Supplementary information for:

A family of human microRNA genes from miniature inverted-repeat transposable elements

Jittima Piriyapongsa and I. King Jordan
School of Biology, Georgia Institute of Technology, Atlanta, GA 30332

Supplementary Table 1. **Made1 homologous human expressed sequence tags (ESTs)[1].**

Hit identifiers	% identity	Length	Mis-matches	Gap openings	Query start	Query end	Hit start	Hit end	E-value	Bit score
gi|23517262|gb|BU674347.1|	94.81	77	4	0	4	80	59	135	7.00E-26	121
gi|11450750|gb|BF438233.1|	93.42	76	5	0	1	76	311	386	7.00E-23	111
gi|18976268|gb|BM668437.1|	93.42	76	5	0	1	76	330	405	7.00E-23	111
gi|19006458|gb|BM693200.1|	93.42	76	5	0	1	76	282	207	7.00E-23	111
gi|19721538|gb|BM996637.1|	93.42	76	5	0	1	76	321	396	7.00E-23	111
gi|23274374|gb|BU608159.1|	93.42	76	5	0	1	76	326	401	7.00E-23	111
gi|2784598|gb|AA743782.1|	93.42	76	5	0	1	76	110	185	7.00E-23	111
gi|2876039|gb|AA804638.1|	93.42	76	5	0	1	76	318	393	7.00E-23	111
gi|3933745|gb|AI290971.1|	93.42	76	5	0	1	76	311	386	7.00E-23	111
gi|4990875|gb|AI702975.1|	93.42	76	5	0	1	76	307	382	7.00E-23	111
gi|5454573|gb|AI832593.1|	93.42	76	5	0	1	76	309	384	7.00E-23	111
gi|7320253|gb|AW615067.1|	93.42	76	5	0	1	76	313	388	7.00E-23	111
gi|8167811|gb|AW976581.1|	93.42	76	5	0	1	76	307	382	7.00E-23	111
gi|8359944|gb|BE042891.1|	93.42	76	5	0	1	76	307	382	7.00E-23	111
gi|52721466|gb|CV371411.1|	95.52	67	3	0	14	80	236	170	3.00E-22	109
gi|32004424|emb|BX492684.1|	92.41	79	6	0	2	80	382	304	3.00E-22	109
gi|2907387|gb|AA833659.1|	93.24	74	5	0	3	76	199	126	1.00E-21	107
gi|52700258|gb|CV350203.1|	93.24	74	5	0	2	75	551	624	1.00E-21	107
gi|6837361|gb|AW340735.1|	93.24	74	5	0	3	76	216	143	1.00E-21	107
gi|7039615|gb|AW469509.1|	93.24	74	5	0	3	76	216	143	1.00E-21	107
gi|3400022|gb|AI073378.1|	91.25	80	7	0	1	80	241	320	2.00E-20	103
gi|46547768|gb|CN478769.1|	91.25	80	7	0	1	80	255	334	2.00E-20	103
gi|20494289|gb|BQ269223.1|	92	75	6	0	1	75	483	409	6.00E-20	101
gi|44842622|gb|CK825697.1|	92	75	6	0	1	75	470	396	6.00E-20	101
gi|45695156|emb|AL519606.3|	94.12	68	3	1	13	80	747	681	3.00E-19	99.6
gi|52811228|gb|CV415725.1|	90.91	77	7	0	4	80	190	266	1.00E-18	97.6
gi|2908283|gb|AA834684.1|	94.12	68	3	1	5	72	137	203	4.00E-18	95.6
gi|13292606|gb|BG399158.1|	90.79	76	7	0	5	80	232	157	4.00E-18	95.6
gi|52653216|gb|CV330002.1|	90.79	76	7	0	5	80	163	88	4.00E-18	95.6
gi|8061011|gb|AW896806.1|	90.79	76	7	0	1	76	289	214	4.00E-18	95.6
gi|52667308|gb|CV344094.1|	92.11	76	5	1	1	76	280	206	4.00E-18	95.6
gi|5438416|gb|AI819337.1|	90	80	8	0	1	80	241	320	4.00E-18	95.6
gi|7946376|gb|AW850859.1|	90	80	8	0	1	80	170	249	4.00E-18	95.6
gi|12766146|gb|BG256330.1|	92.5	80	4	2	1	80	385	308	4.00E-18	95.6
gi|27846682|emb|BX105680.1|	91.55	71	6	0	8	78	364	434	2.00E-17	93.7
gi|3837536|gb|AI242139.1|	91.55	71	6	0	8	78	269	199	2.00E-17	93.7
gi|58568449|dbj|BP395858.1|	90.54	74	7	0	7	80	284	211	6.00E-17	91.7
gi|6602709|emb|AL134522.1|	93.24	74	3	2	4	77	28	99	6.00E-17	91.7

gi|91749404|gb|EB386059.1|	91.3	69	6	0	11	79	162	94	2.00E-16	89.7
gi|32005544|emb|BX493226.1|	92.75	69	4	1	7	75	216	149	2.00E-16	89.7
gi|14321058|gb|BG926535.1|	90.41	73	7	0	2	74	661	589	2.00E-16	89.7
gi|52707894|gb|CV357839.1|	90.41	73	7	0	2	74	127	55	2.00E-16	89.7
gi|82333517|dbj|DA902558.1|	90.91	77	6	1	1	77	193	118	2.00E-16	89.7
gi|12120877|gb|BF772977.1|	90.12	81	7	1	1	80	235	155	2.00E-16	89.7
gi|12120883|gb|BF772983.1|	90.12	81	7	1	1	80	236	156	2.00E-16	89.7
gi|5110886|gb|AI742598.1|	90.12	81	7	1	1	80	241	321	2.00E-16	89.7
gi|1885842|gb|AA250882.1|	92.19	64	5	0	1	64	41	104	1.00E-15	87.7
gi|82341158|dbj|DB016887.1|	92.19	64	5	0	1	64	188	251	1.00E-15	87.7
gi|10200151|gb|BE778953.1|	90.28	72	7	0	9	80	149	220	1.00E-15	87.7
gi|13339103|gb|BG432597.1|	89.47	76	8	0	5	80	516	441	1.00E-15	87.7
gi|13343062|gb|BG436556.1|	89.47	76	8	0	1	76	300	375	1.00E-15	87.7
gi|15164200|emb|AL600694.1|	89.47	76	8	0	1	76	324	399	1.00E-15	87.7
gi|18983536|gb|BM673638.1|	89.47	76	8	0	5	80	116	41	1.00E-15	87.7
gi|19005651|gb|BM692393.1|	89.47	76	8	0	5	80	185	260	1.00E-15	87.7
gi|2834284|gb|AA774950.1|	89.47	76	8	0	1	76	221	146	1.00E-15	87.7
gi|11977833|gb|BF692425.1|	90.79	76	6	1	1	76	386	460	1.00E-15	87.7
gi|13452873|gb|BG491361.1|	90.79	76	6	1	1	76	8	82	1.00E-15	87.7
gi|13580923|gb|BG573270.1|	90.79	76	6	1	1	76	290	364	1.00E-15	87.7
gi|19727271|gb|BQ002371.1|	90.79	76	6	1	1	76	400	326	1.00E-15	87.7
gi|24776874|gb|CA414223.1|	90.79	76	6	1	1	76	400	326	1.00E-15	87.7
gi|27932373|gb|CB106566.1|	90.79	76	6	1	1	76	13	87	1.00E-15	87.7
gi|28365225|gb|CB243581.1|	90.79	76	6	1	1	76	29	103	1.00E-15	87.7
gi|43429246|emb|BX952415.1|	90.79	76	6	1	5	80	89	163	1.00E-15	87.7
gi|43425548|emb|BX951140.1|	88.75	80	9	0	1	80	136	57	1.00E-15	87.7
gi|3038959|gb|AA903836.1|	90	80	7	1	1	80	74	152	1.00E-15	87.7
gi|5543963|gb|AI869995.1|	90	80	7	1	1	80	450	372	1.00E-15	87.7
gi|7668921|gb|AW753989.1|	91.04	67	6	0	12	78	421	487	4.00E-15	85.7
gi|7668972|gb|AW754040.1|	91.04	67	6	0	12	78	421	487	4.00E-15	85.7
gi|8046501|gb|AW884489.1|	88.61	79	9	0	2	80	132	210	4.00E-15	85.7
gi|14466558|gb|BI059028.1|	90	70	7	0	3	72	129	198	1.00E-14	83.8
gi|81125345|dbj|DA460339.1|	90	70	7	0	3	72	349	280	1.00E-14	83.8
gi|27845181|emb|BX102210.1|	89.74	78	7	1	1	77	407	330	1.00E-14	83.8
gi|31915369|emb|BX479525.1|	89.04	73	8	0	4	76	130	202	6.00E-14	81.8
gi|66791763|dbj|BP425510.1|	89.04	73	8	0	4	76	184	256	6.00E-14	81.8
gi|685935|gb|T71414.1|	88.16	76	9	0	1	76	11	86	6.00E-14	81.8
gi|711241|gb|T82953.1|	88.16	76	9	0	1	76	11	86	6.00E-14	81.8
gi|1404173|gb|W88623.1|	87.5	80	10	0	1	80	146	67	6.00E-14	81.8
gi|1891141|gb|AA257012.1|	88.89	81	8	1	1	80	248	168	6.00E-14	81.8
gi|81181343|dbj|DA639796.1|	90.62	64	6	0	1	64	16	79	2.00E-13	79.8
gi|8058080|gb|AW893875.1|	92.19	64	4	1	1	64	471	409	2.00E-13	79.8
gi|14372680|gb|BG954509.1|	90.28	72	6	1	9	80	238	168	2.00E-13	79.8
gi|79163886|dbj|DA105807.1|	90.28	72	6	1	9	80	387	317	2.00E-13	79.8
gi|8623066|gb|BE160345.1|	90.28	72	6	1	9	80	94	164	2.00E-13	79.8
gi|8623148|gb|BE160427.1|	90.28	72	6	1	9	80	94	164	2.00E-13	79.8
gi|2162267|gb|AA448597.1|	88.16	76	9	0	1	76	343	418	2.00E-13	79.8
gi|80799866|dbj|DA505931.1|	88.16	76	9	0	1	76	121	196	2.00E-13	79.8
gi|21855046|gb|BQ716149.1|	89.47	76	7	1	1	76	119	193	2.00E-13	79.8

gi|24805094|gb|CA440674.1|	89.47	76	7	1	1	76	400	326	2.00E-13	79.8
gi|83480277|dbj|DB358036.1|	89.47	76	7	1	1	76	382	308	2.00E-13	79.8
gi|3056341|gb|AA916949.1|	87.5	80	10	0	1	80	230	309	2.00E-13	79.8
gi|1764951|gb|AA181484.1|	88.75	80	8	1	2	80	360	281	2.00E-13	79.8
gi|8054117|gb|AW889912.1|	88.75	80	8	1	1	80	149	227	2.00E-13	79.8
gi|83532058|dbj|DB333866.1|	88.75	80	8	1	1	80	65	143	2.00E-13	79.8
gi|10107714|gb|BE719449.1|	88.73	71	8	0	1	71	623	553	9.00E-13	77.8
gi|504666|dbj|D20846.1|	90.14	71	6	1	1	70	191	121	9.00E-13	77.8
gi|14393270|gb|BG989200.1|	89.33	75	7	1	6	80	304	231	9.00E-13	77.8
gi|31446439|gb|CD514721.1|	89.33	75	7	1	6	80	13	86	9.00E-13	77.8
gi|2617003|gb|AA663012.1|	88.46	78	7	1	5	80	105	28	9.00E-13	77.8
gi|90847359|dbj|DB577513.1|	87.65	81	7	1	1	78	52	132	9.00E-13	77.8
gi|83241952|dbj|DB315742.1|	90	70	6	1	11	79	73	142	4.00E-12	75.8
gi|91749668|gb|EB386323.1|	90	70	6	1	9	77	226	157	4.00E-12	75.8
gi|10918992|dbj|AV761144.1|	88.46	78	8	1	3	79	236	313	4.00E-12	75.8
gi|78737823|dbj|DA326471.1|	88.46	78	8	1	3	80	82	6	4.00E-12	75.8
gi|83199537|dbj|DB235269.1|	85.88	85	6	1	2	80	480	396	4.00E-12	75.8
gi|7668920|gb|AW753988.1|	89.23	65	7	0	14	78	91	27	1.00E-11	73.8
gi|7668971|gb|AW754039.1|	89.23	65	7	0	14	78	91	27	1.00E-11	73.8
gi|83486421|dbj|DB358889.1|	89.23	65	7	0	16	80	329	393	1.00E-11	73.8
gi|2328991|gb|AA558514.1|	87.67	73	9	0	4	76	115	43	1.00E-11	73.8
gi|81156387|dbj|DA383600.1|	87.67	73	9	0	4	76	516	588	1.00E-11	73.8
gi|8165082|gb|AW973998.1|	87.67	73	9	0	4	76	244	172	1.00E-11	73.8
gi|23373989|gb|BU661807.1|	88.16	76	7	1	5	80	85	158	1.00E-11	73.8
gi|3896467|gb|AI274199.1|	88.16	76	7	1	1	74	74	149	1.00E-11	73.8
gi|5054918|gb|AI733805.1|	88.16	76	7	1	1	74	72	147	1.00E-11	73.8
gi|82136433|dbj|DB047679.1|	88.31	77	8	1	1	76	152	228	1.00E-11	73.8
gi|33252132|gb|CF136688.1|	90.62	64	5	1	1	64	30	92	6.00E-11	71.9
gi|46922787|emb|BX405577.2|	87.5	80	9	1	1	80	146	224	6.00E-11	71.9
gi|46233530|emb|AL566894.3|	86.3	73	9	1	1	73	526	455	2.00E-10	69.9
gi|3872647|gb|AI264444.1|	86.84	76	8	1	1	74	72	147	2.00E-10	69.9
gi|33258518|gb|CF143074.1|	86.25	80	7	1	1	80	184	259	2.00E-10	69.9
gi|7111293|gb|AW499536.1|	86.25	80	7	1	1	80	184	259	2.00E-10	69.9
gi|7111295|gb|AW499537.1|	86.25	80	7	1	1	80	184	259	2.00E-10	69.9
gi|7111327|gb|AW499553.1|	86.25	80	7	1	1	80	184	259	2.00E-10	69.9
gi|7116331|gb|AW502136.1|	86.25	80	7	1	1	80	184	259	2.00E-10	69.9
gi|7116335|gb|AW502138.1|	86.25	80	7	1	1	80	184	259	2.00E-10	69.9
gi|3214298|gb|AI004788.1|	85.71	84	8	1	1	80	128	45	2.00E-10	69.9
gi|90648194|dbj|BY797461.2|	85.71	84	8	1	1	80	347	430	2.00E-10	69.9
gi|1486755|gb|AA022674.1|	84.88	86	7	1	1	80	283	198	2.00E-10	69.9
gi|1486863|gb|AA022709.1|	84.88	86	7	1	1	80	101	186	2.00E-10	69.9
gi|83124689|dbj|DB343577.1|	86.49	74	10	0	4	77	405	478	9.00E-10	67.9
gi|90938918|dbj|DB507251.1|	87.84	74	8	1	1	74	74	2	9.00E-10	67.9
gi|694186|gb|T76983.1|	86.84	76	7	1	5	80	204	132	9.00E-10	67.9
gi|83078449|dbj|DB106337.1|	85.9	78	7	1	1	78	458	385	9.00E-10	67.9
gi|83237080|dbj|DB354909.1|	86.25	80	8	1	1	80	218	294	9.00E-10	67.9
gi|12189868|gb|BF837652.1|	88.41	69	7	1	1	69	148	215	4.00E-09	65.9
gi|8167508|gb|AW976282.1|	87.67	73	8	1	5	76	495	423	4.00E-09	65.9
gi|14399447|gb|BG995377.1|	85.71	77	11	0	4	80	227	151	4.00E-09	65.9

gi|81108769|dbj|DA381665.1|	85.9	78	7	1	1	78	487	414	4.00E-09	65.9
gi|83190413|dbj|DB352537.1|	87.65	81	8	2	1	80	312	391	4.00E-09	65.9
gi|83517488|dbj|DB143470.1|	86.84	76	9	1	1	76	116	42	1.00E-08	63.9
gi|82338292|dbj|DB049879.1|	85	80	12	0	1	80	166	87	1.00E-08	63.9
gi|80933029|dbj|DA523524.1|	86.25	80	10	1	1	80	468	390	1.00E-08	63.9

[1] BLASTN was used to search the human EST database with a full length Made1 element query sequence. Only hits that were ≥80% identical over ≥80% of the length of the element are reported. Hit identifiers (Genbank identification numbers and accessions) are shown followed by the BLAST statistics for each query-hit pair.

Supplementary Table 2. **Over-represented GO biological process categories among genes with Made1 derived hsa-mir-548 target sites.**

GO ID[1]	Description[2]	Gene acc[3]	Obs[4]	Exp[5]	*P*-value[6]
GO:0000087	M phase of mitotic cell cycle	ENSG00000130177 ENSG00000086827* ENSG00000004897*	3	0.44	9.42E-03
GO:0007067	mitosis	ENSG00000130177 ENSG00000086827* ENSG00000004897*	3	0.43	9.06E-03
GO:0007088	regulation of mitosis	ENSG00000130177 ENSG00000086827*	2	0.12	6.39E-03
GO:0006917	induction of apoptosis	ENSG00000163161 ENSG00000171132* ENSG00000004468	3	0.44	9.42E-03
GO:0012502	induction of programmed cell death	ENSG00000163161 ENSG00000171132* ENSG00000004468	3	0.44	9.42E-03
GO:0008283	cell proliferation	ENSG00000076716 ENSG00000112038* ENSG00000143125 ENSG00000125657 ENSG00000130177 ENSG00000004897*	6	1.58	4.47E-03
GO:0007059	chromosome segregation	ENSG00000163535 ENSG00000086827*	2	0.11	5.17E-03

[1] GO biological process category ID

[2] Functional description for the GO category

[3] The list of Ensembl gene accessions in the GO category, * indicates genes that are down-regulated in colorectal cancer tissue

[4] Observed gene number in the GO category

[5] Expected gene number in the GO category

[6] *P*-value showing significance of enrichment for the GO category based on the hypergeometric test

Supplementary Table 3. **Over-represented GO biological process categories among genes with miRanda predicted hsa-mir-548 target sites that map to colorectal cancer down-regulated co-expression clusters (*i.e.* 12, 15 & 20 in Figure 6).**

GO ID[1]	**Description[2]**	**Gene acc[3]**	**Obs[4]**	**Exp[5]**	***P*-value[6]**
GO:0007155	cell adhesion	ENSG00000179776 ENSG00000040731 ENSG00000154162 ENSG00000133800 ENSG00000073712 ENSG00000138080 ENSG00000018236 ENSG00000038427 ENSG00000170989 ENSG00000146648 ENSG00000128536 ENSG00000087303 ENSG00000115414 ENSG00000102290 ENSG00000164171 ENSG00000158887 ENSG00000067141 ENSG00000124215 ENSG00000107562 ENSG00000112378 ENSG00000143341 ENSG00000164199 ENSG00000077522 ENSG00000104415 ENSG00000163347 ENSG00000154655 ENSG00000198542	27	10.05	2.61E-06
GO:0016337	cell-cell adhesion	ENSG00000179776 ENSG00000040731 ENSG00000154162 ENSG00000146648 ENSG00000128536 ENSG00000102290 ENSG00000158887 ENSG00000124215 ENSG00000164199 ENSG00000163347	10	3.45	2.48E-03
GO:0007156	homophilic cell adhesion	ENSG00000179776 ENSG00000040731 ENSG00000154162 ENSG00000128536	7	2.16	5.99E-03

		ENSG00000102290 ENSG00000158887 ENSG00000124215			
GO:0031589	cell-substrate adhesion	ENSG00000133800 ENSG00000087303 ENSG00000164171 ENSG00000077522 ENSG00000198542	5	0.9	2.03E-03
GO:0007160	cell-matrix adhesion	ENSG00000133800 ENSG00000087303 ENSG00000164171 ENSG00000077522 ENSG00000198542	5	0.9	2.03E-03
GO:0007154	cell communication	ENSG00000064989 ENSG00000153208 ENSG00000145632 ENSG00000166073 ENSG00000184984 ENSG00000080644 ENSG00000147432 ENSG00000135902 ENSG00000108018 ENSG00000018236 ENSG00000174429 ENSG00000169676 ENSG00000170989 ENSG00000146648 ENSG00000140009 ENSG00000151348 ENSG00000138685 ENSG00000115641 ENSG00000115414 ENSG00000113327 ENSG00000091844 ENSG00000164949 ENSG00000146072 ENSG00000135821 ENSG00000127920 ENSG00000177464 ENSG00000132975 ENSG00000064652 ENSG00000171189 ENSG00000095752 ENSG00000164171 ENSG00000183111 ENSG00000182634	76	51.71	1.01E-04

		ENSG00000113594 ENSG00000101665 ENSG00000116141 ENSG00000124089 ENSG00000143198 ENSG00000158887 ENSG00000067141 ENSG00000134259 ENSG00000170485 ENSG00000133636 ENSG00000165588 ENSG00000169860 ENSG00000167941 ENSG00000115252 ENSG00000154678 ENSG00000172572 ENSG00000113448 ENSG00000108551 ENSG00000156475 ENSG00000156218 ENSG00000144724 ENSG00000115665 ENSG00000166592 ENSG00000107562 ENSG00000196632 ENSG00000196781 ENSG00000105989 ENSG00000175868 ENSG00000152284 ENSG00000182880 ENSG00000164199 ENSG00000104415 ENSG00000049246 ENSG00000124104 ENSG00000078043 ENSG00000165970 ENSG00000149305 ENSG00000170579 ENSG00000198542 ENSG00000137962 ENSG00000198752 ENSG00000064692 ENSG00000198929			
GO:0007267	cell-cell signaling	ENSG00000153208 ENSG00000166073 ENSG00000147432	20	8.45	3.13E-04

		ENSG00000169676 ENSG00000140009 ENSG00000138685 ENSG00000135821 ENSG00000171189 ENSG00000095752 ENSG00000158887 ENSG00000067141 ENSG00000134259 ENSG00000115665 ENSG00000107562 ENSG00000104415 ENSG00000165970 ENSG00000149305 ENSG00000170579 ENSG00000064692 ENSG00000198929			
GO:0019226	transmission of nerve impulse	ENSG00000166073 ENSG00000147432 ENSG00000169676 ENSG00000135821 ENSG00000171189 ENSG00000158887 ENSG00000115665 ENSG00000165970 ENSG00000149305 ENSG00000170579 ENSG00000064692 ENSG00000198929	12	3.79	4.30E-04
GO:0007268	synaptic transmission	ENSG00000166073 ENSG00000147432 ENSG00000169676 ENSG00000135821 ENSG00000171189 ENSG00000158887 ENSG00000115665 ENSG00000165970 ENSG00000149305 ENSG00000170579 ENSG00000064692 ENSG00000198929	12	3.64	3.02E-04
GO:0001505	regulation of neurotransmitter levels	ENSG00000135821 ENSG00000115665 ENSG00000064692 ENSG00000198929	4	0.73	5.84E-03
GO:0007165	signal transduction	ENSG00000064989	65	47.47	2.90E-03

		ENSG00000153208 ENSG00000145632 ENSG00000166073 ENSG00000184984 ENSG00000080644 ENSG00000147432 ENSG00000135902 ENSG00000108018 ENSG00000018236 ENSG00000174429 ENSG00000169676 ENSG00000170989 ENSG00000146648 ENSG00000140009 ENSG00000151348 ENSG00000138685 ENSG00000115641 ENSG00000115414 ENSG00000113327 ENSG00000091844 ENSG00000164949 ENSG00000146072 ENSG00000127920 ENSG00000177464 ENSG00000132975 ENSG00000064652 ENSG00000171189 ENSG00000164171 ENSG00000183111 ENSG00000182634 ENSG00000113594 ENSG00000101665 ENSG00000116141 ENSG00000124089 ENSG00000143198 ENSG00000170485 ENSG00000133636 ENSG00000165588 ENSG00000169860 ENSG00000167941 ENSG00000115252 ENSG00000154678 ENSG00000172572 ENSG00000113448 ENSG00000108551 ENSG00000156475			

		ENSG00000156218 ENSG00000144724 ENSG00000166592 ENSG00000107562 ENSG00000196632 ENSG00000196781 ENSG00000105989 ENSG00000175868 ENSG00000152284 ENSG00000182880 ENSG00000164199 ENSG00000104415 ENSG00000049246 ENSG00000124104 ENSG00000078043 ENSG00000198542 ENSG00000137962 ENSG00000198752			
GO:0051056	regulation of small GTPase mediated signal transduction	ENSG00000174429 ENSG00000183111 ENSG00000198752	3	0.27	2.38E-03
GO:0035023	regulation of Rho protein signal transduction	ENSG00000174429 ENSG00000183111	2	0.06	1.52E-03
GO:0007266	Rho protein signal transduction	ENSG00000174429 ENSG00000183111 ENSG00000137962	3	0.39	6.51E-03
GO:0009966	regulation of signal transduction	ENSG00000145632 ENSG00000174429 ENSG00000091844 ENSG00000183111 ENSG00000165588 ENSG00000167941 ENSG00000196781 ENSG00000152284 ENSG00000198752	9	3.31	6.04E-03
GO:0006575	amino acid derivative metabolism	ENSG00000129596 ENSG00000131480 ENSG00000115665 ENSG00000064692	4	0.82	9.08E-03
GO:0009250	glucan biosynthesis	ENSG00000111713 ENSG00000056998	2	0.15	8.65E-03
GO:0005978	glycogen biosynthesis	ENSG00000111713 ENSG00000056998	2	0.15	8.65E-03
GO:0007417	central nervous system development	ENSG00000061676 ENSG00000171189 ENSG00000170485	6	1.82	9.91E-03

		ENSG00000165588 ENSG00000134595 ENSG00000043355			
GO:0007596	blood coagulation	ENSG00000095752 ENSG00000164171 ENSG00000169860 ENSG00000143341 ENSG00000154655	5	1.29	9.39E-03
GO:0051260	protein homooligomerization	ENSG00000187134 ENSG00000077522	2	0.15	8.65E-03
GO:0050952	sensory perception of electrical stimulus	ENSG00000182634 ENSG00000182880	2	0	0.00E+00
GO:0050978	magnetoreception, using electrical stimulus	ENSG00000182634 ENSG00000182880	2	0	0.00E+00
GO:0050954	sensory perception of mechanical stimulus	ENSG00000153208 ENSG00000115380 ENSG00000131480 ENSG00000140522 ENSG00000143341 ENSG00000164199	6	1.58	5.03E-03
GO:0050979	magnetoreception, using mechanical stimulus	ENSG00000153208 ENSG00000115380 ENSG00000131480 ENSG00000140522 ENSG00000143341 ENSG00000164199	6	0	0.00E+00
GO:0019233	sensory perception of pain	ENSG00000165091 ENSG00000164199 ENSG00000136156	3	0.03	1.41E-06
GO:0050966	detection of mechanical stimulus during sensory perception of pain	ENSG00000165091 ENSG00000164199 ENSG00000136156	3	0	0.00E+00
GO:0051341	regulation of oxidoreductase activity	ENSG00000146648 ENSG00000198929	2	0.06	1.52E-03
GO:0050999	regulation of nitric-oxide synthase activity	ENSG00000146648 ENSG00000198929	2	0.06	1.52E-03

[1] GO biological process category ID

[2] Functional description for the GO category

[3] The list of Ensembl gene accessions in the GO category

[4] Observed gene number in the GO category

[5] Expected gene number in the GO category

[6] *P*-value showing significance of enrichment for the GO category based on the hypergeometric test

Supplementary Table 4. **Putative hsa-mir-548 target genes previously implicated as being involved in colorectal cancer by microarray expression profiling.**

Accn[1]	**Ref**[2]	**Name**[3]	**Status**[4]	**Target**[5]	***P*-value**[6]
ENST00000282050	[1]	ATP synthase alpha chain, mitochondrial precursor (EC 3.6.3.14)	down	a	1.73E-05
ENST00000219660	[1]	Aquaporin-8 (AQP-8)	down	b	0.0050
ENST00000262825	[2]	Cytokine receptor common beta chain precursor (GM-CSF/IL-3/IL-5 receptor common beta-chain) (CD131 antigen) (CDw131)	down	b	0.0006
ENST00000201031	[2]	Transcription factor AP-2 gamma (AP2-gamma) (Activating enhancer- binding protein 2 gamma) (Transcription factor ERF-1)	down	a,b,c,d	0.0006
ENST00000241261	[2]	Tumor necrosis factor ligand superfamily member 10 (TNF-related apoptosis-inducing ligand) (TRAIL protein) (Apo-2 ligand) (Apo-2L) (CD253 antigen)	down	a	0.0315
ENST00000360121	[2]	Leukosialin precursor (Leucocyte sialoglycoprotein) (Sialophorin) (Galactoglycoprotein) (GALGP) (CD43 antigen)	down	a,c	0.0018
ENST00000360876	[2]	Eukaryotic translation initiation factor 3 subunit 9 (eIF-3 eta) (eIF3 p116) (eIF3 p110) (eIF3b) (Prt1 homolog) (hPrt1)	up	a	0.0176
ENST00000368083	[3]	Arginase-1 (EC 3.5.3.1) (Type I arginase) (Liver-type arginase)	down	c	0.0308
ENST00000344548	[3]	Cell division control protein 42 homolog precursor (G25K GTP-binding protein)	down	c	0.0455
ENST00000379328	[3]	Trans-acting T-cell-specific transcription factor GATA-3 (GATA-binding factor 3)	down	a	0.0012
ENST00000285900	[3]	Glutamate receptor 1 precursor (GluR-1) (GluR-A) (GluR-K1) (Glutamate receptor ionotropic, AMPA 1) (AMPA-selective glutamate receptor 1)	down	d	0.0058
ENST00000328245	[3]	Heat shock factor protein 1 (HSF 1) (Heat shock transcription factor 1) (HSTF 1)	down	c	1.03E-05
ENST00000227752	[3]	Interleukin-10 receptor alpha chain precursor (IL-10R-A) (IL-10R1) (CDw210a antigen)	down	d	0.0206
ENST00000371794	[3]	Noelin precursor (Neuronal olfactomedin-related ER localized protein) (Olfactomedin-1)	down	d	0.0410
ENST00000334661	[3]	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase delta 1 (EC 3.1.4.11) (Phosphoinositide phospholipase C) (PLC-delta-1) (Phospholipase C-delta-1) (PLC-III)	down	c	0.0198
ENST00000229390	[3]	Splicing factor, arginine/serine-rich 9 (Pre-mRNA-splicing factor SRp30C)	down	a	0.0005
ENST00000340600	[3]	Suppressor of cytokine signaling 2 (SOCS-2) (Cytokine-inducible SH2 protein 2) (CIS-2) (STAT-induced STAT inhibitor 2) (SSI-2)	down	b	0.0056
ENST00000288207	[3]	G2/mitotic-specific cyclin-B2	up	a,b,c,d	0.0010
ENST00000264161	[3]	Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate-	up	c	0.0121

		-tRNA ligase) (AspRS)			
ENST00000309268	[3]	Elongation factor 1-alpha 1 (EF-1-alpha-1) (Elongation factor 1 A-1) (eEF1A-1) (Elongation factor Tu) (EF-Tu)	up	a	0.0252
ENST00000319974	[3]	no description (ets variant gene 4 (E1A enhancer binding protein, E1AF))	up	a	0.0025
ENST00000302068	[3]	Fibrinogen beta chain precursor [Contains: Fibrinopeptide B]	up	a,c	0.0053
ENST00000341048	[3]	no description (interleukin 6 signal transducer (gp130, oncostatin M receptor))	up	a	0.0004
ENST00000296585	[3]	Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa) (Collagen receptor) (VLA-2 alpha chain) (CD49b antigen)	up	d	0.0088
ENST00000260302	[3]	Collagenase 3 precursor (EC 3.4.24.-) (Matrix metalloproteinase-13) (MMP-13)	up	b	0.0197
ENST00000296930	[3]	Nucleophosmin (NPM) (Nucleolar phosphoprotein B23) (Numatrin) (Nucleolar protein NO38)	up	a	0.0015
ENST00000216392	[3]	Glycogen phosphorylase, liver form (EC 2.4.1.1)	up	b,c,d	0.0467
ENST00000370321	[3]	60S ribosomal protein L5	up	c	0.0032
ENST00000265361	[3]	Semaphorin-3C precursor (Semaphorin E) (Sema E)	up	d	8.05E-05
ENST00000244520	[3]	U1 small nuclear ribonucleoprotein C (U1 snRNP protein C) (U1C protein) (U1-C)	up	a	0.0001
ENST00000273258	[4]	PRA1 family protein 3 (ARL-6-interacting protein 5) (ADP-ribosylation- like factor 6-interacting protein 5) (Aip-5) (Glutamate transporter EAAC1-interacting protein) (GTRAP3-18) (Prenylated Rab acceptor protein 2) (Protein JWa) (Dermal papilla-derived pro	down	b	0.0014
ENST00000323456	[4]	myotubularin related protein 4	down	b	0.0011
ENST00000258428	[4]	DNA repair protein REV1 (EC 2.7.7.-) (Rev1-like terminal deoxycytidyl transferase) (Alpha integrin-binding protein 80) (AIBP80)	down	c	0.0011
ENST00000326361	[4]	Zinc finger protein 639 (Zinc finger protein ZASC1) (Zinc finger protein ANC_2H01)	up	a	0.0015
ENST00000259075	[4]	TRAF family member-associated NF-kappa-B activator (TRAF-interacting protein) (I-TRAF)	up	b,c,d	7.5E-05
ENST00000262462	[4]	Long-chain fatty acid transport protein 6 (Fatty acid transport protein 6) (FATP-6) (Very long-chain acyl-CoA synthetase homolog 1) (VLCSH1) (hVLCS-H1) (Fatty-acid-coenzyme A ligase, very long-chain 2) (Solute carrier family 27 member 6)	up	a,c	6.4E-05
ENST00000307633	[4]	Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase) (HisRS)	up	c	0.0001
ENST00000327304	[4]	Exosome complex exonuclease RRP40 (EC 3.1.13.-) (Ribosomal RNA- processing protein 40) (Exosome component 3) (p10)	up	b,c,d	0.0001
ENST00000370986	[4]	Growth arrest and DNA-damage-inducible protein GADD45 alpha (DNA- damage-inducible transcript	up	a	0.0013

		1) (DDIT1)			
ENST00000160827	[4]	Kinesin-like protein KIF22 (Kinesin-like DNA-binding protein) (Kinesin-like protein 4)	up	a	0.0031
ENST00000230588	[5]	Meprin A subunit alpha precursor (EC 3.4.24.18) (Endopeptidase-2) (N- benzoyl-L-tyrosyl-P-amino-benzoic acid hydrolase subunit alpha) (PABA peptide hydrolase) (PPH alpha)	down	a	0.0012
ENST00000162749	[5]	Tumor necrosis factor receptor superfamily member 1A precursor (p60) (TNF-R1) (TNF-RI) (TNFR-I) (p55) (CD120a antigen) [Contains: Tumor necrosis factor receptor superfamily member 1A, membrane form; Tumor necrosis factor-binding protein 1 (TBPI)]	down	b	0.0023
ENST00000314355	[5]	Cyclin-dependent kinases regulatory subunit 2 (CKS-2)	up	a	0.0136
ENST00000283646	[6]	Ribose-5-phosphate isomerase (EC 5.3.1.6) (Phosphoriboisomerase)	down	a	4.2E-05
ENST00000356245	[6]	Ras-GTPase-activating protein-binding protein 1 (EC 3.6.1.-) (ATP- dependent DNA helicase VIII) (GAP SH3-domain-binding protein 1) (G3BP- 1) (HDH-VIII)	up	b	4.15E-05

[1] Ensembl transcript accession for putative hsa-mir-548 target genes

[2] Publication where the genes involvement in colorectal cancer was originally reported

[3] Name and brief description of the gene

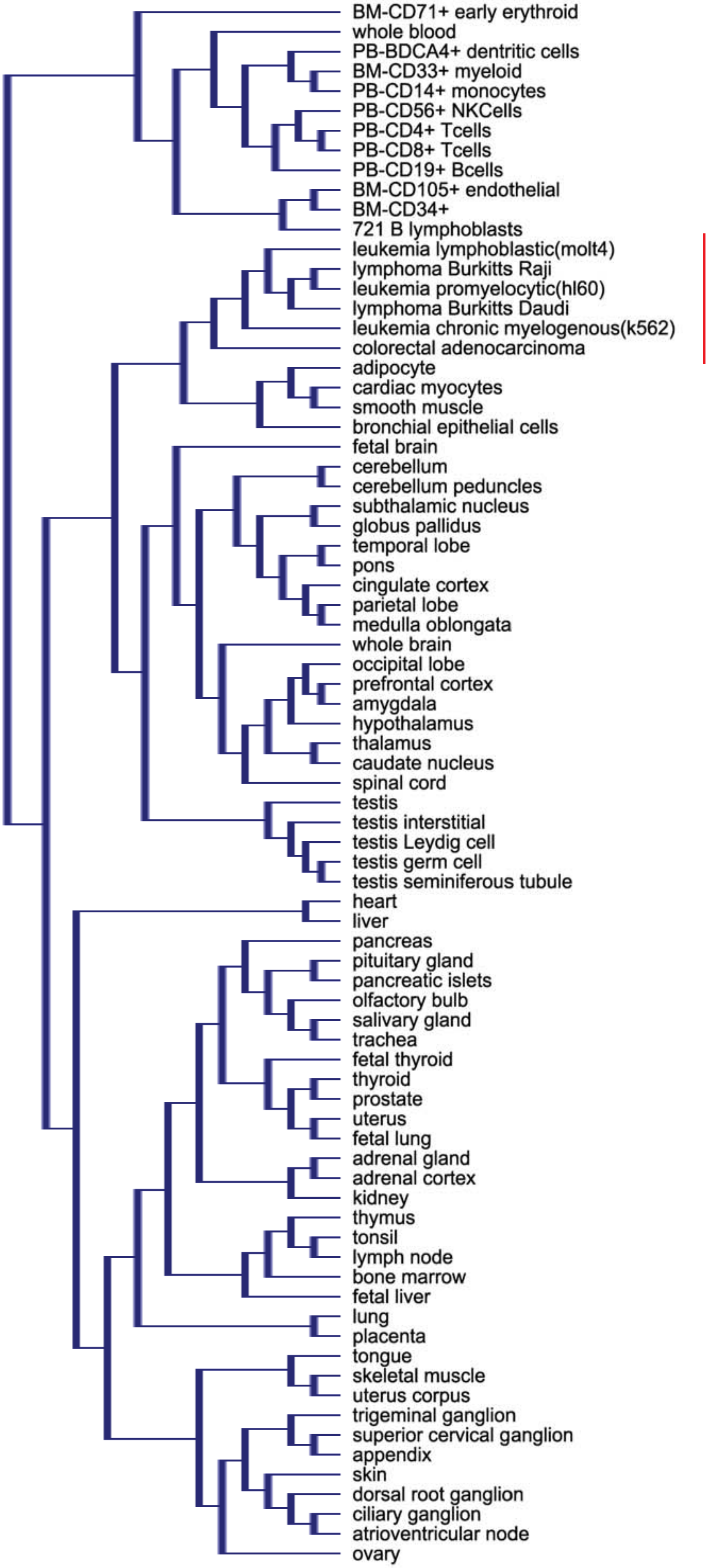
[4] Expression status of the gene (up- or down-regulated) in colorectal cancer relative to normal tissue

[5] Paralog-specific hsa-mir-548 target site

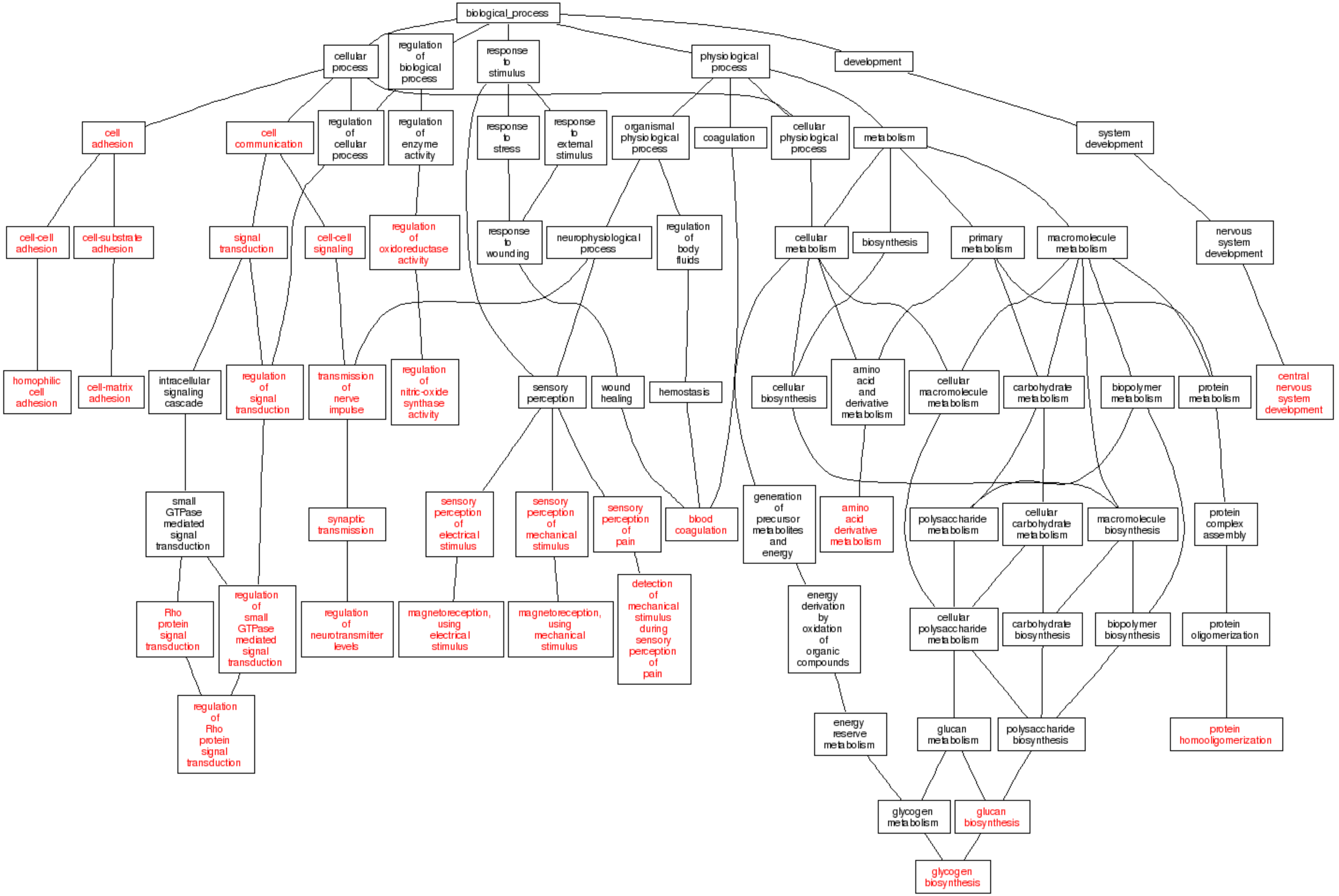
[6] *P*-value associated with the hsa-mir-548 target sites

1. Takemasa I, Higuchi H, Yamamoto H, Sekimoto M, Tomita N, et al. (2001) Construction of preferential cDNA microarray specialized for human colorectal carcinoma: molecular sketch of colorectal cancer. Biochem Biophys Res Commun 285: 1244-1249.
2. Kitahara O, Furukawa Y, Tanaka T, Kihara C, Ono K, et al. (2001) Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia. Cancer Res 61: 3544-3549.
3. Bertucci F, Salas S, Eysteries S, Nasser V, Finetti P, et al. (2004) Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters. Oncogene 23: 1377-1391.
4. Kwon HC, Kim SH, Roh MS, Kim JS, Lee HS, et al. (2004) Gene expression profiling in lymph node-positive and lymph node-negative colorectal cancer. Dis Colon Rectum 47: 141-152.
5. Notterman DA, Alon U, Sierk AJ, Levine AJ (2001) Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. Cancer Res 61: 3124-3130.

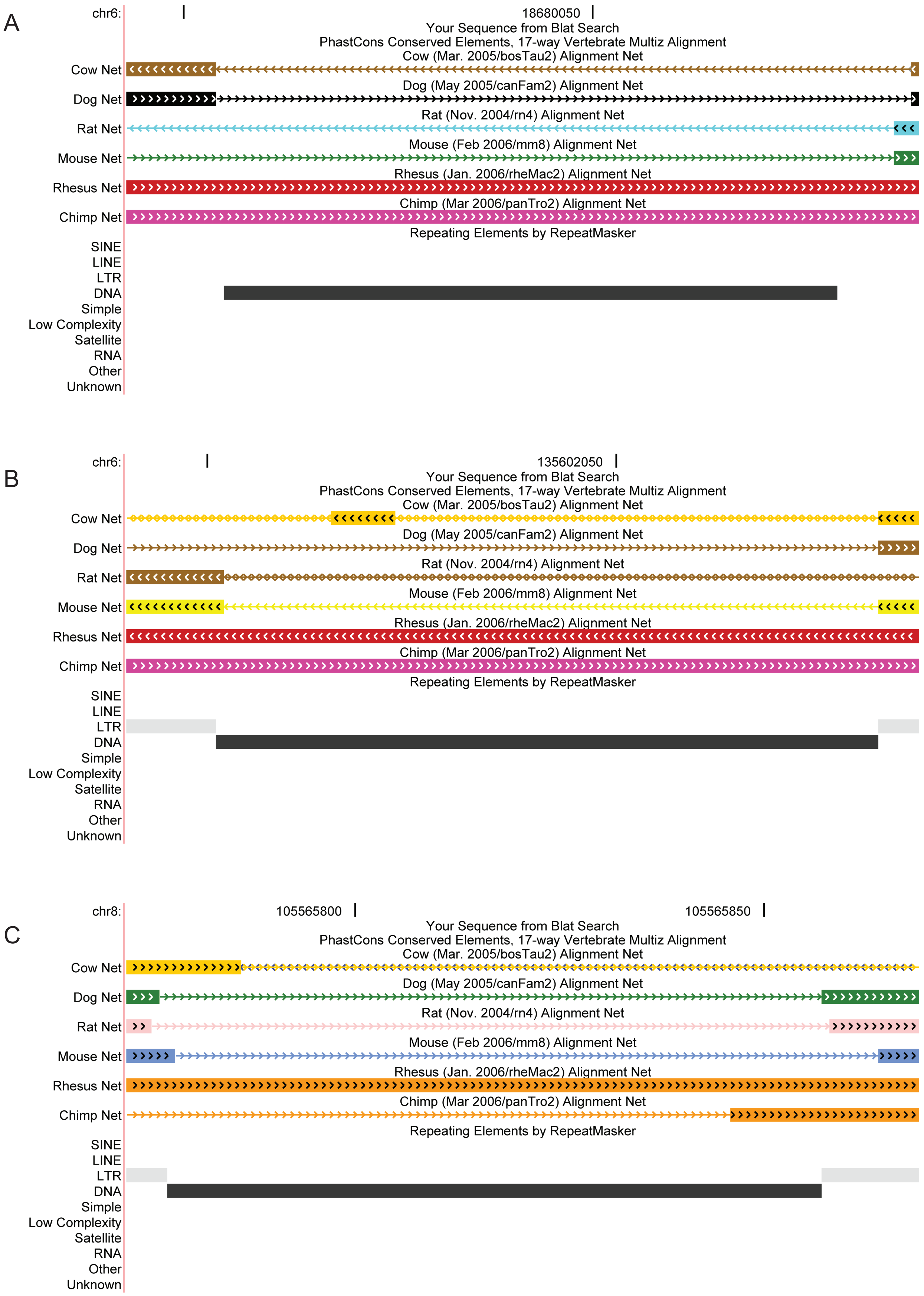
6. Shih W, Chetty R, Tsao MS (2005) Expression profiling by microarrays in colorectal cancer (Review). Oncol Rep 13: 517-524.



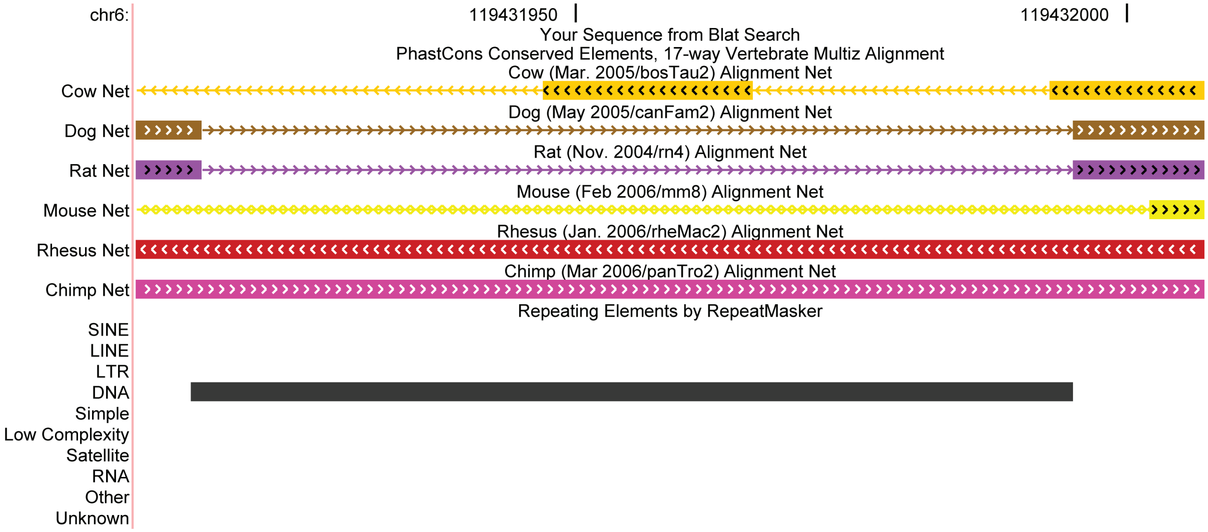
Supplementary Figure 1. **Dendogram showing relationships among tissues from the Novartis Foundation Symatlas microarray dataset.** Cancer tissues are indicated with the red bar.



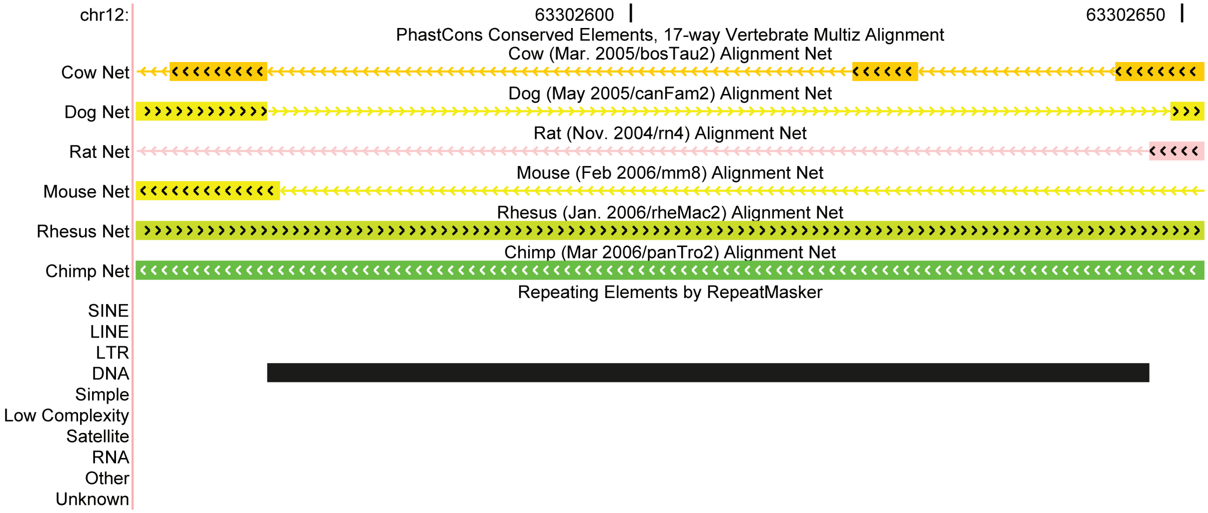
Supplementary Figure 2. **Over-represented GO biological process categories among genes with miRanda predicted hsa-mir-548 target sites that map to colorectal cancer down-regulated co-expression clusters (*i.e.* 12, 15 & 20 in Figure 6).** The portion of the directed acyclic graph (DAG) containing all paths from the root biological process term to the over-represented functional category terms is shown. Over-represented functional categories are indicated in red.

A
chr6:
18680050
Your Sequence from Blat Search
PhastCons Conserved Elements, 17-way Vertebrate Multiz Alignment
Cow (Mar. 2005/bosTau2) Alignment Net
Cow Net
Dog (May 2005/canFam2) Alignment Net
Dog Net
Rat (Nov. 2004/rn4) Alignment Net
Rat Net
Mouse (Feb 2006/mm8) Alignment Net
Mouse Net
Rhesus (Jan. 2006/rheMac2) Alignment Net
Rhesus Net
Chimp (Mar 2006/panTro2) Alignment Net
Chimp Net
Repeating Elements by RepeatMasker
SINE
LINE
LTR
DNA
Simple
Low Complexity
Satellite
RNA
Other
Unknown
B
chr6:
135602050
Your Sequence from Blat Search
PhastCons Conserved Elements, 17-way Vertebrate Multiz Alignment
Cow (Mar. 2005/bosTau2) Alignment Net
Cow Net
Dog (May 2005/canFam2) Alignment Net
Dog Net
Rat (Nov. 2004/rn4) Alignment Net
Rat Net
Mouse (Feb 2006/mm8) Alignment Net
Mouse Net
Rhesus (Jan. 2006/rheMac2) Alignment Net
Rhesus Net
Chimp (Mar 2006/panTro2) Alignment Net
Chimp Net
Repeating Elements by RepeatMasker
SINE
LINE
LTR
DNA
Simple
Low Complexity
Satellite
RNA
Other
Unknown
C
chr8:
105565800
105565850
Your Sequence from Blat Search
PhastCons Conserved Elements, 17-way Vertebrate Multiz Alignment
Cow (Mar. 2005/bosTau2) Alignment Net
Cow Net
Dog (May 2005/canFam2) Alignment Net
Dog Net
Rat (Nov. 2004/rn4) Alignment Net
Rat Net
Mouse (Feb 2006/mm8) Alignment Net
Mouse Net
Rhesus (Jan. 2006/rheMac2) Alignment Net
Rhesus Net
Chimp (Mar 2006/panTro2) Alignment Net
Chimp Net
Repeating Elements by RepeatMasker
SINE
LINE
LTR
DNA
Simple
Low Complexity
Satellite
RNA
Other
Unknown

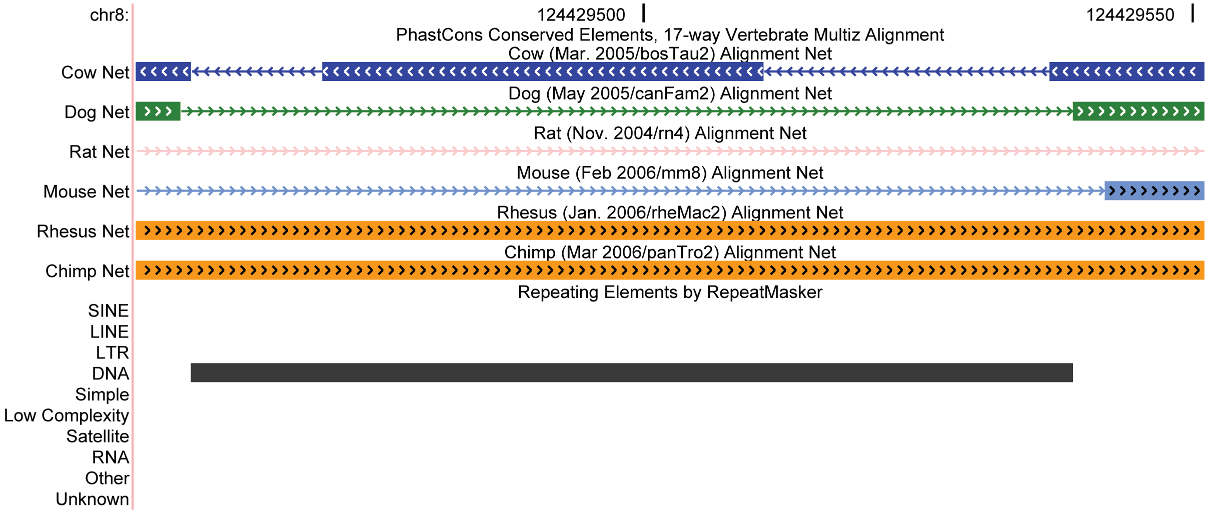
D


chr6:
119431950
119432000
Your Sequence from Blat Search
PhastCons Conserved Elements, 17-way Vertebrate Multiz Alignment
Cow (Mar. 2005/bosTau2) Alignment Net
Cow Net
Dog (May 2005/canFam2) Alignment Net
Dog Net
Rat (Nov. 2004/rn4) Alignment Net
Rat Net
Mouse (Feb 2006/mm8) Alignment Net
Mouse Net
Rhesus (Jan. 2006/rheMac2) Alignment Net
Rhesus Net
Chimp (Mar 2006/panTro2) Alignment Net
Chimp Net
Repeating Elements by RepeatMasker
SINE
LINE
LTR
DNA
Simple
Low Complexity
Satellite
RNA
Other
Unknown

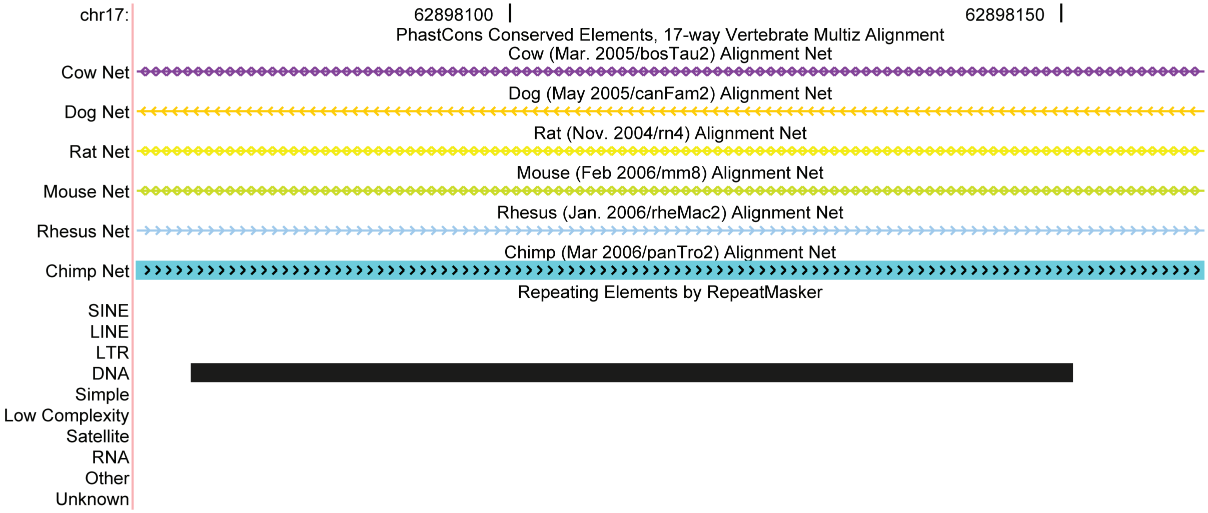
E


chr12:
63302600
63302650
PhastCons Conserved Elements, 17-way Vertebrate Multiz Alignment
Cow (Mar. 2005/bosTau2) Alignment Net
Cow Net
Dog (May 2005/canFam2) Alignment Net
Dog Net
Rat (Nov. 2004/rn4) Alignment Net
Rat Net
Mouse (Feb 2006/mm8) Alignment Net
Mouse Net
Rhesus (Jan. 2006/rheMac2) Alignment Net
Rhesus Net
Chimp (Mar 2006/panTro2) Alignment Net
Chimp Net
Repeating Elements by RepeatMasker
SINE
LINE
LTR
DNA
Simple
Low Complexity
Satellite
RNA
Other
Unknown

F


chr8:
124429500
124429550
PhastCons Conserved Elements, 17-way Vertebrate Multiz Alignment
Cow (Mar. 2005/bosTau2) Alignment Net
Cow Net
Dog (May 2005/canFam2) Alignment Net
Dog Net
Rat (Nov. 2004/rn4) Alignment Net
Rat Net
Mouse (Feb 2006/mm8) Alignment Net
Mouse Net
Rhesus (Jan. 2006/rheMac2) Alignment Net
Rhesus Net
Chimp (Mar 2006/panTro2) Alignment Net
Chimp Net
Repeating Elements by RepeatMasker
SINE
LINE
LTR
DNA
Simple
Low Complexity
Satellite
RNA
Other
Unknown

G



Supplementary Figure 3. **Made1-derived miRNA genes are primate-specific.** Human genomic regions corresponding to Made1-derived miRNA genes are shown: A hsa-mir-548a-1, B hsa-mir-548-a2, C hsa-mir-548-a3, D hsa-mir-548-b, E hsa-mir-548-c, F-hsa-mir-548-d1, G-hsa-mir-548-d2. The UCSC Genome Browser is used to show the location of the Made1 elements (DNA) in the RepeatMasker track. Evolutionary comparisons between the human genome and the corresponding regions in the chimp, rhesus, mouse, rat, dog and cow genomes are shown using the species-specific Net tracks of the Genome Browser. Corresponding Made1 orthologous regions that are present in another species are indicated with a broad line, while regions that are missing in another species are shown with a thin line.