PATIENTS AND CLINICAL RISK FACTORS

For this study, pre-treatment cryopreserved leukemia specimens were available on a representative cohort of 207 of the 272 (76%) patients registered to COG P9906; the clinical and outcome parameters of these 207 patients did not differ significantly from all 272 patients (see Table S1 and Figure S1). As shown in Table S1 and Figure S1, the differences in various characteristics between the entire group (n=272) and the present study cohort (n=207) were examined by the statistical comparisons between the present study cohort and remaining patients (n=65) not included in the present study. Each P-value in Table S1 and Figure S1 is that of the individual test which needs to be adjusted for multiple testing. A simple Bonferroni adjustment multiplies the P-values by the total number of tests (10). After this adjustment, none of the characteristics are significantly different between the entire group and the cohort examined herein, except the test for WBC count when a cutoff value was considered.

Table S1. Comparison of HR-ALL Patients Registered to COG P9906 (n=272) and The Subset of Patients Examined and Modeled for Gene Expression Signatures (n=207)[1]

Characteristics	Not Studied		Studied		Total		p-value (Fisher's exact test)
	N	%	N	%	N	%	
Age							
≥ 10 Yrs	51	78.46	132	63.77	183	67.28	.0335
< 10 Yrs	14	21.54	75	26.23	89	32.72	
Sex							
Male	52	80	137	66.18	189	69.49	.0442
Female	13	20	70	33.82	83	30.51	
WBC (x 1000/µL)							
< 50	52	80	99	47.83	151	55.51	< .0001
≥ 50	13	20	108	52.17	121	44.49	
Race							
Hispanic or Latino	15	23.08	51	24.64	66	24.26	.9638
Others	47	72.31	154	74.39	201	73.90	
Unknown	3	4.61	2	0.97	5	1.84	
MRD at day 29							
Negative	40	61.54	124	59.90	164	60.29	.7550
Positive	19	29.23	67	32.37	86	31.62	
Unknown	6	9.23	16	7.73	22	8.09	
MLL							
Negative	61	93.85	186	89.86	247	90.81	.4617
Positive	4	6.15	21	10.15	25	9.19	
TCF3/PBX1							
Negative	59	90.77	184	88.89	243	89.34	.6384
Positive	5	7.69	23	11.11	28	10.29	
Unknown	1	1.54	0	0	1	0.37	
CNS							
No blasts	54	83.08	160	77.29	214	78.68	.1009
< 5 blasts	3	4.61	26	12.56	29	10.66	
≥ 5 blasts	8	12.31	21	10.15	29	10.66	
Total	65	100	207	100	272	100	

[1] All unknown data were removed before statistical tests were performed.

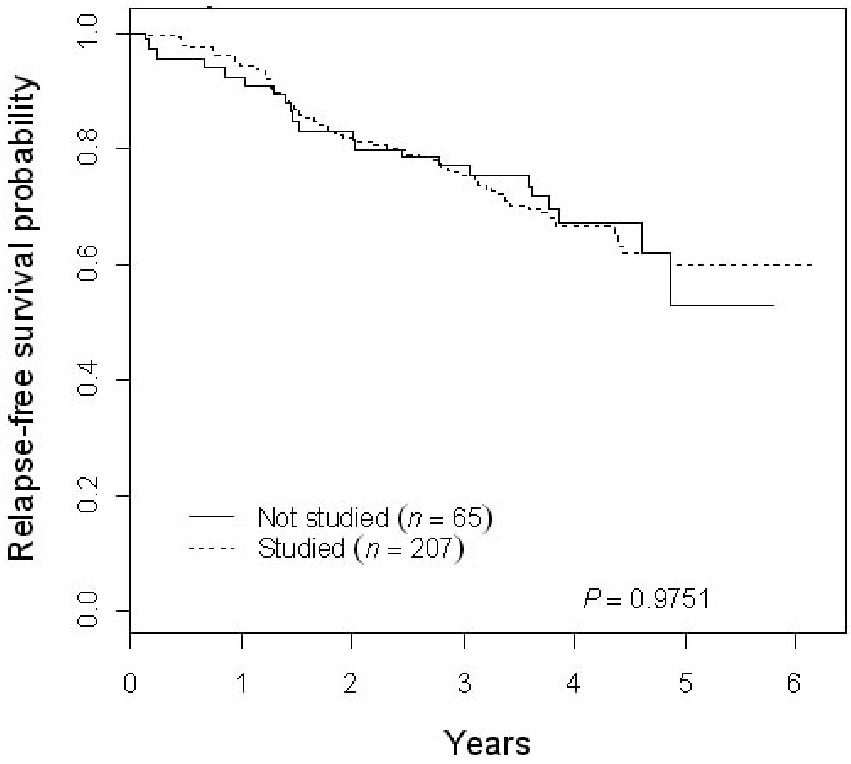

Figure S1. Comparison of relapse free survival between those studied (n=207) and remaining COG P9906 patients not included in this cohort (n=65).

The 207 patient cohort had slight male predominance (66%) and included a subset (23%, 47/201) with blasts in the CNS at diagnosis (CNS2+CNS3). Approximately 35% of the 191 specimens evaluated by flow cytometry on day 29 of induction therapy had subclinical MRD (>0.01% blasts).[3] As shown in Table S2, only MRD at the end of induction therapy and increasing WBC count were significantly associated with decreased relapse free survival (RFS). The significant effect of WBC count as a continuous variable on decreased RFS was no longer seen when the cutoff of 50 K/µL was applied (see Section 7). A trend towards declining RFS was also observed among the 25% of children with Hispanic/Latino ethnicity contained within this cohort. In multivariate analysis, both MRD and WBC count retained significance when adjusted for one another (likelihood ratio test based on COX regression, *P* < .001).

Table S2. Association of Relapse Free Survival with Clinical and Genetic Features in the High Risk ALL Cohort

Characteristic		Association with Relapse Free Survival	
		Hazard Ratio	p-value
Age			
≥ 10 Yrs	132	1	
< 10 Yrs	75	1.152	.561
Age			
Median	13.1 yrs		
Range	1 – 20	.995	.817
Sex			
Male	137	1	
Female	70	0.769	.320
WBC (x 1000/µL)			
Median	62.3		
Range	1 – 959	1.003	< .001
MRD at Day 29			
Negative	124	1	
Positive	67	2.805	< .001
Ethnicity			
Hispanic or Latino	51	1.644	.049
Others	154	1	
MLL			
Positive	21	1.061	.881
Negative	186	1	
TCF3/PBX1			
Positive	23	.704	.409
Negative	184	1	
CNS			
No blasts	160	1	
< 5 blasts	26	0.897	.708
≥ 5 blasts	21		

VALIDATION COHORT

A subset of patients from COG CCG 1961 "Treatment of Patients with Acute Lymphoblastic Leukemia with Unfavorable Features" was used as a validation cohort to determine whether similar clusters were present in a different set of high-risk patients. As described in Bhojwani *et al.*,[4] COG CCG 1961 enrolled a total of 2078 patients with NCI high risk features, i.e. WBC count ≥ 50,000/μL or age ≥10 years old, from September 1996 to May 2002. Microarray data from these 99 patients were analyzed using the methods described in this paper.

DATA PROCESSING

A. Microarray Preparation and Scanning

After RNA quantification, cDNA preparation, and labeling, biotinylated cRNA was fragmented and hybridized to HG_U133_Plus2.0 oligonucleotide microarrays (Affymetrix, Santa Clara, CA) containing 54,675 probe sets. Signals were scanned (Affymetrix GeneChip Scanner) and analyzed with the Affymetrix Microarray Suite (MAS 5.0). Signal intensities and expression data were generated with the Affymetrix GCOS 1.4 software package.

B. Microarray Data Masking

Prior to any intensity analysis, the microarray data were first masked to remove those probes found to be uninformative in a majority of the samples. Removal of these probe pairs improves the overall quality of the data and eliminates many non-specific signals that are shared by a particular sample type (i.e., cross-hybridizing messages present in blood and marrow samples). Each probe pair (across all 207 samples) was evaluated and masked if the mismatch (MM) was greater than the perfect match (PM) in more than 60% of the samples. This mask removed 94,767 probe pairs (15.7% of the 604,258) and had some impact on 38,588 probe sets (71%). As shown in Table S3, the net impact of masking was a significant increase in the number of present calls paralleled by a decrease in the number of absent calls.

The mask removed only seven probe sets (0.01% of the 54,675), all of which represented non-human control genes.

Table S3. Impact of Masking on Affymetrix Statistical Calls (Reported as Percentage of Total Probes: 54,675 raw; 54,668 masked).

	Present	**Marginal**	**Absent**	**No call**
Raw	34.9	1.7	63.3	0
Masked	48.0	3.1	48.9	0 (7)

C. Microarray Data Filtering

Prior to any clustering, the data were filtered to remove probe sets deemed to be unrelated to disease: genes from sex-determining regions of X and Y (which simply correlate with sex), spiked control genes and globin genes (presumed to arise from contaminating normal blood cells). All filtered probe sets were selected based upon their gene symbols or chromosomal location. Table S4 lists the 89 probe sets mapped within sex-determining regions. These include the XIST gene from chromosome X and probe sets from Yp11-Yq11. All probe sets from PAR1 and PAR2 regions of both sex chromosomes are retained. Table S5 lists the 62 Affymetrix spiked control genes. Table S6 lists the twenty excluded globin probe sets with a gene symbol beginning with "HB" and the word "globin" contained within the gene title. After the filtering of these probe sets 54,504 were available for clustering.

Table S4. X- and Y- Specific Transcripts Excluded from the Analysis (89)

Probe Set ID	Gene Symbol	Cytoband
214218_s_at	XIST	Xq13.2
221728_x_at	XIST	Xq13.2
224588_at	XIST	Xq13.2
224589_at	XIST	Xq13.2
224590_at	XIST	Xq13.2
227671_at	XIST	Xq13.2
243712_at	XIST	Xq13.2
201909_at	LOC100133662 /// RPS4Y1	Yp11.3
204409_s_at	EIF1AY	Yq11.222
204410_at	EIF1AY	Yq11.222

205000_at	DDX3Y	Yq11
205001_s_at	DDX3Y /// LOC100130220	Yq11
206279_at	PRKY	Yp11.2
206624_at	LOC100130216 /// USP9Y	Yq11.2
206700_s_at	JARID1D	Yq11|Yq11
206769_at	LOC100130227 /// TMSB4Y	Yq11.221
207063_at	CYorf14	Yq11.222
207246_at	LOC100130829 /// ZFY	Yp11.3
207646_s_at	CDY1 /// CDY1B /// CDY2A /// CDY2B	Yq11.221 /// Yq11.223 /// Yq11.23
207647_at	CDY1	Yq11.23
207703_at	NLGN4Y	Yq11.221
207893_at	LOC100130809 /// SRY	Yp11.3
207909_x_at	DAZ1 /// DAZ2 /// DAZ3 /// DAZ4 /// LOC732447	Yq11.223
207912_s_at	DAZ1 /// DAZ2 /// DAZ3 /// DAZ4 /// LOC732447	Yq11.223
207916_at	RBMY1E	Yq11.223
207918_s_at	LOC728137 /// LOC728395 /// LOC728412 /// TSPY1	Yp11.2
208067_x_at	LOC100130224 /// UTY	Yq11
208220_x_at	AMELY	Yp11.2
208281_x_at	DAZ1 /// DAZ2 /// DAZ3 /// DAZ4 /// LOC732447	Yq11.223
208282_x_at	DAZ1 /// DAZ2 /// DAZ3 /// DAZ4 /// LOC732447	Yq11.223
208307_at	RBMY1A1 /// RBMY1B /// RBMY1D /// RBMY1E /// RBMY1F /// RBMY1J /// RBMY3AP	Yp11.2 /// Yq11.223
208331_at	BPY2	Yq11
208332_at	PRY /// PRY2	Yq11.223
208339_at	XKRY /// XKRY2	Yq11.221
210322_x_at	UTY	Yq11
211149_at	LOC100130224 /// UTY	Yq11
211227_s_at	PCDH11Y	Yp11.2
211460_at	TTTY9A /// TTTY9B	Yq11.221 /// Yq11.222
211461_at	CSPG4LYP1 /// CSPG4LYP2	Yq11.223 /// Yq11.23
211462_s_at	TBL1Y	Yp11.2
214131_at	CYorf15B	Yq11.222

214983_at	TTTY15	Yq11.1
216351_x_at	DAZ1 /// DAZ2 /// DAZ3 /// DAZ4 /// LOC732447	Yq11.223
216374_at	LOC728137 /// LOC728395 /// LOC728412 /// TSPY1	Yp11.2
216544_at	RBMY2FP	Yq11.223
216665_s_at	TTTY2	Yp11.2
216673_at	LOC100101116 /// TTTY1	Yp11.2
216786_at	LOC159110	Yq11.221
216842_x_at	RBM /// RBMY1A1 /// RBMY1B /// RBMY1D /// RBMY1E /// RBMY1F /// RBMY1H /// RBMY1J /// RBMY3AP	Yp11.2 /// Yq11.223 /// Yq11.23
216922_x_at	DAZ1 /// DAZ2 /// DAZ3 /// DAZ4 /// LOC732447	Yq11.223
217049_x_at	PCDH11Y	Yp11.2
217160_at	TSPY1	Yp11.2
217261_at	LOC100101117 /// TTTY2	Yp11.2
222229_x_at	LOC441533	Yp11.2
223645_s_at	CYorf15B	Yq11.222
223646_s_at	CYorf15B	Yq11.222
224003_at	TTTY14	Yq11.222
224007_at	HSFY1 /// HSFY2	Yq11.222
224040_at	TTTY5	Yq11.223
224041_at	TTTY6	Yq11.223
224052_at	HSFY1 /// HSFY2	Yq11.222
224142_s_at	LOC100101118 /// TTTY8	Yp11.2
224143_at	LOC100101118 /// TTTY8	Yp11.2
224174_at	TTTY11	Yp11.2
224195_at	TTTY12	Yp11.2
224292_at	TTTY13	Yq11.223
224293_at	TTTY10	Yq11.221
228492_at	LOC100130216 /// USP9Y	Yq11.2
230760_at	LOC100130829 /// ZFY	Yp11.3
232618_at	CYorf15A	Yq11.222
233151_s_at	TTTY7	Yp11.2
233178_at	TGIF2LY	Yp11.2
234309_at	TTTY7	Yp11.2
234715_at	GOLGA2LY1 /// GOLGA2LY2	Yq11.223
234913_at	TTTY4 /// TTTY4B /// TTTY4C	Yq11.2 /// Yq11.223
234931_at	AYP1p1	Yp11.31

235941_s_at	LOC159110 /// LOC401629 /// LOC401630	Yq11.221
235942_at	LOC401629 /// LOC401630	Yq11.221
236694_at	CYorf15A	Yq11.222
1552952_at	RBMY2FP	Yq11.223
1554125_a_at	NLGN4Y	Yq11.221
1561185_at	TTTY7	Yp11.2
1561390_at	FAM41AY	Yq11.221
1562313_at	BCORL2	Yq11.222
1563420_at	XGPY2	Yp11.31
1565132_at	RBMY3AP	Yp11.2
1565320_at	RBMY3AP	Yp11.2
1570359_at	DDX3Y	Yq11
1570360_s_at	DDX3Y /// LOC100130220	Yq11

Table S5. AFFX Probe Sets Excluded from the Analysis (62)

Probe Set ID
AFFX-BioB-5_at
AFFX-BioB-M_at
AFFX-BioB-3_at
AFFX-BioC-5_at
AFFX-BioC-3_at
AFFX-BioDn-5_at
AFFX-BioDn-3_at
AFFX-CreX-5_at
AFFX-CreX-3_at
AFFX-DapX-5_at
AFFX-DapX-M_at
AFFX-DapX-3_at
AFFX-LysX-5_at
AFFX-LysX-M_at
AFFX-LysX-3_at
AFFX-PheX-5_at
AFFX-PheX-M_at
AFFX-PheX-3_at
AFFX-ThrX-5_at
AFFX-ThrX-M_at
AFFX-ThrX-3_at
AFFX-TrpnX-5_at
AFFX-TrpnX-M_at

AFFX-TrpnX-3_at
AFFX-r2-Ec-bioB-5_at
AFFX-r2-Ec-bioB-M_at
AFFX-r2-Ec-bioB-3_at
AFFX-r2-Ec-bioC-5_at
AFFX-r2-Ec-bioC-3_at
AFFX-r2-Ec-bioD-5_at
AFFX-r2-Ec-bioD-3_at
AFFX-r2-P1-cre-5_at
AFFX-r2-P1-cre-3_at
AFFX-r2-Bs-dap-5_at
AFFX-r2-Bs-dap-M_at
AFFX-r2-Bs-dap-3_at
AFFX-r2-Bs-lys-5_at
AFFX-r2-Bs-lys-M_at
AFFX-r2-Bs-lys-3_at
AFFX-r2-Bs-phe-5_at
AFFX-r2-Bs-phe-M_at
AFFX-r2-Bs-phe-3_at
AFFX-r2-Bs-thr-3_s_at
AFFX-r2-Bs-thr-M_s_at
AFFX-r2-Bs-thr-5_s_at
AFFX-HUMISGF3A/M97935_5_at
AFFX-HUMISGF3A/M97935_MA_at
AFFX-HUMISGF3A/M97935_MB_at
AFFX-HUMISGF3A/M97935_3_at
AFFX-HUMRGE/M10098_5_at
AFFX-HUMRGE/M10098_M_at
AFFX-HUMRGE/M10098_3_at
AFFX-HUMGAPDH/M33197_5_at
AFFX-HUMGAPDH/M33197_M_at
AFFX-HUMGAPDH/M33197_3_at
AFFX-HSAC07/X00351_5_at
AFFX-HSAC07/X00351_M_at
AFFX-HSAC07/X00351_3_at
AFFX-M27830_5_at
AFFX-M27830_M_at
AFFX-M27830_3_at
AFFX-hum_alu_at

Table S6. Globin Probe Sets Excluded from the Analysis (20)

Probe Set ID	Gene Symbol	Cytoband
1562981_at	HBB	11p15.5
204018_x_at	HBA1 /// HBA2	16p13.3
204419_x_at	HBG1 /// HBG2	11p15.5
204848_x_at	HBG1 /// HBG2	11p15.5
205919_at	HBE1	11p15.5
206647_at	HBZ	16p13.3
206834_at	HBD	11p15.5
209116_x_at	HBB	11p15.5
209458_x_at	HBA1 /// HBA2	16p13.3
211696_x_at	HBB	11p15.5
211699_x_at	HBA1 /// HBA2	16p13.3
211745_x_at	HBA1 /// HBA2	16p13.3
213515_x_at	HBG1 /// HBG2	11p15.5
214414_x_at	HBA1 /// HBA2	16p13.3
216036_at	HBBP1	11p15.5
217232_x_at	HBB	11p15.5
217414_x_at	HBA1 /// HBA2	16p13.3
217683_at	HBE1	11p15.5
220807_at	HBQ1	16p13.3
240336_at	HBM	16p13.3

SELECTION OF CLUSTERING PROBE SETS: High CV, ROSE and COPA

A. Selection of High CV Probe Sets

Each of the remaining 54,504 filtered probe sets was ordered by its coefficient of variation (CV = standard deviation/mean). The 254 probe sets with the highest CVs were used for the H clustering.

B. Selection of COPA Probe Sets

The COPA method was applied essentially as described by Tomlins et al.[5] First, the median expression for each probe set was adjusted to zero. Secondly, the median absolute deviation from median (MAD) was calculated and the intensities for each probe set were divided by its MAD. Finally, these MAD-normalized intensities at the 95th percentile were sorted. In order to make the comparison of all clustering methods more comparable, an equal number of probe sets (254) was selected from the top of the sorted list and was used for clustering.

C. Selection of ROSE Probe Sets

ROSE (Recognition of Outlier by Sampling Ends) was developed as an alternative method for outlier detection. In COPA, units of MAD at a fixed point (typically either the 90th or 95th percentile) rank the outliers. This fixed-point threshold confers a size bias for the clusters (higher percentile levels favor smaller groups of outlier signals). More importantly, the ranking of probe sets is by the magnitude of their deviation. Those with the greatest deviations will dominate the top of the list. The potential drawback to this is that larger groups of related samples with outlier signals may be missed if the magnitude of their variance is not extremely high.

In contrast, ROSE applies a single threshold for the magnitude of the deviation and then orders the probe sets by the size of the largest sampled group that satisfies this cutoff. Regardless of the magnitude of the difference from median, all probe sets that satisfy the threshold cutoff and are within the designated size range are considered equal. Details of the ROSE method, as it was applied in this study, follow. The intensity values for each of the 54,504 probe sets were plotted individually in ascending order. The plots were divided into thirds and the intensities from the middle third were used to generate trend lines by least squares fitting. Groups of 2**k* (where *k* is an integer from 2 to one third of the sample size) were sampled from each end of the intensity plots and the median intensities of these groups were compared to the trend lines. The choice of a trend line as the metric, rather than simply median, is meant to reduce the number of probe sets than simply have a high variance, but do not necessarily contain distinct clusters of outlier samples.

Figure S2 illustrates how this is accomplished. Increasing sized groups are sampled from each end until the median intensity of a group fails to exceed the desired threshold. The largest value of *k* at which each probe set surpasses the threshold is recorded. The probe sets are then ordered by their maximum *k* values. In this study a probe set was selected for clustering if k ≥ 6 and the median intensity of the sampled group was at least 7-fold its corresponding point on the trend line. This threshold for *k* was selected in order to enrich for groups in the range of 10 or more members (greater than 5% of the population size). Smaller groups, although still possibly quite interesting, are much less likely to yield statistically significant results. The 7-fold threshold was chosen to minimize the impact of signal noise on probe set selection and also to limit the total number of probe sets to be used for clustering. Only 254 probe sets out of 54,504 (0.5%) satisfied these criteria of 7X threshold and *k* values ≥ 6.

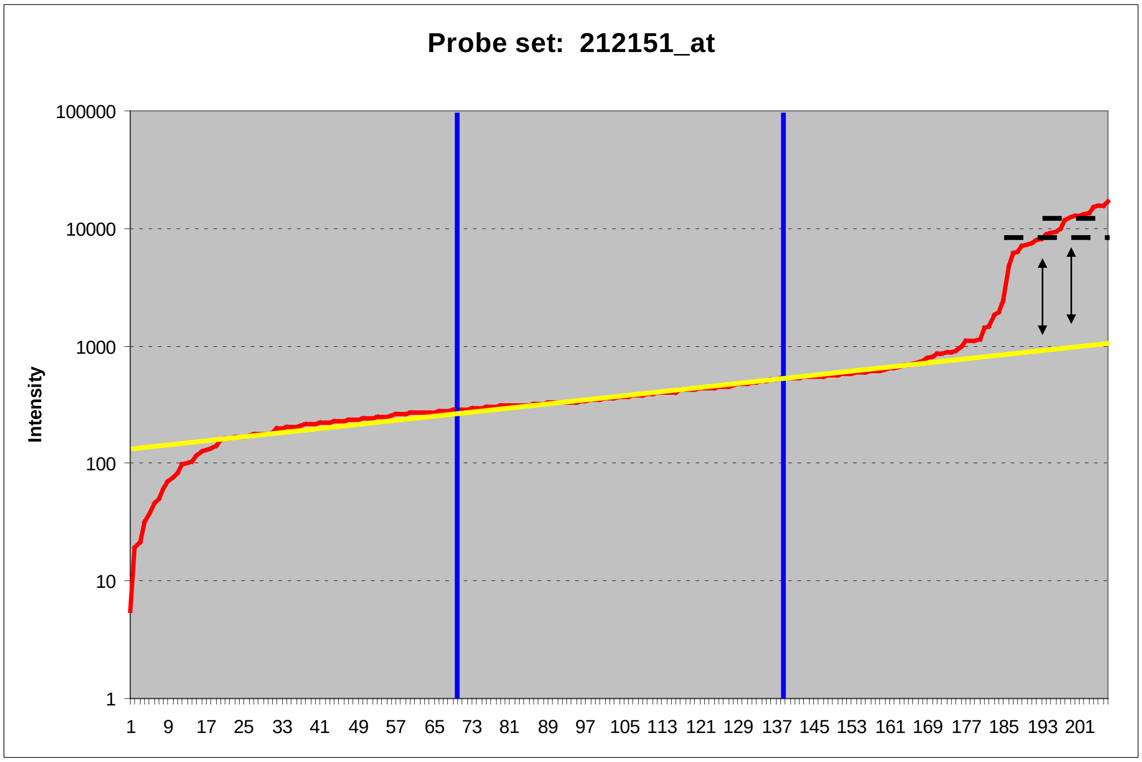

Figure S2. ROSE Methodology. Red line indicates signal intensities for all 207 patient samples for probe 212151_at. Vertical blue lines depict partitioning of samples into thirds. A least-squares curve fit is applied to the middle third of the samples and the resulting trend line is shown in yellow. Different sample groups are illustrated by the dashed lines at the top right. As shown by the double arrowed lines, the median value from each of these groups is compared to the trend line.

D. Outlier Probe Set Selection for CCG 1961 (Validation Cohort)

Masking and filtering was applied to the CCG 1961 data set exactly the same way as in P9906. ROSE used the same 7-fold threshold for intensity and $k \geq 6$. 167 probe sets (0.3% of the 54,504) satisfied these criteria. COPA clustering used the top 167 probe sets at the 95th percentile level. HC used the top 167 probe sets ranked by their CV.

E. Probe Sets Used For Clustering

Table S7A. Probe Sets Used in P9906 and CCG1961
The probe sets common to HC and either COPA or ROSE are shown in bold; those shared between COPA and either HC or ROSE are italicized.

P9906 Probe Sets (254)		
HC	**COPA**	**ROSE**
117_at	38487_at	*38487_at*
1552398_a_at	46665_at	*46665_at*
1553328_a_at	***200800_s_at***	**200799_at**
1553613_s_at	***201105_at***	***200800_s_at***
1554633_a_at	201566_x_at	201012_at
1554892_a_at	201579_at	***201105_at***
1555579_s_at	201656_at	**201215_at**
1557534_at	201669_s_at	*201579_at*
1559477_s_at	***201842_s_at***	*201656_at*
1559696_at	***202178_at***	***201842_s_at***
1559697_a_at	202206_at	***202178_at***
1566772_at	202410_x_at	*202206_at*
200799_at	***202411_at***	202207_at
200800_s_at	***202859_x_at***	202273_at
201105_at	***202917_s_at***	202289_s_at
201215_at	202976_s_at	202336_s_at
201839_s_at	202988_s_at	202409_at
201842_s_at	***203290_at***	***202411_at***
202018_s_at	***203329_at***	***202859_x_at***
202178_at	***203476_at***	202890_at
202411_at	***203535_at***	***202917_s_at***
202859_x_at	***203695_s_at***	*202976_s_at*
202917_s_at	***203757_s_at***	*202988_s_at*

CCG 1961 Probe Sets (167)		
HC	**COPA**	**ROSE**
117_at	***1554008_at***	***1554008_at***
1554008_at	***1555167_s_at***	***1555167_s_at***
1554140_at	1555216_a_at	*1555578_at*
1554655_a_at	*1555578_at*	***1555579_s_at***
1555167_s_at	***1555579_s_at***	*1559394_a_at*
1555579_s_at	***1557534_at***	***1559477_s_at***
1557534_at	***1559280_a_at***	*1560109_s_at*
1559280_a_at	*1559394_a_at*	***1560225_at***
1559477_s_at	***1559477_s_at***	1560483_at
1559696_at	*1560109_s_at*	1560581_at
1559910_at	***1560225_at***	*1565558_at*
1560225_at	***1562903_at***	*200800_s_at*
1562903_at	*1565558_at*	*201579_at*
1567912_s_at	*200800_s_at*	***201842_s_at***
201131_s_at	*201579_at*	*202178_at*
201215_at	***201842_s_at***	*202289_s_at*
201243_s_at	*202178_at*	*202581_at*
201842_s_at	*202289_s_at*	*202890_at*
201843_s_at	202478_at	*203038_at*
202007_at	*202581_at*	***203290_at***
202609_at	*202890_at*	203373_at
203131_at	*203038_at*	203434_s_at
203216_s_at	***203290_at***	*203476_at*

203131_at	203865_s_at	***203290_at***
203153_at	203910_at	***203329_at***
203290_at	203921_at	**203335_at**
203329_at	***203948_s_at***	**203394_s_at**
203335_at	***203949_at***	***203476_at***
203394_s_at	***204066_s_at***	***203535_at***
203476_at	***204069_at***	***203695_s_at***
203535_at	***204114_at***	**203726_s_at**
203695_s_at	***204150_at***	***203757_s_at***
203726_s_at	***204304_s_at***	*203865_s_at*
203757_s_at	204439_at	*203910_at*
203948_s_at	204456_s_at	*203921_at*
203949_at	***204895_x_at***	***203948_s_at***
203973_s_at	***204913_s_at***	***203949_at***
204014_at	***204914_s_at***	**204014_at**
204015_s_at	***204915_s_at***	***204066_s_at***
204066_s_at	***205239_at***	***204069_at***
204069_at	***205253_at***	***204114_at***
204114_at	205347_s_at	***204150_at***
204134_at	205413_at	***204304_s_at***
204150_at	***205489_at***	*204439_at*
204273_at	***205656_at***	**204614_at**
204304_s_at	***205844_at***	***204895_x_at***
204326_x_at	***205899_at***	***204913_s_at***
204351_at	205914_s_at	***204914_s_at***
204363_at	205980_s_at	***204915_s_at***
204469_at	206028_s_at	204999_s_at
204482_at	206040_s_at	205237_at
204614_at	206067_s_at	***205239_at***
204684_at	***206070_s_at***	***205253_at***
204745_x_at	206150_at	205286_at
204895_x_at	206181_at	*205347_s_at*
204913_s_at	***206258_at***	**205402_x_at**
204914_s_at	206298_at	*205413_at*
204915_s_at	***206413_s_at***	**205445_at**
204971_at	***206478_at***	205488_at
205239_at	206637_at	***205489_at***
205253_at	***207110_at***	**205493_s_at**
205402_x_at	207173_x_at	***205656_at***
205405_at	207261_at	***205844_at***

203290_at	*203476_at*	*203695_s_at*
203304_at	*203695_s_at*	*203835_at*
203632_s_at	*203835_at*	*203865_s_at*
204014_at	*203865_s_at*	***204014_at***
204015_s_at	***204014_at***	**204015_s_at**
204066_s_at	***204069_at***	***204069_at***
204069_at	*204114_at*	*204114_at*
204337_at	204304_s_at	*204439_at*
204895_x_at	204416_x_at	***204895_x_at***
205253_at	*204439_at*	204913_s_at
205382_s_at	***204895_x_at***	*204914_s_at*
205413_at	*204914_s_at*	*204915_s_at*
205493_s_at	*204915_s_at*	*204944_at*
205573_s_at	*204944_at*	*205109_s_at*
205627_at	*205109_s_at*	***205253_at***
205857_at	***205253_at***	***205413_at***
205899_at	***205382_s_at***	*205489_at*
205942_s_at	***205413_at***	*205544_s_at*
205951_at	205477_s_at	205592_at
205980_s_at	*205489_at*	***205857_at***
205987_at	*205544_s_at*	*205870_at*
206070_s_at	***205627_at***	***205899_at***
206084_at	***205857_at***	*205936_s_at*
206135_at	*205870_at*	*205946_at*
206204_at	***205899_at***	*206111_at*
206207_at	*205936_s_at*	*206181_at*
206298_at	*205946_at*	***206207_at***
206371_at	*206111_at*	*206413_s_at*
206432_at	***206135_at***	***206756_at***
206741_at	*206181_at*	*208285_at*
206756_at	***206207_at***	***209291_at***
206785_s_at	***206371_at***	*209392_at*
206851_at	*206413_s_at*	*209570_s_at*
207638_at	206710_s_at	*209602_s_at*
207768_at	***206756_at***	*209822_s_at*
207802_at	206881_s_at	***209905_at***
208029_s_at	*208285_at*	*210016_at*
208090_s_at	208470_s_at	**210665_at**
208148_at	***209291_at***	***210683_at***
208605_s_at	*209392_at*	*211306_s_at*

205445_at	207453_s_at	***205899_at***
205489_at	207696_at	**205950_s_at**
205493_s_at	***208303_s_at***	*206028_s_at*
205513_at	***208567_s_at***	*206067_s_at*
205557_at	209087_x_at	***206070_s_at***
205592_at	209101_at	*206181_at*
205593_s_at	***209291_at***	***206258_at***
205614_x_at	209604_s_at	*206298_at*
205656_at	209728_at	**206310_at**
205844_at	209897_s_at	***206413_s_at***
205857_at	***209905_at***	***206478_at***
205858_at	209959_at	**206633_at**
205863_at	***210016_at***	206756_at
205899_at	***210664_s_at***	**206836_at**
205950_s_at	***210665_at***	*207173_x_at*
206070_s_at	211340_s_at	**207651_at**
206172_at	***211657_at***	**207978_s_at**
206207_at	211735_x_at	***208303_s_at***
206258_at	***212062_at***	208553_at
206310_at	212077_at	***208567_s_at***
206413_s_at	***212094_at***	**208937_s_at**
206461_x_at	***212148_at***	*209101_at*
206478_at	***212151_at***	***209291_at***
206633_at	212158_at	**209301_at**
206634_at	212592_at	*209604_s_at*
206749_at	***213005_s_at***	209875_s_at
206836_at	***213150_at***	209892_at
206932_at	213273_at	*209897_s_at*
207110_at	***213317_at***	***209905_at***
207651_at	***213371_at***	***210016_at***
207978_s_at	***213479_at***	210150_s_at
208148_at	213714_at	**210640_s_at**
208173_at	213737_x_at	***210664_s_at***
208303_s_at	***213844_at***	***210665_at***
208567_s_at	214043_at	210869_s_at
208581_x_at	214453_s_at	*211340_s_at*
208937_s_at	214497_s_at	211341_at
209289_at	***214651_s_at***	**211506_s_at**
209290_s_at	215028_at	**211560_s_at**
209291_at	***215177_s_at***	**211597_s_at**

209289_at	*209570_s_at*	211382_s_at
209291_at	*209602_s_at*	*211560_s_at*
209436_at	*209822_s_at*	**211743_s_at**
209687_at	***209905_at***	***212148_at***
209774_x_at	*210016_at*	*212151_at*
209905_at	210432_s_at	*212592_at*
210095_s_at	***210683_at***	**212942_s_at**
210135_s_at	*211306_s_at*	*213005_s_at*
210402_at	***211518_s_at***	213050_at
210546_x_at	*211560_s_at*	***213317_at***
210664_s_at	212094_at	***213371_at***
210665_at	***212148_at***	*213423_x_at*
210683_at	*212151_at*	213906_at
211276_at	*212592_at*	*214020_x_at*
211518_s_at	*213005_s_at*	*214446_at*
211674_x_at	***213150_at***	***214651_s_at***
211719_x_at	***213317_at***	*214978_s_at*
211743_s_at	***213371_at***	*215177_s_at*
212148_at	*213423_x_at*	***216623_x_at***
212554_at	***213558_at***	***217109_at***
212942_s_at	213566_at	***217110_s_at***
213032_at	*214020_x_at*	*217963_s_at*
213150_at	214043_at	*218922_s_at*
213317_at	*214446_at*	*219355_at*
213371_at	***214651_s_at***	*219463_at*
213380_x_at	*214978_s_at*	*219489_s_at*
213418_at	*215177_s_at*	*219840_s_at*
213436_at	***215305_at***	*219855_at*
213479_at	***216623_x_at***	*220276_at*
213558_at	***217109_at***	*220377_at*
213791_at	***217110_s_at***	220922_s_at
213993_at	*217963_s_at*	*222162_s_at*
213994_s_at	*218922_s_at*	***222288_at***
214433_s_at	***219225_at***	***222450_at***
214651_s_at	*219355_at*	223075_s_at
214769_at	*219463_at*	223754_at
214774_x_at	*219489_s_at*	***223786_at***
215108_x_at	*219840_s_at*	***224022_x_at***
215121_x_at	*219855_at*	*224762_at*
215305_at	*220276_at*	225369_at

209301_at	215426_at	***211657_at***
209369_at	215666_at	***212062_at***
209757_s_at	216834_at	*212077_at*
209905_at	217083_at	***212094_at***
210016_at	***217109_at***	***212148_at***
210254_at	217963_s_at	***212151_at***
210640_s_at	218086_at	*212158_at*
210664_s_at	218468_s_at	212192_at
210665_at	218469_at	*212592_at*
210746_s_at	218625_at	***213005_s_at***
211338_at	218804_at	***213150_at***
211456_x_at	218847_at	**213258_at**
211506_s_at	219463_at	***213317_at***
211560_s_at	219489_s_at	213362_at
211597_s_at	219837_s_at	***213371_at***
211634_x_at	220059_at	***213479_at***
211639_x_at	220075_s_at	*213714_at*
211655_at	220377_at	213802_at
211657_at	***220416_at***	**213808_at**
211820_x_at	220638_s_at	***213844_at***
212062_at	220759_at	213880_at
212094_at	221066_at	214146_s_at
212104_s_at	221254_s_at	**214349_at**
212148_at	***221933_at***	214534_at
212151_at	222934_s_at	214537_at
212185_x_at	223121_s_at	***214651_s_at***
212501_at	***223278_at***	**214774_x_at**
212859_x_at	223449_at	***215177_s_at***
213005_s_at	223502_s_at	215182_x_at
213150_at	223720_at	**215379_x_at**
213194_at	223885_at	**215692_s_at**
213258_at	***224215_s_at***	**216623_x_at**
213317_at	225369_at	*217083_at*
213371_at	225436_at	***217109_at***
213418_at	225483_at	**217110_s_at**
213479_at	***225496_s_at***	217276_x_at
213488_at	225660_at	217281_x_at
213791_at	***225681_at***	217284_x_at
213808_at	226282_at	*217963_s_at*

215733_x_at	*220377_at*	225782_at
216320_x_at	220528_at	*225977_at*
216623_x_at	*222162_s_at*	***226034_at***
217109_at	222258_s_at	*226096_at*
217110_s_at	***222288_at***	*226282_at*
217138_x_at	222347_at	*226636_at*
218507_at	***222450_at***	*226913_s_at*
219093_at	223319_at	227006_at
219225_at	223422_s_at	***227289_at***
219525_at	***223786_at***	***227372_s_at***
220225_at	***224022_x_at***	*227377_at*
221731_x_at	*224762_at*	*227441_s_at*
221870_at	*225977_at*	*227949_at*
221901_at	***226034_at***	*228018_at*
222288_at	*226096_at*	*228057_at*
222315_at	*226282_at*	*228116_at*
222450_at	*226636_at*	*228262_at*
222885_at	*226913_s_at*	***228462_at***
223235_s_at	***227289_at***	***228863_at***
223611_s_at	***227372_s_at***	*228994_at*
223612_s_at	*227377_at*	229108_at
223786_at	*227441_s_at*	229247_at
224022_x_at	*227949_at*	***229638_at***
225575_at	*228018_at*	*229975_at*
225842_at	*228057_at*	230030_at
226034_at	*228116_at*	230668_at
226676_at	*228262_at*	*230680_at*
226677_at	***228462_at***	***231040_at***
227174_at	***228863_at***	***231223_at***
227289_at	*228994_at*	*231257_at*
227372_s_at	***229638_at***	231316_at
227481_at	229661_at	231455_at
227758_at	***229963_at***	*231600_at*
228462_at	*229975_at*	231859_at
228766_at	230472_at	***231899_at***
228780_at	*230680_at*	*232010_at*
228863_at	***231040_at***	*232231_at*
229147_at	***231223_at***	*232636_at*
229638_at	*231257_at*	232903_at

213844_at	***226415_at***	*218086_at*
213993_at	***226913_s_at***	218330_s_at
214349_at	***227099_s_at***	*218468_s_at*
214651_s_at	***227289_at***	*218469_at*
214774_x_at	***227439_at***	*218847_at*
215108_x_at	227440_at	*219463_at*
215177_s_at	227441_s_at	219470_x_at
215214_at	227711_at	*219489_s_at*
215379_x_at	***227949_at***	*219837_s_at*
215692_s_at	228017_s_at	**220010_at**
215784_at	***228057_at***	*220059_at*
216320_x_at	***228434_at***	*220377_at*
216336_x_at	***228462_at***	***220416_at***
216401_x_at	228599_at	*221254_s_at*
216491_x_at	***228854_at***	***221933_at***
216560_x_at	***228863_at***	222921_s_at
216623_x_at	228918_at	*222934_s_at*
216853_x_at	229029_at	*223121_s_at*
216874_at	229149_at	**223786_at**
216984_x_at	229233_at	***224215_s_at***
217109_at	229461_x_at	224520_s_at
217110_s_at	***229638_at***	*225436_at*
217143_s_at	***229661_at***	*225483_at*
217148_x_at	229967_at	***225496_s_at***
217165_x_at	229975_at	225597_at
217179_x_at	***229985_at***	***225681_at***
217235_x_at	230030_at	**226084_at**
217258_x_at	230110_at	*226282_at*
217388_s_at	230306_at	***226415_at***
217623_at	230468_s_at	**226676_at**
218145_at	230472_at	226733_at
219093_at	***230537_at***	***226913_s_at***
219360_s_at	230668_at	227006_at
219666_at	230698_at	***227099_s_at***
219714_s_at	230803_s_at	***227289_at***
220010_at	230817_at	***227439_at***
220416_at	231040_at	*227440_at*
221215_s_at	***231223_at***	*227441_s_at*

229934_at	231503_at	234985_at
229963_at	*231600_at*	235343_at
230110_at	***231899_at***	***235557_at***
230372_at	*232010_at*	*235988_at*
230495_at	*232231_at*	236430_at
231040_at	*232636_at*	*236489_at*
231223_at	***235557_at***	237207_at
231899_at	235911_at	*237421_at*
232523_at	*235988_at*	*237466_s_at*
233038_at	*236489_at*	*238617_at*
233463_at	*237421_at*	238778_at
233969_at	*237466_s_at*	*239657_x_at*
235004_at	237974_at	*239964_at*
235557_at	*238617_at*	*240032_at*
235700_at	239610_at	240179_at
235771_at	*239657_x_at*	*240245_at*
236301_at	*239964_at*	240336_at
237802_at	*240032_at*	*240347_at*
238091_at	*240245_at*	*240466_at*
238175_at	*240347_at*	*240496_at*
240758_at	*240466_at*	241506_at
242172_at	*240496_at*	241960_at
243533_x_at	***242172_at***	***242172_at***
243917_at	242747_at	242468_at
243932_at	***243917_at***	***243917_at***

221766_s_at	***231257_at***	***227949_at***
221933_at	231455_at	*228017_s_at*
222288_at	231706_s_at	***228057_at***
223278_at	***231771_at***	**228262_at**
223678_s_at	231899_at	**228297_at**
223786_at	***232231_at***	***228434_at***
223939_at	232530_at	***228462_at***
224215_s_at	***233225_at***	***228854_at***
225496_s_at	233847_x_at	***228863_at***
225681_at	234261_at	*229233_at*
226034_at	234803_at	*229461_x_at*
226084_at	234849_at	***229638_at***
226189_at	234985_at	***229661_at***
226325_at	235284_s_at	*229975_at*
226415_at	235666_at	***229985_at***
226492_at	235721_at	*230110_at*
226621_at	235911_at	**230128_at**
226676_at	***235988_at***	230130_at
226677_at	236430_at	*230472_at*
226757_at	***236489_at***	***230537_at***
226818_at	236633_at	*230698_at*
226913_s_at	236773_at	*230803_s_at*
227099_s_at	236967_at	*230817_at*
227195_at	237069_s_at	*231040_at*
227289_at	237238_at	231166_at
227439_at	237717_x_at	***231223_at***
227697_at	237828_at	***231257_at***
227949_at	237978_at	*231455_at*
228057_at	***238018_at***	231513_at
228262_at	238689_at	***231771_at***
228297_at	238900_at	*231899_at*
228434_at	239361_at	***232231_at***
228462_at	***240179_at***	**232523_at**
228854_at	***240336_at***	**232636_at**
228863_at	***240758_at***	232914_s_at
229638_at	240794_at	***233225_at***
229661_at	241527_at	*234261_at*
229985_at	241535_at	235521_at
230128_at	242172_at	*235666_at*
230255_at	242385_at	*235911_at*
230291_s_at	***242457_at***	***235988_at***

230537_at	***242468_at***	*236430_at*
230788_at	242747_at	***236489_at***
230791_at	***243533_x_at***	*236773_at*
231202_at	244002_at	***238018_at***
231223_at	244155_x_at	*238689_at*
231257_at	***244665_at***	239657_x_at
231771_at	244750_at	***240179_at***
232231_at	244782_at	***240336_at***
232523_at	***1552398_a_at***	***240758_at***
232629_at	1552767_a_at	*241535_at*
232636_at	1553629_a_at	**241960_at**
233225_at	1553963_at	*242172_at*
234830_at	1554343_a_at	*242385_at*
235249_at	1554912_at	***242457_at***
235371_at	1555220_a_at	***242468_at***
235988_at	***1555579_s_at***	***243533_x_at***
236489_at	1555745_a_at	***244665_at***
237471_at	***1557534_at***	*244750_at*
237613_at	1557876_at	***1552398_a_at***
237625_s_at	1559394_a_at	1552511_a_at
238018_at	1559459_at	*1552767_a_at*
238423_at	***1559477_s_at***	*1553629_a_at*
240104_at	1559842_at	*1554343_a_at*
240179_at	1559865_at	**1554633_a_at**
240336_at	1560315_at	***1555579_s_at***
240758_at	1560642_at	*1555745_a_at*
241960_at	1561025_at	1555756_a_at
242457_at	1563868_a_at	***1557534_at***
242468_at	1566825_at	*1559394_a_at*
242541_at	1568603_at	*1559459_at*
243533_x_at	1569591_at	***1559477_s_at***
244463_at	1569663_at	*1561025_at*
244665_at	1570058_at	*1566825_at*

Table S7B. Overlap of Probe Sets Used in Either P9906 or CCG1961

P9906 (254 total)		
	COPA	ROSE
HC	96 (37.8%)	135 (53.1%)
COPA	--	169 (66.5%)
HC & COPA	--	94 (37.0%)

CCG1961 (167 total)		
	COPA	ROSE
HC	55 (32.9%)	46 (27.5%)
COPA	--	130 (77.8%)
HC & COPA	--	42 (25.1%)

Table S7C: Common P9906 and CCG1961 Probe Sets by Method

	HC (1961)	COPA (1961)	ROSE (1961)
HC (9906)	55 (32.9%)	56 (33.5%)	59 (35.3%)
COPA (9906)	36 (21.6%)	66 (39.5%)	68 (40.7%)
ROSE (9906)	45 (26.9%)	75 (44.9%)	77 (46.1%)

OVERLAP OF P9906 CLUSTERS DEFINED BY EACH METHOD

Each of the three clustering methods in P9906 identified predominantly the same samples even though they shared only 37% of the probe sets (Table S7B). As in shown in Table S8, the overall identity of samples across all three methods is 86.5%. The primary factor responsible for this being lower than ~90% is that HC and ROSE identified a cluster 4, while COPA did not. All 23 of the patients with *TCF3-PBX1* translocations were grouped into cluster 1 by all three methods, as were 19 of the 21 patients with *MLL* translocations. Even though the remaining clusters lacked known underlying translocations they were also very highly conserved.

Table S8. Identity of Membership in P9906 Clusters

	Cluster								
	1	2	3	4	5	6	7	8	Overall
HC v COPA	19	23	8	0	9	19	88	19	89.4%
HC v ROSE	20	23	8	10	9	19	82	22	93.2%
COPA v ROSE	20	23	10	0	10	21	82	20	89.9%
HC v COPA v ROSE	19	23	8	0	9	19	82	19	86.5%

PROBESETS ASSOCIATED WITH ROSE CLUSTERS (BY MEDIAN RANK ORDER)

The top 100 median rank order probe sets for each ROSE cluster are given. Percentile denotes the ranking of the median cluster rank order relative to the maximum possible. Bold font indicates that these probe sets were also among the 254 outliers selected for clustering. Probe sets marked with an asterisk (including several PCDH17, GAB1, GPR110, AGAP1 and CD99) indicate those for which Affymetrix does not specify a gene, however the probe sets were mapped using the UCSC Genome Browser (http://genome.ucsc.edu/) between exons of the indicated genes. Those with a question mark were also lacking Affymetrix gene data, but were mapped within 10 kb of the indicated gene using the UCSC Genome Browser.

Table S9. Top 100 Rank Order Genes Defining ROSE Cluster 1 (R1)

Probeset	Percentile	Symbol	EntrezID	Cytoband
219463_at	**100**	**C20orf103**	**24141**	**20p12**
205899_at	**100**	**CCNA1**	**8900**	**13q12.3-q13**
235479_at	100	CPEB2	132864	4p15.33
226939_at	100	CPEB2	132864	4p15.33
241706_at	100	CPNE8	144402	12q12
236921_at	100	EMB*	---	5q11.1
222603_at	100	ERMP1	79956	9p24
213147_at	100	HOXA10	3206	7p15-p14
213150_at	**100**	**HOXA10**	**3206**	**7p15-p14**
235521_at	**100**	**HOXA3**	**3200**	**7p15-p14**
214651_s_at	**100**	**HOXA9**	**3205**	**7p15-p14**
209905_at	**100**	**HOXA9**	**3205**	**7p15-p14**
215163_at	100	IGF2BP2*	---	3q27.2
226789_at	100	LOC647121	647121	1p11.2
202890_at	**100**	**MAP7**	**9053**	**6q23.3**
238498_at	100	MAP7?	---	6q23.3
204069_at	**100**	**MEIS1**	**4211**	**2p14-p13**
242172_at	**100**	**MEIS1**	**4211**	**2p14-p13**
1559477_s_at	**100**	**MEIS1**	**4211**	**2p14-p13**
219033_at	100	PARP8	79668	5q11.1
204304_s_at	**100**	**PROM1**	**8842**	**4p15.32**
242414_at	100	QPRT	23475	16p11.2
204044_at	100	QPRT	23475	16p11.2
1568589_at	100	REEP3*	---	10q21.3
231899_at	**100**	**ZC3H12C**	**85463**	**11q22.3**
220416_at	**99.5**	**ATP8B4**	**79895**	**15q21.2**
225841_at	99.5	C1orf59	113802	1p13.3
227877_at	99.5	C5orf39	389289	5p12
212063_at	99.5	CD44	960	11p13
213844_at	**99.5**	**HOXA5**	**3202**	**7p15-p14**
218847_at	**99.5**	**IGF2BP2**	**10644**	**3q27.2**
201163_s_at	99.5	IGFBP7	3490	4q12
201105_at	**99.5**	**LGALS1**	**3956**	**22q13.1**
228412_at	99.5	LOC643072	643072	2q24.2
240180_at	99.5	MAP7?	---	6q23.3
201153_s_at	99.5	MBNL1	4154	3q25
1558111_at	99.5	MBNL1	4154	3q25
1556658_a_at	99.5	MBNL1*	---	3q25.2

238558_at	99.5	MBNL1*	---	3q25.2
244008_at	99.5	PARP8?	---	5q11.1
204082_at	99.5	PBX3	5090	9q33-q34
230480_at	99.5	PIWIL4	143689	11q21
232231_at	**99.5**	**RUNX2**	**860**	**6p21**
211769_x_at	99.5	SERINC3	10955	20q13.1-q13.3
226415_at	**99.5**	**VAT1L**	**57687**	**16q23.1**
203827_at	99.5	WIPI1	55062	17q24.2
242023_at	99	ABHD4	63874	14q11.2
202603_at	99	ADAM10*	---	15q22.1
215925_s_at	99	CD72	971	9p13.3
228365_at	99	CPNE8	144402	12q12
214297_at	99	CSPG4	1464	15q24.2
200046_at	99	DAD1	1603	14q11-q12
227002_at	99	FAM78A	286336	9q34
235291_s_at	99	FLJ32255	643977	5p12
238712_at	99	FOXP1*	---	3p14.1
204417_at	99	GALC	2581	14q31
235173_at	99	hCG_1806964	401093	3q25.1
201162_at	99	IGFBP7	3490	4q12
232544_at	99	IGFBP7*	---	4q12
241391_at	99	JMJD1C*	---	10q21.2
1557534_at	**99**	**LOC339862**	**339862**	**3p24.3**
1556657_at	99	MBNL1*	---	3q25.2
219988_s_at	99	RNF220	55182	1p34.1
221473_x_at	99	SERINC3	10955	20q13.1-q13.3
206506_s_at	99	SUPT3H	8464	6p21.1-p21.3
213836_s_at	99	WIPI1	55062	17q24.2
218581_at	98.5	ABHD4	63874	14q11.2
214895_s_at	98.5	ADAM10	102	15q2|15q22
212174_at	98.5	AK2	204	1p34
203562_at	98.5	FEZ1	9638	11q24.2
235753_at	98.5	HOXA7	3204	7p15-p14
213910_at	98.5	IGFBP7	3490	4q12
1569041_at	98.5	JMJD1C*	---	10q21.2
203836_s_at	98.5	MAP3K5	4217	6q22.33
203837_at	98.5	MAP3K5	4217	6q22.33
201152_s_at	98.5	MBNL1	4154	3q25
235879_at	98.5	MBNL1	4154	3q25
225202_at	98.5	RHOBTB3	22836	5q15
227719_at	98.5	SMAD9	4093	13q12-q14

225959_s_at	98.5	ZNRF1	84937	16q23.1
223382_s_at	98.5	ZNRF1	84937	16q23.1
210783_x_at	98	CLEC11A	6320	19q13.3
232645_at	98	LOC153684	153684	5p12
241681_at	98	MBNL1*	---	3q25.2
202976_s_at	**98**	**RHOBTB3**	**22836**	**5q15**
227611_at	98	TARSL2	123283	15q26.3
209825_s_at	98	UCK2	7371	1q23
223383_at	98	ZNRF1	84937	16q23.1
36553_at	97.5	ASMTL	8623	Xp22.3; Yp11.3
224848_at	97.5	CDK6	1021	7q21-q22
213379_at	97.5	COQ2	27235	4q21.23
209101_at	**97.5**	**CTGF**	**1490**	**6q23.1**
218147_s_at	97.5	GLT8D1	55830	3p21.1
218468_s_at	**97.5**	**GREM1**	**26585**	**15q13-q15**
227235_at	97.5	GUCY1A3	2982	4q31.3-q33|4q31.1-q31.2
206289_at	97.5	HOXA4	3201	7p15-p14
227384_s_at	97.5	LOC727820	727820	1q21.1
203537_at	97.5	PRPSAP2	5636	17p11.2-p12
226168_at	97.5	ZFAND2B	130617	2q35
225962_at	97.5	ZNRF1	84937	16q23.1

Table S10. Top 100 Rank Order Genes Defining ROSE Cluster 2 (R2)

Probeset	Percentile	Symbol	EntrezID	Cytoband
227440_at	**100**	**ANKS1B**	**56899**	**12q23.1**
227441_s_at	**100**	**ANKS1B**	**56899**	**12q23.1**
227439_at	**100**	**ANKS1B**	**56899**	**12q23.1**
234261_at	**100**	**ANKS1B***	**---**	**12q23.1**
243533_x_at	**100**	**ANKS1B***	**---**	**12q23.1**
202206_at	**100**	**ARL4C**	**10123**	**2q37.1**
229247_at	100	FBLN7	129804	2q13
239657_x_at	**100**	**FOXO6**	**100132074**	**1p34.1**
202106_at	100	GOLGA3	2802	12q24.33
213005_s_at	**100**	**KANK1**	**23189**	**9p24.3**
207110_at	100	KCNJ12	3768	17p11.2
232289_at	100	KCNJ12	3768	17p11.2
208567_s_at	**100**	**KCNJ12 /// LOC100131509 /// LOC100134444**	**100131509 /// 100134444 /// 3768**	**17p11.2**
213909_at	100	LRRC15	131578	3q29
206028_s_at	**100**	**MERTK**	**10461**	**2q14.1**

211913_s_at	100	MERTK	10461	2q14.1
238778_at	100	MPP7	143098	10p11.23
212789_at	100	NCAPD3	23310	11q25
212148_at	**100**	**PBX1**	**5087**	**1q23**
212151_at	**100**	**PBX1**	**5087**	**1q23**
205253_at	**100**	**PBX1**	**5087**	**1q23**
227949_at	**100**	**PHACTR3**	**116154**	**20q13.32**
231095_at	100	PITPNC1*	---	17q24.2
202178_at	**100**	**PRKCZ**	**5590**	**1p36.33-p36.2**
223693_s_at	100	RADIL	55698	7p22.1
222513_s_at	100	SORBS1	10580	10q23.3-q24.1
225235_at	100	TSPAN17	26262	5q35.3
225483_at	**100**	**VPS26B**	**112936**	**11q25**
224022_x_at	100	WNT16	51384	7q31
202207_at	**99.5**	**ARL4C**	**10123**	**2q37.1**
202208_s_at	99.5	ARL4C	10123	2q37.1
206255_at	99.5	BLK	640	8p23-p22
223786_at	**99.5**	**CHST6**	**4166**	**16q22**
205489_at	**99.5**	**CRYM**	**1428**	**16p13.11-p12.3**
205159_at	99.5	CSF2RB	1439	22q13.1
212538_at	99.5	DOCK9	23348	13q32.3
229655_at	99.5	FAM19A5	25817	22q13.32
206404_at	99.5	FGF9	2254	13q11-q12
209558_s_at	99.5	HIP1R	9026	12q24
38340_at	99.5	HIP1R	9026	12q24
235911_at	**99.5**	**K03200***	**---**	**3q29**
204114_at	**99.5**	**NID2**	**22795**	**14q21-q22**
1562235_s_at	99.5	PBX1*	---	1q23.3
229414_at	99.5	PITPNC1	26207	17q24.2
231040_at	**99.5**	**RORB?**	**---**	**9q21.13**
46665_at	**99.5**	**SEMA4C**	**54910**	**2q11.2**
206181_at	**99.5**	**SLAMF1**	**6504**	**1q22-q23**
239427_at	99.5	SLAMF1?	---	1q23.3
203940_s_at	99.5	VASH1	22846	14q24.3
230306_at	99.5	VPS26B	112936	11q25
221113_s_at	99.5	WNT16	51384	7q31
226233_at	99	B3GALNT2	148789	1q42.3
201615_x_at	99	CALD1	800	7q33
209570_s_at	99	D4S234E	27065	4p16.3
229892_at	99	EP400NL	347918	12q24.33
206070_s_at	**99**	**EPHA3**	**2042**	**3p11.2**
237094_at	99	FAM19A5	25817	22q13.32
227676_at	99	FAM3D	131177	3p14.2
201579_at	**99**	**FAT1**	**2195**	**4q35**
204225_at	99	HDAC4	9759	2q37.3
1566030_at	99	PHACTR3*	---	20q13.32
242385_at	**99**	**RORB**	**6096**	**9q22**
221669_s_at	98.5	ACAD8	27034	11q25

205083_at	98.5	AOX1	316	2q33
225313_at	98.5	C20orf177	63939	20q13.2-q13.33
201616_s_at	98.5	CALD1	800	7q33
209569_x_at	98.5	D4S234E	27065	4p16.3
212371_at	98.5	FAM152A	51029	1q44
229770_at	98.5	GLT1D1	144423	12q24.32
226949_at	98.5	GOLGA3	2802	12q24.33
204202_at	98.5	IQCE	23288	7p22.2
213358_at	98.5	KIAA0802	23255	18p11.22
210150_s_at	**98.5**	**LAMA5**	**3911**	**20q13.2-q13.3**
238451_at	98.5	MPP7	143098	10p11.23
219155_at	98.5	PITPNC1	26207	17q24.2
215807_s_at	98.5	PLXNB1	5364	3p21.31
225728_at	98.5	SORBS2	8470	4q35.1
217650_x_at	98.5	ST3GAL2	6483	16q22.1
1554340_a_at	98	C1orf187	374946	1p36.22
212077_at	**98**	**CALD1**	**800**	**7q33**
220373_at	98	DCHS2	54798	4q32.1
232204_at	98	EBF1	1879	5q34
201718_s_at	98	EPB41L2	2037	6q23
201719_s_at	98	EPB41L2	2037	6q23
231455_at	**98**	**FLJ42418**	**400941**	**2p25.2**
219271_at	98	GALNT14	79623	2p23.1
214265_at	98	ITGA8	8516	10p13
235666_at	**98**	**ITGA8?**	**---**	**10p13**
209760_at	98	KIAA0922	23240	4q31.3
226796_at	98	LOC116236	116236	17q11.2
228262_at	**98**	**MAP7D2**	**256714**	**Xp22.12**
212845_at	98	SAMD4A	23034	14q22.2
202796_at	98	SYNPO	11346	5q33.1
222752_s_at	98	TMEM206	55248	1q32.3
227733_at	98	TMEM63C	57156	14q24.3
242957_at	98	VWCE	220001	11q12.2
224516_s_at	97.4	CXXC5	51523	5q31.3
220911_s_at	97.4	KIAA1305	57523	14q12
213136_at	97.4	PTPN2	5771	18p11.3-p11.2
202478_at	97.4	TRIB2	28951	2p25.1-p24.3

Table S11. Top 100 Rank Order Genes Defining ROSE Cluster 3 (R3)

Probeset	Percentile	Symbol	EntrezID	Cytoband
244463_at	100	ADAM23	8745	2q33
240143_at	100	ADAM23*	---	2q33.3
213808_at	**100**	**ADAM23***	**---**	**2q33.3**
204129_at	100	BCL9	607	1q21
213050_at	100	COBL	23242	7p12.1

205659_at	100	HDAC9	9734	7p21.1
230968_at	100	HDAC9?	---	7p21.1
217869_at	100	HSD17B12	51144	11p11.2
1557252_at	100	HSD17B12*	---	11p11.2
216028_at	100	HSD17B12?	---	11p11.2
242616_at	100	HSD17B12?	---	11p11.2
230128_at	**100**	**IGL@**	**3535**	**22q11.1-q11.2**
204686_at	100	IRS1	3667	2q36
206765_at	100	KCNJ2	3759	17q23.1-q24.2
203726_s_at	**100**	**LAMA3**	**3909**	**18q11.2**
224823_at	100	MYLK	4638	3q21
202555_s_at	100	MYLK	4638	3q21
216012_at	100	PDE4D*	---	5q12.1
205632_s_at	100	PIP5K1B	8395	9q13
204469_at	100	PTPRZ1	5803	7q31.3
212104_s_at	100	RBM9	23543	22q13.1
213243_at	100	VPS13B	157680	8q22.2
226325_at	99.5	ADSSL1	122622	14q32.33
1552496_a_at	99.5	COBL	23242	7p12.1
219518_s_at	99.5	ELL3	80237	15q15.3
231513_at	**99.5**	**KCNJ2***	**---**	**17q24.3**
221584_s_at	99.5	KCNMA1	3778	10q22.3
213568_at	99.5	OSR2	116039	8q22.2
202780_at	99.5	OXCT1	5019	5p13.1
239832_at	99.5	PIP5K1B*	---	9q21.11
213309_at	99.5	PLCL2	23228	3p24.3
216218_s_at	99.5	PLCL2	23228	3p24.3
203020_at	99.5	RABGAP1L	9910	1q24
203097_s_at	99.5	RAPGEF2	9693	4q32.1
218137_s_at	99.5	SMAP1	60682	6q13
223246_s_at	99.5	STRBP	55342	9q33.3
225496_s_at	**99.5**	**SYTL2**	**54843**	**11q14**
1554803_s_at	99.5	TRIM72	493829	16p11.2
206046_at	99	ADAM23	8745	2q33
203865_s_at	**99**	**ADARB1**	**104**	**21q22.3**
206167_s_at	99	ARHGAP6	395	Xp22.3
219517_at	99	ELL3	80237	15q15.3
45572_s_at	99	GGA1	26088	22q13.31
204891_s_at	99	LCK	3932	1p34.3
204890_s_at	99	LCK	3932	1p34.3
222322_at	99	PDE4D*	---	5q12.1

203038_at	99	PTPRK	5796	6q22.2-q22.3
213982_s_at	99	RABGAP1L	9910	1q24
238894_at	99	RABGAP1L*	---	1q25.1
203096_s_at	99	RAPGEF2	9693	4q32.1
215992_s_at	99	RAPGEF2	9693	4q32.1
232739_at	99	SPIB	6689	19q13.3-q13.4
220613_s_at	99	SYTL2	54843	11q14
212350_at	99	TBC1D1	23216	4p14
203588_s_at	99	TFDP2	7029	3q23
219520_s_at	99	WWC3	55841	Xp22.32
227173_s_at	98.5	BACH2	60468	6q15
241871_at	98.5	CAMK4	814	5q21.3
206806_at	98.5	DGKI	9162	7q32.3-q33
205425_at	98.5	HIP1	3092	7q11.23
215946_x_at	98.5	IGLL3	91353	22q11.2|22q11.23
225963_at	98.5	KLHDC5	57542	12p11.22
234608_at	98.5	LAMA3	3909	18q11.2
217140_s_at	98.5	LOC100133724 /// VDAC1	100133724 /// 7416	5q31
213502_x_at	98.5	LOC91316	91316	22q11.23
205826_at	98.5	MYOM2	9172	8p23.3
244387_at	98.5	PDE4D*	---	5q12.1
1565762_at	98.5	RABGAP1L*	---	1q25.1
205590_at	98.5	RASGRP1	10125	15q14
232914_s_at	**98.5**	**SYTL2**	**54843**	**11q14**
244043_at	98.5	TFDP2?	---	3q23
223750_s_at	98.5	TLR10	81793	4p14
212038_s_at	98.5	VDAC1	7416	5q31
243734_x_at	98.5	VWC2?	---	7p12.2
243526_at	98.5	WDR86	349136	7q36.1
234033_at	98	---	---	4q32.1
203263_s_at	98	ARHGEF9	23229	Xq11.1
213238_at	98	ATP10D	57205	4p12
221234_s_at	98	BACH2	60468	6q15
218285_s_at	98	BDH2	56898	4q24
235952_at	98	DGKH-1*	---	13q14.11
234912_at	98	DKFZP547L112	81787	15q11.2
213186_at	98	DZIP3	9666	3q13.13
50277_at	98	GGA1	26088	22q13.31
242952_at	98	HDAC9*	---	7p21.1
214836_x_at	98	IGKC	3514	2p12
237625_s_at	98	IGKC*	---	2p12

225961_at	98	KLHDC5	57542	12p11.22
230551_at	98	KSR2	283455	12q24.22-q24.23
205386_s_at	98	MDM2	4193	12q14.3-q15
222350_at	97.5	BTBD3	22903	20p12.2
229715_at	97.5	BTBD6	90135	14q32
202946_s_at	97.5	IGKC	3514	2p12
225389_at	97.5	KCNJ11?	---	11p15.1
214669_x_at	97.5	LOC729082	729082	15q15.1
225332_at	97.5	NBPF1*	---	1q21.1
213273_at	97.5	ODZ4	26011	11q14.1
235802_at	97.5	PLD4	122618	14q32.33
218526_s_at	97.5	RANGRF	29098	17p13
230597_at	97.5	SLC7A3	84889	Xq13.1

Table S12. Top 100 Rank Order Genes Defining ROSE Cluster 4 (R4)

Probeset	Rank	Symbol	EntrezID	Cytoband
210356_x_at	100.0%	MS4A1	931	11q12
217418_x_at	100.0%	MS4A1	931	11q12
205401_at	99.5%	AGPS	8540	2q31.2
228592_at	99.5%	MS4A1	931	11q12
241774_at	99.5%	---	---	---
218941_at	99.5%	FBXW2	26190	9q34
225114_at	99.0%	AGPS	8540	2q31.2
202123_s_at	99.0%	ABL1	25	9q34.1
203476_at	99.0%	TPBG	7162	6q14-q15
214783_s_at	98.5%	ANXA11	311	10q23
202947_s_at	98.5%	GYPC	2995	2q14-q21
225833_at	98.5%	DAGLB	221955	7p22.1
225073_at	98.5%	PPHLN1	51535	12q12
212730_at	98.5%	SYNM	23336	15q26.3
227846_at	98.5%	GPR176	11245	15q14-q15.1
223991_s_at	98.5%	GALNT2 /// LOC100132910	100132910 /// 2590	18q12.2 /// 1q41-q42
208195_at	98.0%	TTN	7273	2q31
233713_at	98.0%	---	---	---
217788_s_at	98.0%	GALNT2	2590	1q41-q42
224830_at	98.0%	NUDT21	11051	16q13

226832_at	98.0%	---	---	---
202273_at	98.0%	PDGFRB	5159	5q31-q32
225376_at	98.0%	C20orf11	54994	20q13.33
225281_at	98.0%	C3orf17	25871	3q13.2
201096_s_at	98.0%	ARF4	378	3p21.2-p21.1
203948_s_at	97.5%	MPO	4353	17q23.1
1558017_s_at	97.5%	---	---	---
203949_at	97.5%	MPO	4353	17q23.1
1555392_at	97.5%	LOC100128868	100128868	7q31.2
227541_at	97.5%	WDR20	91833	14q32.31
1567458_s_at	97.5%	RAC1	5879	7p22
213920_at	97.5%	CUX2	23316	12q24.11-q24.12
224734_at	97.5%	HMGB1	3146	13q12
206673_at	97.5%	GPR176	11245	15q14-q15.1
224636_at	97.5%	ZFP91	80829	11q12
235232_at	97.5%	GMEB1	10691	1p35.3
208762_at	97.5%	SUMO1	7341	2q33
36612_at	97.0%	FAM168A	23201	11q13.4
225240_s_at	97.0%	MSI2	124540	17q22
336_at	97.0%	TBXA2R	6915	19p13.3
223101_s_at	97.0%	ARPC5L	81873	9q33.3
209049_s_at	97.0%	ZMYND8	23613	20q13.12
217940_s_at	97.0%	CARKD	55739	13q34
216508_x_at	97.0%	CTCFL /// HMGB1 /// HMGB1L1 /// HMGB1L10 /// LOC100132863	100130561 /// 100132863 /// 10357 /// 140690 /// 3146	13q12 /// 20q13.31 /// 20q13.32 /// 22q12.1 /// 9q33.2
201266_at	97.0%	TXNRD1	7296	12q23-q24.1
212286_at	97.0%	ANKRD12	23253	18p11.22
200618_at	97.0%	LASP1	3927	17q11-q21.3
227577_at	97.0%	EXOC8	149371	1q42.2
203068_at	97.0%	KLHL21	9903	1p36.31
217787_s_at	97.0%	GALNT2	2590	1q41-q42
239930_at	97.0%	GALNT2	2590	1q41-q42

227700_x_at	97.0%	ATAD3A	55210	1p36.33
225694_at	97.0%	CRKRS	51755	17q12
202514_at	97.0%	DLG1	1739	3q29
226115_at	97.0%	AHCTF1	25909	1q44
1562948_at	97.0%	---	---	---
225456_at	97.0%	MED1	5469	17q12-q21.1
208821_at	97.0%	SNRPB	6628	20p13
212204_at	97.0%	TMEM87A	25963	15q15.1
231124_x_at	97.0%	LY9	4063	1q21.3-q22
218118_s_at	97.0%	TIMM23	10431	10q11.21-q11.23
212272_at	96.5%	LPIN1	23175	2p25.1
220684_at	96.5%	TBX21	30009	17q21.32
216836_s_at	96.5%	ERBB2	2064	17q11.2-q12|17q21.1
232521_at	96.5%	PCSK7	9159	11q23-q24
205839_s_at	96.5%	BZRAP1	9256	17q22-q23
218031_s_at	96.5%	FOXN3	1112	14q31.3
226640_at	96.5%	DAGLB	221955	7p22.1
213514_s_at	96.5%	DIAPH1	1729	5q31
225494_at	96.5%	DYNLL2	140735	17q22
213222_at	96.5%	PLCB1	23236	20p12
212594_at	96.5%	PDCD4	27250	10q24
201133_s_at	96.5%	PJA2	9867	5q21.3
235463_s_at	96.5%	LASS6	253782	2q24.3
200047_s_at	96.5%	YY1	7528	14q
201407_s_at	96.5%	PPP1CB	5500	2p23
1552931_a_at	96.5%	PDE8A	5151	15q25.3
242467_at	96.5%	---	---	---
213860_x_at	96.5%	CSNK1A1	1452	5q32
212927_at	96.5%	SMC5	23137	9q21.11
227237_x_at	96.5%	ATAD3B /// LOC732419	732419 /// 83858	1p36.33

200775_s_at	96.5%	HNRNPK	3190	9q21.32-q21.33
210203_at	96.5%	CNOT4	4850	7q22-qter
214352_s_at	96.5%	KRAS	3845	12p12.1
1555772_a_at	96.5%	CDC25A	993	3p21
212696_s_at	96.5%	RNF4	6047	4p16.3
235233_s_at	96.5%	GMEB1	10691	1p35.3
225535_s_at	96.5%	TIMM23	10431	10q11.21-q11.23
1555762_s_at	96.5%	RBM15	64783	1p13
204735_at	96.5%	PDE4A	5141	19p13.2
228599_at	96.0%	MS4A1	931	11q12
212511_at	96.0%	PICALM	8301	11q14
207681_at	96.0%	CXCR3	2833	Xq13
224912_at	96.0%	TTC7A	57217	2p21
218447_at	96.0%	C16orf61	56942	16q23.2
204206_at	96.0%	MNT	4335	17p13.3
227433_at	96.0%	KIAA2018	205717	3q13.2
224617_at	96.0%	ROD1	9991	9q32
1560339_s_at	96.0%	NAP1L4	4676	11p15.5
201015_s_at	96.0%	JUP	3728	17q21

Table S13. Top 100 Rank Order Genes Defining ROSE Cluster 5 (R5)

Probeset	Percentile	Symbol	EntrezID	Cytoband
202804_at	100	ABCC1	4363	16p13.1
204638_at	100	ACP5	54	19p13.3-p13.2
205423_at	100	AP1B1	162	22q12|22q12.2
212062_at	**100**	**ATP9A**	**10079**	**20q13.2**
216129_at	100	ATP9A	10079	20q13.2
236226_at	100	BTLA	151888	3q13.2
209498_at	100	CEACAM1	634	19q13.2
222786_at	100	CHST12	55501	7p22
218927_s_at	100	CHST12	55501	7p22
219500_at	100	CLCF1	23529	11q13.3
1556385_at	100	CLCF1*	---	11q13.1
201445_at	100	CNN3	1266	1p22-p21
228297_at	**100**	**CNN3***	**---**	**1p21.3**
228585_at	100	ENTPD1	953	10q24
1554903_at	100	FRMD8	83786	11q13

1554905_x_at	100	FRMD8	83786	11q13
227964_at	100	FRMD8	83786	11q13
230788_at	100	GCNT2	2651	6p24.2
202032_s_at	100	MAN2A2	4122	15q26.1
209703_x_at	100	METTL7A	25840	12q13.13
226531_at	100	ORAI1	84876	12q24.31
60471_at	100	RIN3	79890	14q32.12
207735_at	100	RNF125	54941	18q12.1
229661_at	**100**	**SALL4**	**57167**	**20q13.13-q13.2**
222088_s_at	100	SLC2A14 /// SLC2A3	144195 /// 6515	12p13.3 /// 12p13.31
202498_s_at	100	SLC2A3	6515	12p13.3
202499_s_at	100	SLC2A3	6515	12p13.3
213083_at	100	SLC35D2	11046	9q22.32
215447_at	100	TFPI	7035	2q32
231775_at	100	TNFRSF10A	8797	8p21
227595_at	100	ZMYM6	9204	1p34.2
243121_x_at	99.5	---	---	19q13.41
223646_s_at	99.5	CYorf15B	84663	Yq11.222
203139_at	99.5	DAPK1	1612	9q34.1
211214_s_at	99.5	DAPK1	1612	9q34.1
223306_at	99.5	EBPL	84650	13q12-q13
209474_s_at	99.5	ENTPD1	953	10q24
209473_at	99.5	ENTPD1	953	10q24
229280_s_at	99.5	FLJ22536	401237	6p22.3
228188_at	99.5	FOSL2	2355	2p23.3
AFFX-HUMGAPDH/M33197_5_at	99.5	GAPDH	2597	12p13
204689_at	99.5	HHEX	3087	10q23.33
1552623_at	99.5	HSH2D	84941	19p13.11
207761_s_at	99.5	METTL7A	25840	12q13.13
207132_x_at	99.5	PFDN5	5204	12q12
1557948_at	99.5	PHLDB3	653583	19q13.31
213362_at	**99.5**	**PTPRD**	**5789**	**9p23-p24.3**
227983_at	99.5	RILPL2	196383	12q24.31
219457_s_at	99.5	RIN3	79890	14q32.12
211474_s_at	99.5	SERPINB6	5269	6p25
223196_s_at	99.5	SESN2	83667	1p35.3
216236_s_at	99.5	SLC2A14 /// SLC2A3	144195 /// 6515	12p13.3 /// 12p13.31
202497_x_at	99.5	SLC2A3	6515	12p13.3

227594_at	99.5	ZMYM6	9204	1p34.2
202805_s_at	99	ABCC1	4363	16p13.1
213346_at	99	C13orf27	93081	13q33.1
223527_s_at	99	CDADC1	81602	13q14.2
213060_s_at	99	CHI3L2	1117	1p13.3
203277_at	99	DFFA	1676	1p36.3-p36.2
208887_at	99	EIF3G	8666	19p13.2
219016_at	99	FASTKD5	60493	20p13
218034_at	99	FIS1	51024	7q22.1
225163_at	99	FRMD4A	55691	10p13
239606_at	99	GCNT2A*	---	6p24.2
230348_at	99	LATS2	26524	13q11-q12
209332_s_at	99	MAX	4149	14q23
227379_at	99	MBOAT1	154141	6p22.3
217980_s_at	99	MRPL16	54948	11q12-q13.1
238082_at	99	PLEKHA2*	---	8p11.23
232473_at	99	PRPF18	8559	10p13
220330_s_at	99	SAMSN1	64092	21q11
223917_s_at	99	SLC39A3	29985	19p13.3
219257_s_at	99	SPHK1	8877	17q25.2
203544_s_at	99	STAM	8027	10p14-p13
213258_at	**99**	**TFPI**	**7035**	**2q32**
210664_s_at	**99**	**TFPI**	**7035**	**2q32**
210665_at	**99**	**TFPI**	**7035**	**2q32**
201379_s_at	99	TPD52L2	7165	20q13.2-q13.3
212481_s_at	99	TPM4	7171	19p13.1
235094_at	99	TPM4*	---	19p13.2
212923_s_at	98.5	C6orf145	221749	6p25.2
206120_at	98.5	CD33	945	19q13.3
1559916_a_at	98.5	CHST12*	---	7p22.2
1554464_a_at	98.5	CRTAP	10491	3p22.3
209774_x_at	98.5	CXCL2	2920	4q21
225168_at	98.5	FRMD4A	55691	10p13
213453_x_at	98.5	GAPDH	2597	12p13
209604_s_at	**98.5**	**GATA3**	**2625**	**10p15**
209602_s_at	98.5	GATA3	2625	10p15
204000_at	98.5	GNB5	10681	15q21.2
233877_at	98.5	GOLIM4*	---	3q26.2
203395_s_at	98.5	HES1	3280	3q28-q29
214950_at	98.5	IL9R /// LOC729486	3581 /// 729486	16p13.3 /// Xq28 and Yq12
213923_at	98.5	RAP2B	5912	3q25.2

238091_at	98.5	RPH3AL*	---	17p13.3
236501_at	98.5	SALL4	57167	20q13.13-q13.2
223195_s_at	98.5	SESN2	83667	1p35.3
227518_at	98.5	SLC35E1	79939	19p13.11
243981_at	98.5	STK4	6789	20q11.2-q13.2
212369_at	98.5	ZNF384	171017	12p12

Table S14. Top 100 Rank Order Genes Defining ROSE Cluster 6 (R6)

Probeset	Percentile	Symbol	EntrezID	Cytoband
242457_at	**100**	**---**	**---**	**5q21.1**
204066_s_at	**100**	**AGAP1**	**116987**	**2q37**
233038_at	100	AGAP1*	---	2q37.2
233225_at	**100**	**AGAP1***	**---**	**2q37.2**
235968_at	100	AGAP1*	---	2q37.2
240758_at	**100**	**AGAP1***	**---**	**2q37.2**
228240_at	100	AGAP1?	---	2q37.2
206756_at	**100**	**CHST7**	**56548**	**Xp11.23**
200614_at	100	CLTC	1213	17q11-qter
231166_at	**100**	**GPR155**	**151556**	**2q31.1**
228863_at	**100**	**PCDH17**	**27253**	**13q21.1**
227289_at	**100**	**PCDH17**	**27253**	**13q21.1**
205656_at	**100**	**PCDH17**	**27253**	**13q21.1**
230537_at	**100**	**PCDH17?**	**---**	**13q21.1**
203335_at	**100**	**PHYH**	**5264**	**10p13**
1555579_s_at	**100**	**PTPRM**	**5797**	**18p11.2**
203329_at	**100**	**PTPRM**	**5797**	**18p11.2**
1554343_a_at	**100**	**STAP1**	**26228**	**4q13.2**
220059_at	**100**	**STAP1**	**26228**	**4q13.2**
211890_x_at	99.5	CAPN3	825	15q15.1-q21.1
219470_x_at	**99.5**	**CCNJ**	**54619**	**10pter-q26.12**
229091_s_at	99.5	CCNJ	54619	10pter-q26.12
239956_at	99.5	CHST2?	---	3q23
1552398_a_at	**99.5**	**CLEC12A /// CLEC12B**	**160364 /// 387837**	**12p13.2**
219821_s_at	99.5	GFOD1	54438	6pter-p22.1
239533_at	99.5	GPR155	151556	2q31.1
202409_at	**99.5**	**IGF2 /// INS-IGF2**	**3481 /// 723961**	**11p15.5**
230179_at	99.5	LOC285812	285812	6p23
202819_s_at	99.5	TCEB3	6924	1p36.1
232081_at	99	ABCG1?	---	21q22.3
1561786_at	99	AGAP1*	---	2q37.2

1559280_a_at	99	AK092578*	---	4q32.3
1554486_a_at	99	C6orf114	85411	6p23
1558621_at	99	CABLES1	91768	18q11.2
203921_at	**99**	**CHST2**	**9435**	**3q24**
209087_x_at	99	MCAM	4162	11q23.3
211340_s_at	**99**	**MCAM**	**4162**	**11q23.3**
223130_s_at	99	MYLIP	29116	6p23-p22.3
228098_s_at	99	MYLIP	29116	6p23-p22.3
226814_at	98.5	ADAMTS9	56999	3p14.3-p14.2
238987_at	98.5	B4GALT1	2683	9p13
225499_at	98.5	c20orf74?	---	20p11.23
1556593_s_at	98.5	CHST2?	---	3q23
231600_at	98.5	CLEC12B	387837	12p13.2
214683_s_at	98.5	CLK1	1195	2q33
201656_at	**98.5**	**ITGA6**	**3655**	**2q31.1**
202746_at	98.5	ITM2A	9452	Xq13.3-Xq21.2
210869_s_at	**98.5**	**MCAM**	**4162**	**11q23.3**
1569484_s_at	98.5	MDN1	23195	6q15
228097_at	98.5	MYLIP	29116	6p23-p22.3
229407_at	98.5	SDK1	221935	7p22.2
209593_s_at	98.5	TOR1B	27348	9q34
222281_s_at	98	c1orf186*	---	1q32.1
239826_at	98	CABLES1*	---	18q11.2
214475_x_at	98	CAPN3	825	15q15.1-q21.1
210944_s_at	98	CAPN3	825	15q15.1-q21.1
1556592_at	98	CHST2?	---	3q23
211623_s_at	98	FBL	2091	19q13.1
234339_s_at	98	GLTSCR2	29997	19q13.3
225330_at	98	IGF1R	3480	15q26.3
212978_at	98	LRRC8B	23507	1p22.2
215692_s_at	**98**	**MPPED2**	**744**	**11p13**
205413_at	**98**	**MPPED2**	**744**	**11p13**
223129_x_at	98	MYLIP	29116	6p23-p22.3
232280_at	98	SLC25A29	123096	14q32.2
202818_s_at	98	TCEB3	6924	1p36.1
225127_at	98	TMEM181	57583	6q25.3
241535_at	**97.5**	**---**	**---**	**2p25.3**
233867_at	97.5	AKAP13*	---	15q25.3
212702_s_at	97.5	BICD2	23299	9q22.31
224435_at	97.5	C10orf57 /// C10orf58	80195 /// 84293	10q22.3 /// 10q23.1
242406_at	97.5	c1orf186*	---	1q32.1

230954_at	97.5	C20orf112	140688	20q11.1-q11.23
220331_at	97.5	CYP46A1	10858	14q32.1
204836_at	97.5	GLDC	2731	9p22
215177_s_at	**97.5**	**ITGA6**	**3655**	**2q31.1**
230591_at	97.5	LOC729887	729887	16q24.1
227805_at	97.5	MAP1D?	---	2q31.1
209086_x_at	97.5	MCAM	4162	11q23.3
223627_at	97.5	MEX3B	84206	15q25.2
220319_s_at	97.5	MYLIP	29116	6p23-p22.3
223096_at	97.5	NOP5/NOP58	51602	2q33.1
243612_at	97.5	NSD1	64324	5q35.2-q35.3
214620_x_at	97.5	PAM	5066	5q14-q21
202336_s_at	**97.5**	**PAM**	**5066**	**5q14-q21**
242664_at	97.5	PTPRM*	---	18p11.23
226342_at	97.5	SPTBN1	6711	2p21
229594_at	97.5	SPTY2D1	144108	11p15.1
239361_at	97	CABLES1*	---	18q11.2
220450_at	97	---	---	4q31.22
204567_s_at	97	ABCG1	9619	21q22.3
229720_at	97	BAG1	573	9p12
243409_at	97	FOXL1	2300	16q24
202747_s_at	97	ITM2A	9452	Xq13.3-Xq21.2
212658_at	97	LHFPL2	10184	5q14.1
225611_at	97	LOC100128443 /// MAST4	100128443 /// 375449	5q12.3
212239_at	97	PIK3R1	5295	5q13.1
226143_at	97	RAI1	10743	17p11.2
1552329_at	97	RBBP6	5930	16p12.2
225305_at	97	SLC25A29	123096	14q32.2

Table S15. Top 100 Rank Order Genes Defining ROSE Cluster 8 (R8)

Probeset	Rank	Symbol	EntrezID	Cytoband
238689_at	**100.0**	**GPR110**	**266977**	**6p12.3**
235988_at	**100.0**	**GPR110**	**266977**	**6p12.3**
236489_at	**100.0**	**GPR110?**	**---**	**6p12.3**
217109_at	**100.0**	**MUC4**	**4585**	**3q29**
217110_s_at	**99.5**	**MUC4**	**4585**	**3q29**
205795_at	99.5	NRXN3	9369	14q31
216565_x_at	99.0	---	---	1p36.11
214022_s_at	99.0	IFITM1	8519	11p15.5
201601_x_at	99.0	IFITM1	8519	11p15.5
204895_x_at	**99.0**	**MUC4**	**4585**	**3q29**

206873_at	98.5	CA6	765	1p36.2
201028_s_at	98.5	CD99	4267	Xp22.32; Yp11.3
242051_at	98.5	CD99?	---	Xp22.32; Yp11.3
240586_at	98.5	ENAM	10117	4q13.3
212592_at	**98.5**	**IGJ**	**3512**	**4q21**
223304_at	98.5	SLC37A3	84255	7q34
1569666_s_at	98.5	SLC37A3*	---	7q34
238063_at	98.5	TMEM154	201799	4q31.3
207900_at	98.0	CCL17	6361	16q13
201029_s_at	98.0	CD99	4267	Xp22.32; Yp11.3
214907_at	98.0	CEACAM21	90273	19q13.2
201315_x_at	98.0	IFITM2	10581	11p15.5
222154_s_at	98.0	LOC26010	26010	2q33.1
211675_s_at	98.0	MDFIC	29969	7q31.1-q31.2
239272_at	98.0	MMP28	79148	17q11-q21.1
212183_at	98.0	NUDT4 /// NUDT4P1	11163 /// 440672	12q21 /// 1q21.1
212181_s_at	98.0	NUDT4 /// NUDT4P1	11163 /// 440672	12q21 /// 1q21.1
220024_s_at	98.0	PRX	57716	19q13.13-q13.2
207426_s_at	98.0	TNFSF4	7292	1q25
208303_s_at	**97.4**	**CRLF2**	**64109**	**Xp22.3; Yp11.3**
205983_at	97.4	DPEP1	1800	16q24.3
207651_at	**97.4**	**GPR171**	**29909**	**3q25.1**
213371_at	**97.4**	**LDB3**	**11155**	**10q22.3-q23.2**
1559315_s_at	97.4	LOC144481	144481	12q22
226382_at	97.4	LOC283070	283070	10p14
229334_at	97.4	RUFY3	22902	4q13.3
225244_at	97.4	SNAP47	116841	1q42.13
203372_s_at	97.4	SOCS2	8835	12q
244721_at	97.4	TP53INP1	94241	8q22
218862_at	96.9	ASB13	79754	10p15.1
206150_at	96.9	CD27	939	12p13
218013_x_at	96.9	DCTN4	51164	5q31-q32
219777_at	96.9	GIMAP6	474344	---
233884_at	96.9	HIVEP3	59269	1p34
203435_s_at	96.9	MME	4311	3q25.1-q25.2
239273_s_at	96.9	MMP28	79148	17q11-q21.1
202149_at	96.9	NEDD9	4739	6p25-p24
205259_at	96.9	NR3C2	4306	4q31.1
215021_s_at	96.9	NRXN3	9369	14q31
236750_at	96.9	NRXN3*	---	14q31.1
228696_at	96.9	SLC45A3	85414	1q32.1

223741_s_at	96.9	TTYH2	94015	17q25.1
219141_s_at	96.4	AMBRA1	55626	11p11.2
230161_at	96.4	CD99*	---	Xp22.32; Yp11.3
223377_x_at	96.4	CISH	1154	3p21.3
229114_at	96.4	GAB1	2549	4q31.21
1552316_a_at	96.4	GIMAP1	170575	7q36.1
229649_at	96.4	NRXN3	9369	14q31
226433_at	96.4	RNF157	114804	17q25.1
220454_s_at	96.4	SEMA6A	57556	5q23.1
225660_at	96.4	SEMA6A	57556	5q23.1
230747_s_at	96.4	TTC39C	125488	18q11.2
1555194_at	96.4	TTC39C*	---	18q11.2
203756_at	95.9	ARHGEF17	9828	11q13.4
242579_at	95.9	BMPR1B	658	4q22-q24
212974_at	95.9	DENND3	22898	8q24.3
217967_s_at	95.9	FAM129A	116496	1q25
226002_at	95.9	GAB1	2549	4q31.21
207375_s_at	95.9	IL15RA	3601	10p15-p14
208071_s_at	95.9	LAIR1	3903	19q13.4
210644_s_at	95.9	LAIR1	3903	19q13.4
215020_at	95.9	NRXN3	9369	14q31
238297_at	95.9	PHACTR1*	---	6p24.1
210830_s_at	95.9	PON2	5445	7q21.3
203373_at	95.9	SOCS2	8835	12q
225912_at	95.9	TP53INP1	94241	8q22
225108_at	95.4	AGPS	8540	2q31.2
229975_at	**95.4**	**BMPR1B**	**658**	**4q22-q24**
202910_s_at	95.4	CD97	976	19p13
216605_s_at	95.4	CEACAM21	90273	19q13.2
229604_at	95.4	CMAH	8418	6p21.32
1556037_s_at	95.4	HHIP	64399	4q28-q32
244764_at	95.4	HIVEP3*	---	1p34.2
222762_x_at	95.4	LIMD1	8994	3p21.3
236632_at	95.4	LOC646576	646576	4q31.22
240457_at	95.4	NEURL1B*	---	5q35.1
1553995_a_at	95.4	NT5E	4907	6q14-q21
219812_at	95.4	PVRIG	79037	7q22.1
52731_at	94.9	AMBRA1	55626	11p11.2
236766_at	94.9	C8orf38*	---	8q22.1
221223_x_at	94.9	CISH	1154	3p21.3
209210_s_at	94.9	FERMT2	10979	14q22.2

238880_at	94.9	GTF3A	2971	13q12.3-q13.1
212203_x_at	94.9	IFITM3	10410	11p15.5
209695_at	94.9	LOC100131062 /// PTP4A3	100131062 /// 11156	8q24.3
51146_at	94.9	PIGV	55650	1p36.11
219238_at	94.9	PIGV	55650	1p36.11
48106_at	94.9	SLC48A1	55652	12q13.11
226838_at	94.9	TTC32	130502	2p24.1
230643_at	94.9	WNT9A	7483	1q42

Table S16. Top 100 Rank Order Genes Associated with Unclustered ROSE Samples (R7)

Probeset	Percentile	Symbol	EntrezID	Cytoband
220230_s_at	96.2	CYB5R2	51700	11p15.4
212188_at	93.7	KCTD12	115207	13q22.3
242593_at	93.1	---	---	?
1564878_at	93.1	---	---	12q24.23-q24.31
227435_at	93.1	KIAA2018	205717	3q13.2
226869_at	93.1	MEGF6	1953	1p36.3
200866_s_at	93.1	PSAP	5660	10q21-q22
212956_at	93.1	TBC1D9	23158	4q31.21
205987_at	91.8	CD1C	911	1q22-q23
229288_at	91.8	EPHA7	2045	6q16.1
229716_at	91.2	---	---	1p36.12
1556682_s_at	91.2	AUTS2*	---	7q11.22
226640_at	91.2	DAGLB	221955	7p22.1
238533_at	91.2	EPHA7	2045	6q16.1
204396_s_at	91.2	GRK5	2869	10q24-qter
240413_at	91.2	PYHIN1	149628	1q23.1
213164_at	91.2	SLC5A3	6526	21q22.12
242644_at	91.2	TMC8	147138	17q25.3
237946_at	90.6	---	---	11p15.4
229967_at	90.6	CMTM2	146225	16q21
221773_at	90.6	ELK3	2004	12q23
205718_at	90.6	ITGB7	3695	12q13.13
212192_at	**90.6**	**KCTD12**	**115207**	**13q22.3**
1559263_s_at	90.6	PPIL4 /// ZC3H12D	340152 /// 85313	6q24-q25 /// 6q25.1
218613_at	90.6	PSD3	23362	8pter-p23.3
203355_s_at	90.6	PSD3	23362	8pter-p23.3
221808_at	90.6	RAB9A	9367	Xp22.2
227210_at	90.6	SFMBT2?	---	10p14

202912_at	89.9	ADM	133	11p15.4
205290_s_at	89.9	BMP2	650	20p12
219837_s_at	**89.9**	**CYTL1**	**54360**	**4p16-p15**
213316_at	89.9	KIAA1462	57608	10p11.23
210629_x_at	89.9	LST1	7940	6p21.3
220122_at	89.9	MCTP1	79772	5q15
214735_at	89.9	PIP3-E	26034	6q25.2
209568_s_at	89.9	RGL1	23179	1q25.3
226207_at	89.9	RILPL1	353116	12q24.31
212944_at	89.9	SLC5A3	6526	21q22.12
207777_s_at	89.9	SP140	11262	2q37.1
226080_at	89.9	SSH2	85464	17q11.2
230590_at	89.9	SSH2*	---	17q11.2
223375_at	89.9	TBC1D22B	55633	6p21.2
224967_at	89.9	UGCG	7357	9q31
213618_at	89.3	ARAP2	116984	4p14
203923_s_at	89.3	CYBB	1536	Xp21.1
225833_at	89.3	DAGLB	221955	7p22.1
214574_x_at	89.3	LST1	7940	6p21.3
207339_s_at	89.3	LTB	4050	6p21.3
217418_x_at	89.3	MS4A1	931	11q12
200871_s_at	89.3	PSAP	5660	10q21-q22
216748_at	89.3	PYHIN1	149628	1q23.1
204688_at	89.3	SGCE	8910	7q21-q22
204328_at	89.3	TMC6	11322	17q25.3
227353_at	89.3	TMC8	147138	17q25.3
233596_at	89.3	UIMC1*	---	5q35.2
229040_at	88.7	BC40064*	---	21q22.3
203922_s_at	88.7	CYBB	1536	Xp21.1
204057_at	88.7	IRF8	3394	16q24.1
218656_s_at	88.7	LHFP	10186	13q12
211101_x_at	88.7	LILRA2	11027	19q13.4
239062_at	88.7	LOC100131096	100131096	17q25.3
206940_s_at	88.7	LOC100131317 /// POU4F1	100131317 /// 5457	13q31.1
211581_x_at	88.7	LST1	7940	6p21.3
244230_at	88.7	MEF2C*	---	5q14.3
1569136_at	88.7	MGAT4A	11320	2q12
1569931_at	88.7	NCOR2*	---	12q24.31
241387_at	88.7	PTK2*	---	8q24.3
41220_at	88.7	SEPT9*	10801	17q25.2-q25.3
208657_s_at	88.7	SEPT9*	10801	17q25.2-q25.3

231837_at	88.7	USP28	57646	11q23
1552678_a_at	88.7	USP28	57646	11q23
236635_at	88.7	ZNF667	63934	19q13.43
231418_at	88.1	---	---	11q12.2
229041_s_at	88.1	BC40064*	---	21q22.3
205289_at	88.1	BMP2	650	20p12
37170_at	88.1	BMP2K	55589	4q21.21
225828_at	88.1	DAGLB	221955	7p22.1
214966_at	88.1	GRIK5	2901	19q13.2
1555349_a_at	88.1	ITGB2	3689	21q22.3
227433_at	88.1	KIAA2018	205717	3q13.2
232935_at	88.1	LHFP*	---	13q13.3
215633_x_at	88.1	LST1	7940	6p21.3
214181_x_at	88.1	LST1	7940	6p21.3
242191_at	88.1	NBPF10 /// RP11-94I2.2	100132406 /// 200030	1q21.1
209949_at	88.1	NCF2	4688	1q25
206370_at	88.1	PIK3CG	5294	7q22.3
203038_at	88.1	PTPRK	5796	6q22.2-q22.3
204319_s_at	88.1	RGS10	6001	10q25
220922_s_at	88.1	SPANXA1 /// SPANXA2 /// SPANXB1 /// SPANXB2 /// SPANXC /// SPANXF1	100133171 /// 171490 /// 30014 /// 64663 /// 728695 /// 728712	Xq27.1
230970_at	88.1	SSH2*	---	17q11.2
222942_s_at	88.1	TIAM2	26230	6q25.2
214958_s_at	88.1	TMC6	11322	17q25.3
204881_s_at	88.1	UGCG	7357	9q31
221765_at	88.1	UGCG	7357	9q31
220586_at	87.4	CHD9	80205	16q12.2
229268_at	87.4	FAM105B	90268	5p15.2
225140_at	87.4	KLF3	51274	4p14
244741_s_at	87.4	MGC9913	386759	19q13.43
231199_at	87.4	NAT13*	---	3q13.2
235652_at	87.4	SCML1*	---	Xp22.2

Table S17. Top 100 Ross[1] BCR-ABL Probe Sets Compared to ROSE Clustering and Top Rank Order

Probe Set ID	Gene Symbol	Cytoband	ROSE Clustering	Rank Order Group
224811_at	---	---		
226345_at	---	---		
240173_at	---	---		
240499_at	---	---		
202123_s_at	ABL1	9q34.1		R4

209321_s_at	ADCY3	2p23.3		
223075_s_at	AIF1L	9q34.13-q34.3		
214255_at	ATP10A	15q11.2		
219218_at	BAHCC1	17q25.3		
229975_at	BMPR1B	4q22-q24	Yes	R8
242579_at	BMPR1B	4q22-q24	Yes	R8
201310_s_at	C5orf13	5q22.1		
200655_s_at	CALM1	14q24-q31		
205467_at	CASP10	2q33-q34		
200951_s_at	CCND2	12p13		
200953_s_at	CCND2	12p13		
206150_at	CD27	12p13		R8
201028_s_at	CD99	Xp22.32; Yp11.3		R8
201029_s_at	CD99	Xp22.32; Yp11.3		R8
242051_at	CD99*	---		R8
202717_s_at	CDC16	13q34		
212862_at	CDS2	20p13		
213385_at	CHN2	7p15.3		
204576_s_at	CLUAP1	16p13.3		
201445_at	CNN3	1p22-p21	Yes	R5
228297_at	CNN3*	---	Yes	R5
201906_s_at	CTDSPL	3p21.3		
218013_x_at	DCTN4	5q31-q32		R8
222488_s_at	DCTN4	5q31-q32		R8
209365_s_at	ECM1	1q21		
217967_s_at	FAM129A	1q25		R8
202771_at	FAM38A	16q24.3		
222729_at	FBXW7	4q31.3		
219871_at	FLJ13197	4p14		
218084_x_at	FXYD5	19q12-q13.1		
216033_s_at	FYN	6q21		
64064_at	GIMAP5	7q36.1		
229367_s_at	GIMAP6	---		
235988_at	GPR110	6p12.3	Yes	R8
238689_at	GPR110	6p12.3	Yes	R8
236489_at	GPR110*	---	Yes	R8
202947_s_at	GYPC	2q14-q21		R4
203089_s_at	HTRA2	2p12		
208881_x_at	IDI1	10p15.3		
212203_x_at	IFITM3	11p15.5		R8
212592_at	IGJ	4q21	Yes	R8

222868_s_at	IL18BP	11q13		
202794_at	INPP1	2q32		
205376_at	INPP4B	4q31.21		
201656_at	ITGA6	2q31.1	Yes	R6
205055_at	ITGAE	17p13		
229139_at	JPH1	8q21		
208071_s_at	LAIR1	19q13.4		R8
205269_at	LCP2	5q33.1-qter		
205270_s_at	LCP2	5q33.1-qter		
222762_x_at	LIMD1	3p21.3		R8
215617_at	LOC26010	2q33.1		R8
222154_s_at	LOC26010	2q33.1		R8
241812_at	LOC26010	2q33.1		R8
225799_at	LOC541471 /// NCRNA00152	2p11.2 /// 2q13		
238488_at	LRRC70	5q12.1		
203005_at	LTBR	12p13		
239273_s_at	MMP28	17q11-q21.1		R8
217110_s_at	MUC4	3q29	Yes	R8
218966_at	MYO5C	15q21		
205259_at	NR3C2	4q31.1		R8
212298_at	NRP1	10p12		
239519_at	NRP1*	---		
204004_at	PAWR	12q21		
201876_at	PON2	7q21.3		R8
210830_s_at	PON2	7q21.3		R8
213093_at	PRKCA	17q22-q23.2		
218764_at	PRKCH	14q22-q23		
220024_s_at	PRX	19q13.13-q13.2		R8
219938_s_at	PSTPIP2	18q12		
200863_s_at	RAB11A	15q21.3-q22.31		
200864_s_at	RAB11A	15q21.3-q22.31		
209229_s_at	SAPS1	19q13.42		
215028_at	SEMA6A	5q23.1		R8
223449_at	SEMA6A	5q23.1		R8
225660_at	SEMA6A	5q23.1		R8
225913_at	SGK269	15q24.3		
204429_s_at	SLC2A5	1p36.2		
204430_s_at	SLC2A5	1p36.2		
48106_at	SLC48A1	12q13.11		R8
225244_at	SNAP47	1q42.13		R8

200665_s_at	SPARC	5q31.3-q32		
212458_at	SPRED2	2p14		
203217_s_at	ST3GAL5	2p11.2		
216985_s_at	STX3	11q12.1		
220684_at	TBX21	17q21.32		R4
219315_s_at	TMEM204	16p13.3		
203508_at	TNFRSF1B	1p36.3-p36.2		
207196_s_at	TNIP1	5q32-q33.1		
200742_s_at	TPP1	11p15		
202369_s_at	TRAM2	6p21.1-p12		
202242_at	TSPAN7	Xp11.4		
212242_at	TUBA4A	2q35		
218348_s_at	ZC3H7A	16p13-p12		
228046_at	ZNF827	4q31.22		

Table S18. Genes/Probe Sets Common to Rank Order and *BCR-ABL1*-like Signature[2]

BCR-ABL up-regulated

Gene	Cluster
216565_x_at	R8
ABL1	R4
AGPS	R4/R8
CA6	R8
CD97	R8
CD99	R8
CNN3	R5
DCTN4	R8
GIMAP6	R8
GYPC	R4
HIVEP2	R6
IFITM1	R8
IFITM3	R8
IGJ	R8
IL2RA	R6
LIMD1	R8
MMP28	R8
MUC4	R8
PON2	R8
PRX	R8
SEMA6A	R8

BCR-ABL down-regulated

Gene	Cluster
BACH2	R2
CSF2RB	R3
CYP46A1	R6
IRS1	R2
KIAA0922	R3
LY9	R4
PHYH	R6
WWC3	R2

SLC5A3	R7
TBXA2R	R4

Table S19. Genes/Probe Sets Defining CCG 1961 ROSE Clusters

Bold font indicates that these genes or probe sets were also among the best ROSE rank order from P9906. Bold and italicized (in R9) indicates that these were associated with the *TEL-AML1* translocation defining genes from the Ross paper. Italicized probe sets (in R10) indicate those that were also outlier probe sets in P9906 but were not among the top rank order for any cluster in that cohort. Asterisks indicate those for which Affymetrix does not specify a gene, however the probe sets were mapped using the UCSC Genome Browser (http://genome.ucsc.edu/) between exons of the specified genes. Question marks highlight those that also lack Affymetrix gene data, but were mapped within 10 kb of the indicated gene using the UCSC Genome Browser.

R1		R2		R6		R8		R9		R10	
1559477_s_at	**MEIS1**	1560109_s_at	NUB1*	**1555578_at**	**PTPRM**	1554008_at	OSMR	1560225_at	CNR1	203835_at	LRRC32
204069_at	**MEIS1**	**201579_at**	**FAT1**	**205413_at**	**MPPED2**	**204895_x_at**	**MUC4**	203373_at	SOCS2	205870_at	BDKRB2
205899_at	**CCNA1**	**202178_at**	**PRKCZ**	**206756_at**	**CHST7**	205544_s_at	CR2	204913_s_at	SOX11	225977_at	PCDH18
209905_at	**HOXA9**	202289_s_at	TACC2	209291_at	ID4	**212592_at**	**IGJ**	204914_s_at	SOX11	227372_s_at	BAIAP2L1
214651_s_at	**HOXA9**	204014_at	DUSP4	**227289_at**	**PCDH17**	**213371_at**	**LDB3**	204915_s_at	SOX11	*228057_at*	*DDIT4L*
232231_at	**RUNX2**	204015_s_at	DUSP4	**228863_at**	**PCDH17**	213906_at	MYBL1	***205109_s_at***	***ARHGEF4***	230030_at	HS6ST2
242172_at	**MEIS1**	**204114_at**	**NID2**	**231600_at**	**CLEC12B**	**217109_at**	**MUC4**	213050_at	COBL	230680_at	SLITRK4?
		205253_at	**PBX1**			**217110_s_at**	**MUC4**	***213317_at***	***CLIC5***	*232636_at*	*SLITRK4*
		205489_at	**CRYM**			228994_at	CCDC24	213423_x_at	TUSC3		
		206181_at	SLAMF1			**229975_at**	**BMPR1B**	218922_s_at	LASS4		
		209570_s_at	**D4S234E**			**235988_at**	**GPR110**	219855_at	NUDT11		
		211382_s_at	TACC2			**236489_at**	**GPR110?**	220276_at	RERGL		
		212148_at	**PBX1**					220922_s_at	SPANXA1//		
		212151_at	**PBX1**					227377_at	IGF2BP1		
		213005_s_at	**KANK1**					228116_at	---		
		222450_at	PMEPA1					229108_at	RASGRF2?		
		223786_at	**CHST6**					231455_at	FLJ42418		
		224022_x_at	**WNT16**					236430_at	TMED6		
		225369_at	ESAM					238617_at	KIF26B?		
		226034_at	DUSP4?					240245_at	RBMS3*		
		226913_s_at	SOX8					***243917_at***	***CLIC5***		
		227949_at	**PHACTR3**								
		228262_at	**MAP7D2**								
		229247_at	FBLN7								
		231040_at	**RORB?**								
		231257_at	TCERG1L								
		234985_at	LDLRAD3								
		238778_at	**MPP7**								
		239657_x_at	**FOXO6**								

GENOME-WIDE COPY NUMBER VARIATION ASSOCIATION WITH ROSE CLUSTER GROUPS

Table S20. Copy Number Analysis (CNA) Variations Associated with ROSE Clusters[1]
The CNA variations are shown along with their membership in each ROSE cluster. FET indicates the *P* value for this results as determined by Fisher's Exact Test. CNA variations are sorted in ascending order by *P* values.

	R1	**R2**	**R3**	**R4**	**R5**	**R6**	**R7**	**R8**	**Total**	**FET**	
Cases evaluated	20	22	12	13	10	21	76	24		**p-value**	**Comments**
1q (gain)[2]	0	14	1	1	0	0	1	0	17	<0.0001	R2 contains *TCF3-PBX1*
IKZF1	1	0	0	3	3	6	24	22	59	<0.0001	
CDKN2A/B	4	9	11	11	1	5	40	15	96	<0.0001	
TCF3	0	14	0	0	2	2	2	0	20	<0.0001	R2 contains *TCF3-PBX1*
ERG	0	0	0	1	0	8	0	0	9	<0.0001	
VPREB1	0	0	0	5	1	8	23	14	51	<0.0001	
B cell pathway	5	17	5	12	4	12	54	23	132	<0.0001	
B pathway w/VPREB1	5	17	5	12	5	14	56	24	138	<0.0001	
PAX5	1	9	4	11	0	3	28	7	63	<0.0001	
JAK1/2 **(mutation)**	0	0	0	0	0	0	2/5	1/11	0	<0.0001	
EBF1	0	0	0	0	0	0	4	9	13	0.0001	
TBL1XR1	0	0	3	0	1	1	0	0	5	0.0005	
NUP160-PTPRJ	0	0	0	0	0	0	0	4	4	0.0028	
ETV6	1	0	2	2	4	1	14	0	24	0.0055	
IL3RA-CSF2RA	0	0	0	0	1	0	6	7	14	0.0064	High *CRLF2* expression
C20orf94	0	0	0	1	1	0	7	7	16	0.0102	
DMD	0	5	1	0	2	3	3	0	14	0.0109	
RAG1/2	1	0	0	0	0	0	1	5	7	0.0144	
ADD3	0	1	0	2	0	0	7	7	17	0.0156	
GRIK2	0	2	0	2	2	0	2	4	12	0.0195	
NF1	1	1	0	0	2	0	0	1	5	0.0269	
ARMC2-SESN1	0	2	0	1	2	0	3	5	13	0.0297	
ADARB2	0	0	0	0	0	2	0	2	4	0.0658	
BTLA-CD200	0	0	0	1	0	0	5	5	11	0.0764	
BTG1	0	0	1	1	2	2	8	6	20	0.0843	
ELF1	0	5	0	1	1	0	5	1	13	0.0971	
IL1RAP	0	0	2	0	0	0	1	0	3	0.1205	
FLNB	0	0	0	0	0	2	1	2	5	0.2306	

[1] All p-values are derived from Fisher's Exact Test.
[2] All abnormalities are losses unless otherwise indicated

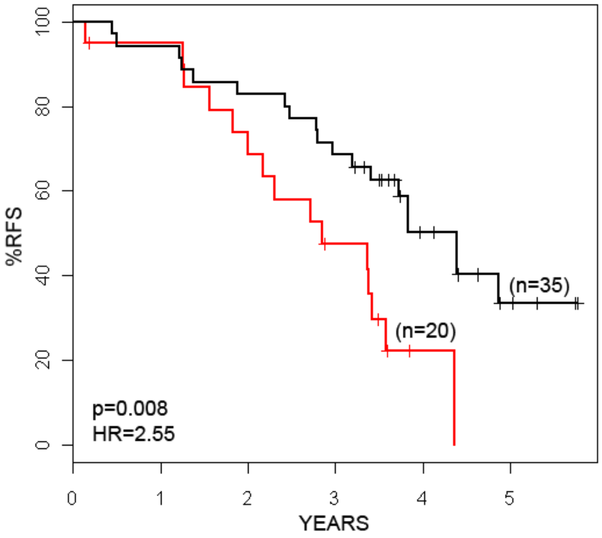

Figure S3. Survival of *IKZF1*-positive patients in R8 compared to not-R8. *IKZF1*-positive patients were divided into those in cluster 8 (red line) and those in other clusters (black line). The p-value and hazard ratio for this comparison are given in the lower left panel.

CLUSTERING OF CCG 1961 WITH R8 PROBE SETS

In order to further illustrate that the probe sets identified by ROSE in P9906 and CCG 1961 are highly related and associated with outcome, several of the R8 probe sets from P9906 that were both among the outliers selected and also included in the top rank order of R8 were used to generate a heat map with CCG 1961 samples. Figure S4 shows the results of this clustering and highlights in black those samples that relapse. It is clear that a small subset of these probe sets (representing only 4 genes: *MUC4*, *IGJ*, *GPR110* and *LDB3*) identify the majority of patients who relapse.

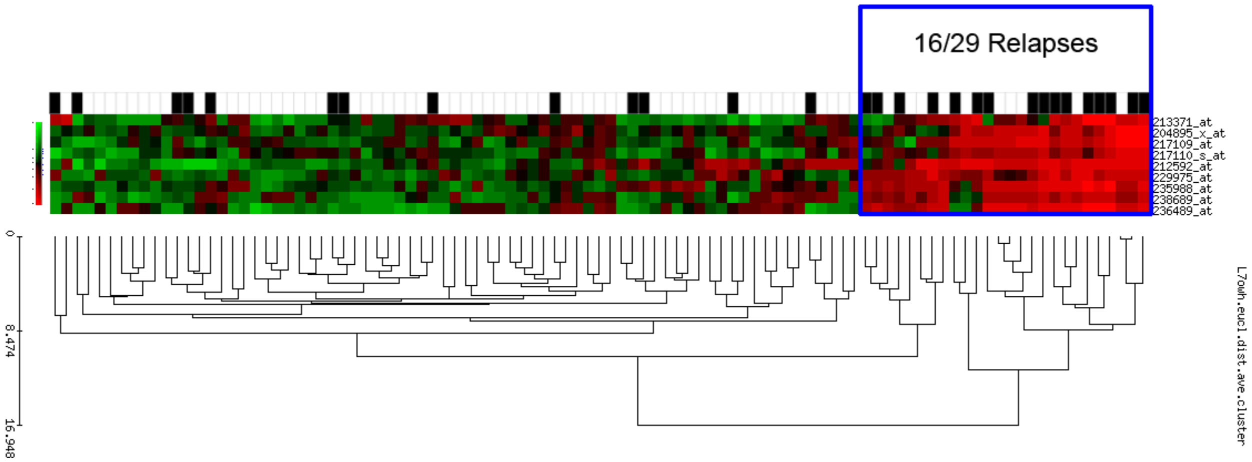

Figure S4. Heat map of CCG 1961 Samples Using P9906 R8 Probe Sets. Nine probe sets, representing four genes, were selected from the most highly conserved from the ROSE P9906 clustering. These nine probe sets were among the 254 used for P9906 clustering and also were among the top rank order probe sets for R8. Black boxes across the top of the heat map indicate patients who relapsed. Sixteen of the 29 relapses (55%) occur within a group of samples having greatly elevated expression of two or more of these genes.

REFERENCES

1. Ross ME, Zhou XD, Song GC, et al. Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. *Blood*. 2003;102(8):2951-2959.

2. Mullighan CG, Su X, Zhang J, et al. Deletion of *IKZF1* and prognosis in acute lymphoblastic leukemia. *N Engl J Med*. 2009;360(5):470-480.

3. Borowitz MJ, Devidas M, Hunger SP, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia and its relationship to other prognostic factors: a Children's Oncology Group study. *Blood*. 2008;111(12):5477-5485.

4. Bhojwani D, Kang H, Menezes RX, et al. Gene expression signatures predictive of early response and outcome in high-risk childhood acute lymphoblastic leukemia: a Children's Oncology Group Study on behalf of the Dutch Childhood Oncology Group and the German Cooperative Study Group for Childhood Acute Lymphoblastic Leukemia. *J Clin Oncol*. 2008;26(27):4376-4384.

5. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of *TMPRSS2* and *ETS* transcription factor genes in prostate cancer. *Science*. 2005;310(5748):644-648.