmu-miR-23b	overexpre
15105502	Mus musculus	mouse	HOXB8	NM_010461	15208	1	Mus musculus	mmu-miR-196a	
15361871	Mus musculus	mouse	HOXB8	NM_010461	15209	1	Mus musculus	mmu-miR-196a	overexpre
15538371	Mus musculus	mouse	Mtpn	NM_008098	15210	1	Mus musculus	mmu-miR-375	overexpre
15538371	Mus musculus	mouse	Mtpn	NM_008098	14489	1	Mus musculus	mmu-miR-375	underexpre
15538371	Mus musculus	mouse	Mtpn	NM_008098	14489	1	Mus musculus	mmu-miR-375	overexpre
15685193	Mus musculus	mouse	Mapk14	NM_011951	26416		Homo sapiens	hsa-miR-124	
15685193	Mus musculus	mouse	Mapk14	NM_011951	26416		Homo sapiens	hsa-miR-124	
15685193	Mus musculus	mouse	Mapk14	NM_011951	26416		Homo sapiens	hsa-miR-124	overexpre
15685193	Mus musculus	mouse	Mapk14	NM_011951	26416		Homo sapiens	hsa-miR-124	
15806104	Mus musculus	mouse	Jak2	NM_008413	16452		Mus musculus	mmu-miR-375	overexpre
15806104	Mus musculus	mouse	C1qbp	NM_007573	12261		Mus musculus	mmu-miR-375	
15806104	Mus musculus	mouse	Usp1	NM_146144	230484		Mus musculus	mmu-miR-375	
15806104	Mus musculus	mouse	Adipor2	NM_197985	68465		Mus musculus	mmu-miR-375	overexpre
15806104	Mus musculus	mouse	Mtpn	NM_008098	14489		Mus musculus	mmu-miR-124	overexpre
15806104	Mus musculus	mouse	Mtpn	NM_008098	14489		Mus musculus	mmu-let-7b	overexpre
15806104	Mus musculus	mouse	Mtpn	NM_008098	14489		Mus musculus	mmu-miR-375	overexpre
15854907	Mus musculus	mouse	Rtl1/Peg11	NM_184109	353326	1	Mus musculus	mmu-miR-434-3p	
15854907	Mus musculus	mouse	Rtl1/Peg11	NM_184109	353326	1	Mus musculus	mmu-miR-434-5p	
15854907	Mus musculus	mouse	Rtl1/Peg11	NM_184109	353326	1	Mus musculus	mmu-miR-127	
15854907	Mus musculus	mouse	Rtl1/Peg11	NM_184109	353326	1	Mus musculus	mmu-miR-433	

miRNAs added to pathway



miRNA and gene expression microarray analysis

Raw data

Arrayanalysis.org

Raw data Quality Control

Data pre-processing

Evaluation of the effect of normalization

QC report and export of pre-processed (normalized) data

Happy with the QC ?

NO..

Remove faulty array(s) and re-compute QC step.

YES!!

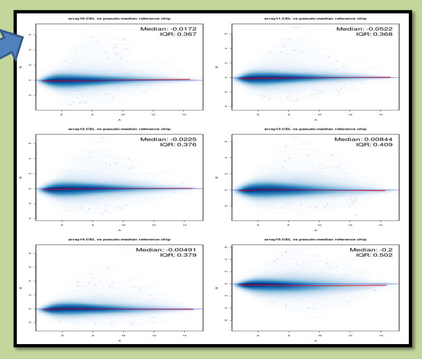
Pre-processed data

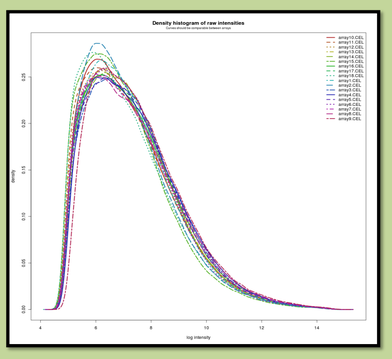
Statistical Analysis

Analyzed data

PathVisio

Visualization

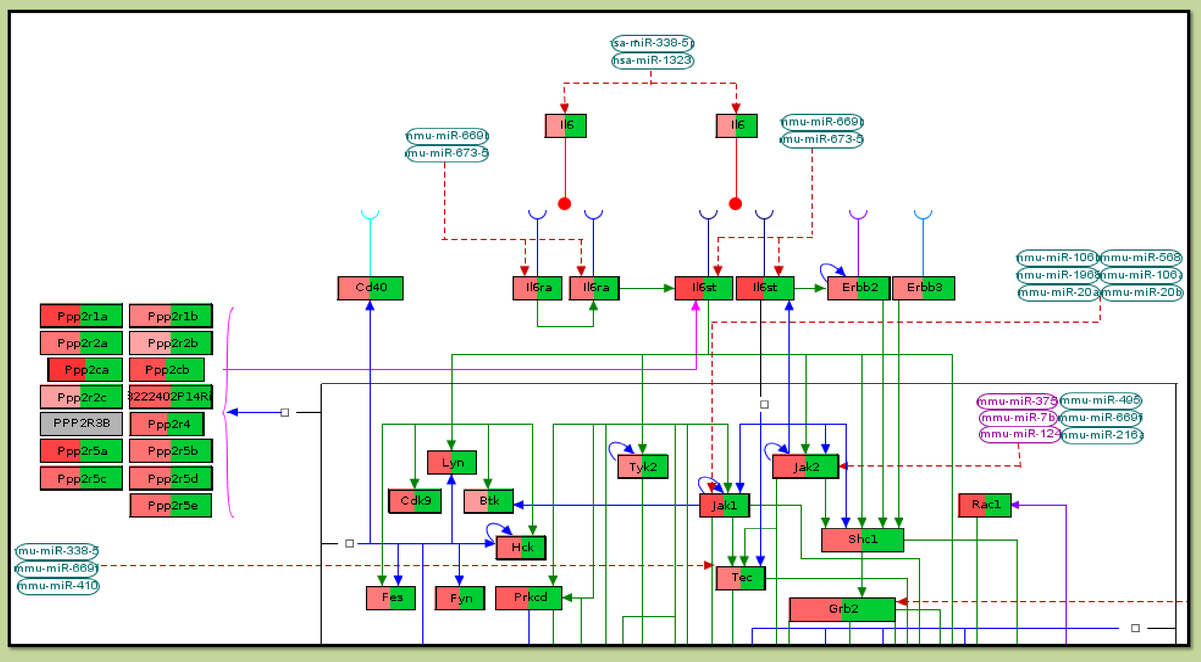




ENSG_ID	array10.CEL	array11.CEL	array12.CEL	array13.CEL	array14.CEL	array15.CEL	array16.CEL	array17.CEL	array18.CEL	array1.CEL
ENSMUSG00000000001	8.9199779249	8.7961905002	8.9127551853	8.6716489206	8.4382715742	9.0814408089	9.725217581	9.6447850105	10.0128244847	8.2289533344
ENSMUSG00000000003	2.6162043686	2.6987160142	2.8599964961	2.7720045265	2.6958679099	2.5485142888	2.6317330477	2.6184537427	2.7864127402	2.6603563678
ENSMUSG00000000028	6.2165516024	6.0526838647	6.2490643718	6.5133258812	6.4129094747	6.2452225504	6.7214129767	6.6506429436	6.6017041348	6.2529906451
ENSMUSG00000000031	6.9999941689	6.909977274	7.2116743871	7.5845957725	7.3753109654	7.5643778958	7.2627342146	7.3312946848	7.2483667708	6.8417102614
ENSMUSG00000000037	3.8823458415	3.9244648174	3.8161733186	3.6727891512	3.8539677954	3.7422416233	3.8062835609	3.8887345237	3.7185932422	4.3090139062
ENSMUSG00000000049	4.3987560967	4.545449555	4.2017556854	4.3689111139	4.0379253517	4.269729279	4.3722956626	4.0528831418	4.2682148859	3.8952272363
ENSMUSG00000000056	9.3784495857	9.4201328388	9.2710916709	9.2840153111	9.2133181733	8.8367390624	8.6518016856	8.0906268327	8.9274109513	9.6442144938
ENSMUSG00000000058	7.6915007051	7.9005960257	7.6566183168	7.673193252	7.5037768756	7.0779585716	7.1087462359	6.9644259461	7.2540751349	7.8802779438
ENSMUSG00000000078	8.2735663368	8.2147266184	8.3986635177	8.3399542699	8.2042458368	8.9957830522	9.2941422184	9.4600317477	9.128715906	8.3455151792
ENSMUSG00000000085	7.5788085167	7.5137700694	7.5426778776	7.3887044088	7.4921510076	7.1747049393	6.8843222266	6.9095628914	7.6768909447	7.4006526119
ENSMUSG00000000088	11.6841776645	11.7556266689	11.7141184631	11.7057368533	11.8221260999	11.7468334598	11.3372382097	11.0069492806	11.4021879075	11.3990637849
ENSMUSG00000000093	6.7831181997	6.9533174114	6.8609436753	6.9883545022	7.0068047702	7.1439775641	6.7066637696	6.5244351246	6.5686202091	6.9828361798
ENSMUSG00000000094	5.009790933	5.0800722005	5.189680284	5.2511612989	5.0540224809	5.5720710455	5.3031004844	5.0568054142	4.8862265853	5.0849426853
ENSMUSG00000000103	2.6247529414	2.8486687446	3.179659946	2.6330386383	2.8371972153	2.7812597105	2.6621009474	2.8572720209	2.7480073132	2.8739247895
ENSMUSG00000000120	6.4899416216	6.3679749553	6.4273180303	6.5218712708	6.5321886509	6.5426156061	6.3253737025	6.5959513108	6.3775100781	6.3750407574
ENSMUSG00000000125	5.4166948007	5.5109482023	5.402118148	5.6567579265	5.3680154609	5.533131759	5.2943248923	5.46478137	5.7210589724	5.6210761393
ENSMUSG00000000126	6.9925168506	6.7501878054	6.8360899327	6.9257430805	6.6078724223	6.9373698602	6.545719914	6.7669333545	6.8218107738	6.7261104742
ENSMUSG00000000127	7.2769294905	7.329999791	7.3617708702	7.2263813916	7.2402039548	6.9196427905	7.2223806754	7.1832806551	7.3153436628	7.4940893518
ENSMUSG00000000131	8.0742045678	8.0580741434	8.2248305362	8.1160326361	8.1452445362	8.0434865128	8.6602785176	8.8403147128	8.3586338934	8.2529391958
ENSMUSG00000000134	6.9568544209	7.0303730062	7.000482053	7.0089673621	6.7389190074	6.8685974578	7.5728064498	7.7306364536	7.3132180248	6.6665957277
ENSMUSG00000000142	6.1024092284	6.1132299621	5.9821649245	6.3969424675	6.3142696343	6.4101516663	6.1571715134	6.0660706601	6.0770949699	6.2734816307
ENSMUSG00000000148	7.3881264257	7.2478207513	7.2444904325	7.3123902916	7.2661738645	7.3865242346	7.3030731212	7.2691180873	7.2782601213	7.1160841217
ENSMUSG00000000149	9.5800720952	9.6902950333	9.5240294856	9.6961186737	9.730995343	9.2833449041	9.1711987054	9.2142881066	9.0437488839	9.764997253
ENSMUSG00000000154	6.3681824678	6.374083189	6.5232974648	6.2793026174	6.4014069807	6.4573219895	6.372609164	6.4389336131	6.5699172947	6.4036771244
ENSMUSG00000000157	5.8444626463	5.5975550648	5.6308740641	5.9477870687	5.5687979249	5.9231740871	5.5668246081	5.7398997464	5.8077959899	5.518525593
ENSMUSG00000000159	5.3886058927	5.4040975223	5.4530270637	5.4383273262	5.5240007294	5.8216625605	5.6725941329	5.2844072996	5.3849988032	5.5736325961
ENSMUSG00000000167	5.1604967828	4.9806707883	5.0489632875	5.4536074113	5.2948673526	5.2165231679	5.3736891813	5.3461303295	5.1269076769	5.2069961998
ENSMUSG00000000168	10.6076252507	10.75237765	10.7041295066	10.6502481779	10.5317085436	10.236833591	9.7888450147	9.3475732447	10.0561152168	10.9285580155
ENSMUSG00000000171	11.6659977879	11.7010665683	11.5356437886	11.5367776674	11.3745250185	11.4070404709	10.9129230714	10.5866887396	11.1902824897	11.5386381904
ENSMUSG00000000182	6.4451598207	6.5437090507	6.4033019853	6.7362636581	6.669971465	6.4773132381	6.6850048671	6.6003404026	6.6752250886	6.6243665795
ENSMUSG00000000183	4.8150122729	4.9116371221	5.0990422272	4.8663375262	4.8225359101	4.6916211581	4.7535659258	4.8732995262	4.7755061731	4.6662419105

ensembl_gene_id	logFC	Fold Change	AveExpr	t	P.Value	adj.P.Val	B	external_gene_id	description
ENSMUSG00000064339	14.30805551	20284.01336	14.29426	228.1954	5.52E-32	8.89E-28	51.16379096	mt-Rnr2	mitochondrially encoded 16S rRNA Gene [Source:MGI Symbol;Acc:MGI:102492]
ENSMUSG00000064337	14.1105095	17688.3211	14.07061	223.4642	7.97E-32	8.89E-28	51.08718614	mt-Rnr1	mitochondrially encoded 12S rRNA Gene [Source:MGI Symbol;Acc:MGI:102493]
ENSMUSG00000060904	13.9335587	15644.36919	13.86305	205.4166	3.51E-31	2.61E-27	50.7522802	Arl1	ADP-ribosylation factor-like 1 Gene [Source:MGI Symbol;Acc:MGI:99436]
ENSMUSG00000064341	14.23740172	19314.56712	14.20478	197.1253	7.23E-31	4.03E-27	50.57181153	mt-Nd1	mitochondrially encoded NADH dehydrogenase 1 Gene [Source:MGI Symbol;Acc:MGI
ENSMUSG00000064363	14.17125877	18449.04819	14.15128	191.7156	1.18E-30	5.26E-27	50.44338174	mt-Nd4	mitochondrially encoded NADH dehydrogenase 4 Gene [Source:MGI Symbol;Acc:MGI
ENSMUSG00000064370	14.04303748	16880.11929	14.03432	187.3761	1.76E-30	6.56E-27	50.33363767	mt-Cytb	mitochondrially encoded cytochrome b Gene [Source:MGI Symbol;Acc:MGI:102501]
ENSMUSG00000068614	13.85680944	14835.94744	13.93738	171.2771	8.56E-30	2.66E-26	49.86590261	Actc1	actin, alpha, cardiac muscle 1 Gene [Source:MGI Symbol;Acc:MGI:87905]
ENSMUSG00000026414	13.76356982	13907.44535	13.66149	170.2214	9.55E-30	2.66E-26	49.8314575	Tnnt2	troponin T2, cardiac Gene [Source:MGI Symbol;Acc:MGI:104597]
ENSMUSG00000036427	11.7919188	3545.857386	11.79911	161.4785	2.41E-29	5.98E-26	49.52543257	Gpi1	glucose phosphate isomerase 1 Gene [Source:MGI Symbol;Acc:MGI:95797]
ENSMUSG00000064345	14.07298945	17234.23275	13.99699	157.7929	3.62E-29	8.07E-26	49.3843765	mt-Nd2	mitochondrially encoded NADH dehydrogenase 2 Gene [Source:MGI Symbol;Acc:MGI
ENSMUSG00000030695	12.97672184	8060.881177	12.91075	150.4486	8.36E-29	1.70E-25	49.07919778	Aldoa	aldolase A, fructose-bisphosphate Gene [Source:MGI Symbol;Acc:MGI:87994]
ENSMUSG00000064367	13.62235442	12610.64479	13.57775	148.3716	1.07E-28	1.99E-25	48.98657822	mt-Nd5	mitochondrially encoded NADH dehydrogenase 5 Gene [Source:MGI Symbol;Acc:MGI
ENSMUSG00000023456	12.53184861	5921.917242	12.58163	143.68	1.88E-28	3.22E-25	48.76614443	Tpi1	triosephosphate isomerase 1 Gene [Source:MGI Symbol;Acc:MGI:98797]
ENSMUSG00000031518	13.40800299	10869.53795	13.25328	140.5158	2.78E-28	4.43E-25	48.60811381	Cox5b	cytochrome c oxidase, subunit Vb Gene [Source:MGI Symbol;Acc:MGI:88475]
ENSMUSG00000081721	13.60557215	12464.80033	13.6326	137.3742	4.13E-28	6.15E-25	48.4431841	Gm13882	predicted gene 13882 Pseudogene [Source:MGI Symbol;Acc:MGI:3650052]
ENSMUSG00000030470	12.3875359	5358.211813	12.39216	135.9512	4.96E-28	6.92E-25	48.3657106	Csrp3	cysteine and glycine-rich protein 3 Gene [Source:MGI Symbol;Acc:MGI:1330824]
ENSMUSG00000024953	11.78016764	3517.092686	11.71073	134.6592	5.87E-28	7.70E-25	48.29381743	Prdx5	peroxiredoxin 5 Gene [Source:MGI Symbol;Acc:MGI:1859821]
ENSMUSG00000039367	9.325746428	641.6961105	9.386639	131.8202	8.54E-28	1.06E-24	48.13043241	Sec24c	Sec24 related gene family, member C (S. cerevisiae) Gene [Source:MGI Symbol;Acc:M
ENSMUSG00000023944	11.56814549	3036.39863	11.86277	131.244	9.22E-28	1.08E-24	48.09633608	Hsp90ab1	heat shock protein 90 alpha (cytosolic), class B member 1 Gene [Source:MGI Symbol;A
ENSMUSG00000071659	9.431016617	690.2698293	9.627433	126.8515	1.68E-27	1.87E-24	47.82547621	Hnrnpul2	heterogeneous nuclear ribonucleoprotein U-like 2 Gene [Source:MGI Symbol;Acc:MG
ENSMUSG00000034075	9.592898558	772.236298	9.749955	123.4861	2.69E-27	2.86E-24	47.60407212	Zdhhc5	zinc finger, DHHC domain containing 5 Gene [Source:MGI Symbol;Acc:MGI:1923573]
ENSMUSG00000026965	8.617872535	392.8602891	8.695374	122.6952	3.01E-27	3.04E-24	47.5501824	Anapc2	anaphase promoting complex subunit 2 Gene [Source:MGI Symbol;Acc:MGI:2139135]
ENSMUSG00000024661	13.64958513	12850.9298	13.49879	122.4177	3.13E-27	3.04E-24	47.53109878	Fth1	ferritin heavy chain 1 Gene [Source:MGI Symbol;Acc:MGI:95588]
ENSMUSG00000020849	10.22057313	1193.161097	10.40911	122.0041	3.33E-27	3.09E-24	47.50249653	Ywhae	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon

miRNAs visualized in the pathway



5155525, Affy
5156267, Affy
5196772, Affy
5222775, Affy
5229068, Affy
5294120, Affy
5330375, Affy
5332609, Affy
5383046, Affy
5460675, Affy
5519783, Affy
5546820, Affy
5558876, Affy
5566205, Affy
5582844, Affy
5591395, Affy
5608034, Affy

Expression data

Gene id on mapp: 13866

Sample name	ENSMUSG00000062312
logFC	7.34644804622696
Fold Change	162.742590044607
AveExpr	7.2395545997626
t	56.1975772547657
P.Value	2.67769165341193E-21
adj.P.Val	2.12425125904386E-20
B	38.3254791221448
external_gene_id	Erbb2
description	v-erb-b2 erythroblastic leukemia viral oncogene ho

Websites used : www.wikipathways.org ; www.bridgedb.org ; www.pathvisio.org ; www.arrayanalysis.org .

Nature Precedings - doi:10.1038/npre.2011.5918.1 : Posted 21 Apr 2011