

**Table S1. List of identification tags and descriptions**

Identification Tag name	No. of references	No. of differentially expressed features	Identification Tag description
<i>11q23</i>	5	559	11q23 abnormalities
<i>abnormal cytogenetics</i>	1	150	abnormal cytogenetics
<i>adult</i>	12	7254	greater than or equal to 18 years of age
<i>ALL</i>	3	212	acute lymphocytic leukemia
<i>AML</i>	25	15697	acute myeloid leukemia
<i>aneuploid</i>	1	64	aneuploid chromosome number
<i>CBF</i>	2	2410	core binding factor translocation
<i>CD34+CD38- fraction</i>	1	411	CD34+CD38- cell fraction from AML
<i>CD34+CD38+ fraction</i>	1	411	CD34+CD38+ cell fraction from AML
<i>CEBPA mutation</i>	2	402	<i>CEBPA</i> gene mutation
<i>CEBPA silenced</i>	1	304	<i>CEBPA</i> gene silenced
<i>control</i>	2	55	CD34+ cells from normal patient
<i>del(7q)</i>	1	48	deletion of long arm of chromosome 7
<i>euploid</i>	1	64	euploid chromosome number
<i>EVI1 overexpression</i>	1	28	<i>EVI1</i> gene overexpression
<i>FAB</i>	1	213	Associated with FAB, unspecified subtype
<i>FAB-M3</i>	1	20	FAB subtype M3
<i>FAB-M3v</i>	1	20	FAB subtype M3v
<i>FAB-M4</i>	1	40	FAB subtype M4
<i>FAB-M5</i>	1	40	FAB subtype M5
<i>FAB-M7</i>	1	100	FAB subtype M7
<i>FLT3 mutation</i>	4	663	<i>FLT3</i> gene mutation
<i>FLT3-ALM</i>	1	23	<i>FLT3</i> gene activating loop mutations
<i>FLT3-ITD</i>	4	355	<i>FLT3</i> gene internal tandem duplications in juxtamembrane region
<i>FLT3-TKD</i>	1	84	<i>FLT3</i> gene point mutation in second tyrosine kinase domain
<i>good prognosis</i>	11	2177	increased overall survival or disease free survival or response to therapy
<i>high centrosome aberrations</i>	1	21	high number of abnormal centrosomes
<i>high RUNX3-ATRX ratio</i>	1	13	high <i>RUNX3</i> gene to <i>ATRX</i> gene ratio
<i>human</i>	25	15809	human sample
<i>inv(16)</i>	9	1748	<i>inv(16)</i>
<i>low centrosome aberrations</i>	1	21	low number of abnormal centrosomes
<i>low RUNX3-ATRX ratio</i>	1	13	low <i>RUNX3</i> gene to <i>ATRX</i> gene ratio
<i>microarray</i>	25	15686	microarray experiment
<i>MLL fusion gene</i>	2	944	<i>MLL</i> gene chimeric fusions
<i>MLL-PTD</i>	1	19	<i>MLL</i> gene partial tandem duplications
<i>monocytic</i>	1	50	monocytic differentiation
<i>normal cytogenetics</i>	6	925	cytogenetically normal (CN) AML
<i>NPM1 mutation</i>	5	1802	<i>NPM1</i> gene mutation
<i>NRAS-PM</i>	1	67	<i>NRAS</i> gene mutation
<i>patient</i>	25	15809	patient sample
<i>pediatric</i>	5	4297	less than 18 years of age
<i>poor prognosis</i>	12	2336	decreased overall survival or disease free survival or response to therapy
<i>prognosis</i>	2	138	associated with overall survival or disease free survival or response to therapy
<i>RT-PCR</i>	7	123	RT-PCR experiment
<i>t(11;19)</i>	1	10	<i>t(11;19)</i>
<i>t(15;17)</i>	9	401	<i>t(15;17)</i>
<i>t(8;14)</i>	1	15	<i>t(8;14)</i>
<i>t(8;21)</i>	9	1671	<i>t(8;21)</i>
<i>t(9;11)</i>	1	144	<i>t(9;11)</i>
<i>t(9;22)</i>	1	16	<i>t(9;22)</i>
<i>T-ALL</i>	1	15	T-cell acute lymphocytic leukemia
<i>trisomy 8</i>	1	14	trisomy 8

Definitions of identification tags with associated studies and number of features.