

Supplemental Information for:

**CLASSIFICATION OF PEDIATRIC ACUTE LYMPHOBLASTIC
LEUKEMIA BY GENE EXPRESSION PROFILING**

Mary E. Ross¹, Xiaodong Zhou², Guangchun Song², Sheila A. Shurtleff², Kevin Girtman², W.
Kent Williams², Hsi-Che Liu², Rami Mahfouz², Susana C. Raimondi², Noel Lenny², Anami
Patel², & James R. Downing^{2,*}

Table of Contents

Section I: Patient Dataset

Diagnostic ALL samples used for class prediction
Subgroup distribution of ALL cases

Section II: Methods

Hybridization of microarrays
Statistical methods

Section III: Genetic Subtype Discriminating Genes

Top 100 chi-square probe sets selected for *BCR-ABL* decision tree format
Top 100 chi-square probe sets selected for *E2A-PBX1* decision tree format
Top 100 chi-square probe sets selected for Hyperdiploid >50 decision tree format
Top 100 chi-square probe sets selected for *MLL* decision tree format
Top 100 chi-square probe sets selected for T-ALL decision tree format
Top 100 chi-square probe sets selected for *TEL-AML1* decision tree format

Top 100 chi-square probe sets selected for *BCR-ABL* parallel format
Top 100 chi-square probe sets selected for *E2A-PBX1* parallel format
Top 100 chi-square probe sets selected for Hyperdiploid >50 parallel format
Top 100 chi-square probe sets selected for *MLL* parallel format
Top 100 chi-square probe sets selected for T-ALL parallel format
Top 100 chi-square probe sets selected for *TEL-AML1* parallel format

Section IV: Diagnostic Accuracy

Training and test set results
Cross comparison of supervised learning algorithms

Section V: Comparison of Expression Profiles and Real-time PCR (Taqman)

Section VI: References

I: Patient Dataset

132 cases of pediatric ALL were selected from the original 327 diagnostic bone marrow aspirates¹ to reanalyze on the higher density U133A and B microarrays. The selection of cases was based on having sufficient numbers of each subtype to build accurate class predictions, rather than reflecting the actual frequency of these groups in the pediatric population. The list of samples that were used in this reanalysis (Table S1), as well as the subtype distribution (Table S2) are shown below.

| | | |
|----------------------|---------------------|-----------------|
| BCR-ABL-#1 | Hyperdip>50-C18 | Pseudodip-#6 |
| BCR-ABL-#2 | Hyperdip>50-C21 | Pseudodip-C2-N |
| BCR-ABL-#3 | Hyperdip>50-C22 | Pseudodip-C3 |
| BCR-ABL-#4 | Hyperdip>50-C23 | Pseudodip-C5 |
| BCR-ABL-#5 | Hyperdip>50-C27-N | Pseudodip-C6 |
| BCR-ABL-#6 | Hyperdip>50-C32 | Pseudodip-C7 |
| BCR-ABL-#7 | Hyperdip>50-R4 | Pseudodip-C9 |
| BCR-ABL-#8 | Hyperdip47-50-C14-N | Pseudodip-C14 |
| BCR-ABL-#9 | Hyperdip47-50-C3-N | Pseudodip-C16-N |
| BCR-ABL-Hyperdip-#10 | Hypodip-#2 | Pseudodip-R1-N |
| BCR-ABL-C1 | Hypodip-2M#1 | T-ALL-#5 |
| BCR-ABL-R1 | Hypodip-C2 | T-ALL-#6 |
| BCR-ABL-R2 | Hypodip-C5 | T-ALL-#7 |
| BCR-ABL-R3 | MLL-#1 | T-ALL-#8 |
| BCR-ABL-Hyperdip-R5 | MLL-#2 | T-ALL-#10 |
| E2A-PBX1-#5 | MLL-#3 | T-ALL-C2 |
| E2A-PBX1-#6 | MLL-#4 | T-ALL-C6 |
| E2A-PBX1-#9 | MLL-#5 | T-ALL-C7 |
| E2A-PBX1-#10 | MLL-#6 | T-ALL-C11 |
| E2A-PBX1-#12 | MLL-#7 | T-ALL-C15 |
| E2A-PBX1-#13 | MLL-#8 | T-ALL-C19 |
| E2A-PBX1-2M#1 | MLL-2M#1 | T-ALL-C21 |
| E2A-PBX1-C2 | MLL-2M#2 | T-ALL-R5 |
| E2A-PBX1-C3 | MLL-C1 | T-ALL-R6 |
| E2A-PBX1-C4 | MLL-C2 | TEL-AML1-#6 |
| E2A-PBX1-C5 | MLL-C3 | TEL-AML1-#9 |
| E2A-PBX1-C6 | MLL-C4 | TEL-AML1-#10 |
| E2A-PBX1-C7 | MLL-C5 | TEL-AML1-#14 |
| E2A-PBX1-C9 | MLL-C6 | TEL-AML1-2M#1 |
| E2A-PBX1-C10 | MLL-R1 | TEL-AML1-2M#2 |

| | | |
|-----------------|--------------|--------------|
| E2A-PBX1-C11 | MLL-R2 | TEL-AML1-C4 |
| E2A-PBX1-C12 | MLL-R3 | TEL-AML1-C5 |
| E2A-PBX1-R1 | MLL-R4 | TEL-AML1-C6 |
| Hyperdip>50-#8 | Normal-C1-N | TEL-AML1-C26 |
| Hyperdip>50-#12 | Normal-C2-N | TEL-AML1-C28 |
| Hyperdip>50-#14 | Normal-C3-N | TEL-AML1-C30 |
| Hyperdip>50-C1 | Normal-C4-N | TEL-AML1-C31 |
| Hyperdip>50-C4 | Normal-C7-N | TEL-AML1-C32 |
| Hyperdip>50-C6 | Normal-C8 | TEL-AML1-C33 |
| Hyperdip>50-C8 | Normal-C9 | TEL-AML1-C34 |
| Hyperdip>50-C11 | Normal-C11-N | TEL-AML1-C37 |
| Hyperdip>50-C13 | Normal-R1 | TEL-AML1-C38 |
| Hyperdip>50-C15 | Normal-R2-N | TEL-AML1-C40 |
| Hyperdip>50-C16 | Pseudodip-#5 | TEL-AML1-R3 |

Table Key: The nomenclature used in this paper is identical to that used in Yeoh et. al.,¹ and thus should facilitate cross comparisons between the datasets. The nomenclature indicates disease status at the time of the initial study and has not been updated as this dataset was not selected to address the issue of outcome. No analysis has been performed in this study to identify expression profiles associated with outcome.

Subtype Name-C# Dx Sample of patient in CCR

Subtype Name-R# Dx Sample of patient who developed a hematologic relapse

Subtype Name-# Dx Sample used for subgroup classification only

Subtype Name-2M# Dx Sample of patient who later developed 2nd AML

Subtype Name-N Dx Sample in novel group

Table S2. Subgroup distribution of ALL cases

| Subgroup | Training Set | Test Set |
|------------------|--------------|----------|
| <i>BCR-ABL</i> | 11 | 4 |
| <i>E2A-PBX1</i> | 13 | 5 |
| Hyperdiploid >50 | 13 | 4 |
| <i>MLL</i> | 15 | 5 |
| T-ALL | 12 | 2 |
| <i>TEL-AML1</i> | 15 | 5 |
| Other | 21 | 7 |
| Total | 100 | 32 |

II: Methods

Hybridization of microarrays

Hybridization solutions from our previous U95A study had been stored at -80°C since their initial use. These solutions were thawed at 45°C, then microcentrifuged for 2 minutes to remove

any insoluble material from the mixture. The hybridization solutions were added to U133A chips and allowed to hybridize for 16 hours at 45°C. At the end of the incubation period, the hybridization solution was removed from each U133A chip and refrozen. Subsequently, the hybridizations were thawed and hybridized to the U133B chip.

A non-stringent wash buffer (6X SSPE, 0.01% Tween 20) was added to each chip cassette after the hybridization solution was removed and the cassette allowed to equilibrate to room temperature. The microarray cassettes were then placed on the fluidics station and the antibody amplification protocol performed. The arrays were washed at 25°C with the non-stringent buffer followed by a more stringent wash at 50°C with 100 mM MES, 0.1M NaCl₂, 0.01% Tween 20. The arrays were then stained with Streptavidin Phycoerythrin (SAPE, Molecular Probes, Eugene, OR) for 10 minutes at 25°C. Following another non-stringent wash, the arrays were hybridized for 10 minutes at 25°C with an antibody solution (100 mM MES, 1 M [Na⁺], 0.05% Tween 20, 2 mg/ml BSA, 0.1 mg/ml goat IgG, and 3 µg/ml biotinylated antibody). This solution was removed and the cassettes restained with the SAPE solution.

Arrays were scanned on a laser confocal scanner (Agilent, Palo Alto, CA) and then analyzed with Affymetrix Microarray Suite 5.0 (MAS 5.0). Detection values (present, marginal or absent) were determined by default parameters, and signal values were scaled by global methods to a target value of 500. After completing the scans, the arrays were visually inspected for defects and Affymetrix internal controls were utilized to monitor the success of hybridization, washing, and staining procedures.

Statistical methods

The chi-square metric and the *k*-NN and ANN supervised learning algorithms have been previously described. For more information see <http://www.stjuderesearch.org/data/ALL1/>. The

SVM supervised learning algorithm that was used in this study is available as part of the software package Rv 1.6.0.

To determine the performance of each model using ANN, a confidence threshold was built for each diagnostic subtype utilizing a modification of the method described by Khan et al.² Models were built based on a decision tree format where each level of the decision tree contains only two possible distinctions – class and non-class (for example, T verses non-T). At each level, using only samples in the training set, 3 ANN models were built by 3-fold cross validation. The training set samples were then shuffled and 3 additional ANN models were built. This model building process was repeated for a total of 100 times at each step of the decision tree. Then an empirical probability distribution for the ANN output node value was built only for subtype under study, for example, T-ALL at the first step of the decision tree. Only nodal values greater than 0.5 for each subtype were included. For each individual sample in the training set, the 100 validation subtype node values were averaged and compared to threshold. Individual samples were assigned to the subtype under study only when its average subtype nodal value was greater than the 95% confidence threshold. For samples in the test set, subtype nodal values are averaged from all models generated in the 3-fold cross validation. A sample is assigned to the class under study when the average subtype nodal value is greater than the 95% confidence level defined on the training set. A sample not assigned to the subtype will progress to the next level of the decision tree, where the entire process is repeated.

III: Genetic Subtype Discriminating Genes

The following tables contain a listing of the top 100 probe sets for each diagnostic subtypes ranked by their chi-square value (Tables S3-S8). Each table contains the Affymetrix U133 series probe set number, a gene description, gene symbol, chromosomal location, and primary GenBank reference. Chi-square values were calculated utilizing only the samples in the training set in a differential diagnosis decision tree format as discussed in the text and illustrated in Figure 2 (Tables S3-S8) or by a parallel approach (Tables S9-S14). The calculation of the fold change was done in a parallel format using the total data set and comparing the mean signal value in the class versus the mean signal value in the non-class. The last column indicates whether this gene had previously been identified as a class discriminator using the U95Av2 data (old) or identified as a class discriminator only using the U133 data (new) (Tables S3-S8).

Table S3. Top 100 chi-square probe sets selected for *BCR-ABL* in decision tree format

| | U133 probe set | Gene description | Gene symbol | Location | GenBank Reference | Chi-square value | BCR-ABL above/below mean | Fold change | old or new |
|---|-----------------------|---|--------------------|-----------------|--------------------------|-------------------------|---------------------------------|--------------------|-------------------|
| 1 | 241812_at | EST FLJ39877 | FLJ39877 | 2 | AV648669 | 47.4 | Above | 5.2 | new |
| 2 | 201876_at | Paraoxonase 2 | PON2 | 7q21.3 | NM_000305.1 | 47.2 | Above | 18.7 | old |
| 3 | 201028_s_at | Antigen identified by monoclonal antibodies 12E7, F21 and O13 | MIC2 | Xp22.32 | U82164.1 | 44.3 | Above | 2.6 | old |
| 4 | 200953_s_at | Cyclin D2 | CCND2 | 12p13 | NM_001759.1 | 42.3 | Above | 3.5 | old |
| 5 | 202947_s_at | glycophorin C (Gerbich blood group) | GYPC | 2q14-q21 | NM_002101.2 | 42.3 | Above | 3.1 | old |
| 6 | 223449_at | Semaphorin 6A | SEMA6A | 5q23.1 | AF225425.1 | 42.3 | Above | 4.3 | new |
| 7 | 201029_s_at | Antigen identified by monoclonal antibodies 12E7, F21 and O13 | MIC2 | Xp22.32 | NM_002414.1 | 41.2 | Above | 2.4 | old |

| | | | | | | | | | |
|----|-------------|--|---------------|----------------|-------------|------|-------|------|-----|
| 8 | 204429_s_at | Solute carrier family 2 (facilitated glucose/fructose transporter), member 5 | SLC2A5 | 1p36.2 | BE560461 | 41.2 | Above | 5 | old |
| 9 | 210830_s_at | Paraoxonase 2 | PON2 | 7q21.3 | AF001602.1 | 41.2 | Above | 23.6 | old |
| 10 | 215028_at | Semaphorin 6A | SEMA6A | 5 | AB002438.1 | 41.2 | Above | 4.5 | new |
| 11 | 220024_s_at | Periaxin | PRX | 19q13.13-q13.2 | NM_020956.1 | 41.2 | Above | 8.2 | new |
| 12 | 201906_s_at | HYA22 protein | HYA22 | 3p21.3 | NM_005808.1 | 41.1 | Above | 43.4 | old |
| 13 | 209365_s_at | Extracellular matrix protein 1 | ECM1 | 1q21 | U65932.1 | 41.1 | Above | 6 | old |
| 14 | 238689_at | GPR110 G protein-coupled receptor 110 | GPR110 | 6 | BG426455 | 41.1 | Above | 10.9 | new |
| 15 | 222154_s_at | DKFZP564A2416 unknown protein with a histone H5 signature. | DKFZP564A2416 | 2q33.1 | AK002064.1 | 40.4 | Above | 12.4 | new |
| 16 | 218084_x_at | FXYP domain-containing ion transport regulator 5 | FXYP5 | 19q12-q13.1 | NM_014164.2 | 38 | Above | 1.5 | new |
| 17 | 212242_at | Tubulin, alpha 1 (testis specific) | TUBA1 | 2q36.2 | AL565074 | 37 | Above | 3.2 | old |
| 18 | 201445_at | Calponin 3, acidic | CNN3 | 1p22-p21 | NM_001839.1 | 36.3 | Above | 10.8 | old |
| 19 | 202771_at | KIAA0233 gene product | KIAA0233 | 16q24.3 | NM_014745.1 | 36.3 | Above | 1.9 | old |
| 20 | 212298_at | Neuropilin 1 | NRP1 | 10p12 | BE620457 | 36.3 | Above | 13.8 | new |
| 21 | 212458_at | FLJ21897 | FLJ21897 | 2 | AW138902 | 36.3 | Above | 2.4 | new |
| 22 | 222488_s_at | Dynactin 4 (p62) | DCTN4 | 5q31-q32 | BE218028 | 36.3 | Above | 3.6 | new |
| 23 | 222762_x_at | LIM domains containing 1 | LIMD1 | 3p21.3 | AU144259 | 36.3 | Above | 2.6 | new |
| 24 | 200951_s_at | Cyclin D2 | CCND2 | 12p13 | NM_001759.1 | 35.3 | Above | 12.7 | old |
| 25 | 204430_s_at | Solute carrier family 2 (facilitated glucose/fructose transporter), member 5 | SLC2A5 | 1p36.2 | NM_003039.1 | 35.3 | Above | 5.1 | old |
| 26 | 205467_at | Caspase 10, apoptosis-related cysteine protease | CASP10 | 2q33-q34 | NM_001230.1 | 35.3 | Above | 3.6 | old |
| 27 | 225660_at | Semaphorin 6A | SEMA6A | 5q23.1 | W92748 | 35.3 | Above | 3.3 | new |
| 28 | 225913_at | FLJ21140 (Ser/Thr protein kinase) | FLJ21140 | 15 | AK025943.1 | 35.3 | Above | 2.9 | new |
| 29 | 236489_at | EST | | 6 | AI282097 | 35.3 | Above | 16.7 | new |
| 30 | 240173_at | EST | | 4 | AI732969 | 35.3 | Above | 10.3 | new |
| 31 | 240499_at | EST | | 10 | AA482221 | 35.3 | Above | 1.3 | new |
| 32 | 201310_s_at | P311 protein. Similar to gastrin/cholecystokinin type B receptor. | P311 | 5q21.3 | NM_004772.1 | 35.2 | Below | 2.2 | new |
| 33 | 215617_at | FLJ11754 | FLJ11754 | 2 | AU145711 | 35.2 | Above | 14.4 | new |
| 34 | 242579_at | EST | | 4 | AA935461 | 35.2 | Above | 10.2 | new |
| 35 | 202717_s_at | CDC16 cell division cycle 16 homolog | CDC16 | 13q34 | NM_003903.1 | 34.4 | Above | 1.1 | new |
| 36 | 205055_at | Integrin, alpha E (antigen CD103, human mucosal lymphocyte antigen 1) | ITGAE | 17p13 | NM_002208.3 | 34.4 | Below | 2.1 | new |
| 37 | 217967_s_at | Chromosome 1 ORF 24 | C1orf24 | 1q25 | AF288391.1 | 34.4 | Above | 3.2 | new |
| 38 | 201656_at | Integrin, alpha 6 | ITGA6 | 2q31.1 | NM_000210.1 | 33.9 | Above | 2.8 | new |
| 39 | 207196_s_at | Nef-associated factor 1 | NAF1 | 5q32-q33.1 | NM_006058.1 | 32.2 | Above | 1.4 | new |
| 40 | 219315_s_at | hypothetical protein FLJ23058 | FLJ20898 | 16p13.12 | NM_024600.1 | 32.2 | Above | 5.3 | new |
| 41 | 202123_s_at | V-abl Abelson murine leukemia viral oncogene | ABL1 | 9q34.1 | NM_005157.2 | 31.4 | Above | 1.8 | old |

| | | | | | | | | | |
|----|-------------|---|---------------|----------------|-------------|------|-------|------|-----|
| 42 | 219938_s_at | homolog 1 proline-serine-threonine phosphatase interacting protein 2 | PSTPIP2 | 18q12 | NM_024430.1 | 31.2 | Above | 5 | new |
| 43 | 228046_at | EST;DKFZp434P0235 | DKFZp434P0235 | 4 | AA741243 | 31.2 | Above | 1.1 | new |
| 44 | 64064_at | Immune associated nucleotide 4 like 1 | IAN4L1 | 7q36 | AI435089 | 30.9 | Above | 3.3 | new |
| 45 | 222729_at | F-box and WD-40 domain protein 7 (archipelago homolog, Drosophila) | FBXW7 | 4q31.23 | BE551877 | 30.5 | Above | 2.4 | new |
| 46 | 229975_at | EST | | 4 | AI826437 | 30.5 | Above | 9.1 | new |
| 47 | 200864_s_at | RAB11A, member RAS oncogene family | RAB11A | 15q21.3-q22.31 | NM_004663.1 | 29.7 | Above | 1.4 | old |
| 48 | 203089_s_at | Protease, serine, 25 | PRSS25 | 2p12 | NM_013247.1 | 29.7 | Above | 1.7 | new |
| 49 | 205376_at | Inositol polyphosphate-4-phosphatase, type II | INPP4B | 4q31.1 | NM_003866.1 | 29.7 | Above | 12.4 | new |
| 50 | 209229_s_at | KIAA1115 protein | KIAA1115 | 19q13.42 | BC002799.1 | 29.7 | Above | 1.3 | new |
| 51 | 219871_at | Hypothetical protein FLJ13197 | FLJ13197 | 4p14 | NM_024614.1 | 29.7 | Above | 14.5 | new |
| 52 | 222868_s_at | Interleukin 18 binding protein | IL18BP | 11q13 | AI521549 | 29.7 | Above | 7.1 | new |
| 53 | 235988_at | GPR110 G protein-coupled receptor 110 | GPR110 | 6p12.3 | AA746038 | 29.7 | Above | 15.8 | new |
| 54 | 239273_s_at | Matrix metalloproteinase 28 | MMP28 | 17q11-q21.1 | AI927208 | 29.7 | Above | 90.5 | new |
| 55 | 206150_at | Tumor necrosis factor receptor superfamily, member 7 | TNFRSF7 | 12p13 | NM_001242.1 | 29.5 | Above | 3.2 | old |
| 56 | 212203_x_at | Interferon induced transmembrane protein 3 (I-8U) | IFITM3 | 8q13.1 | BF338947 | 29.5 | Above | 2.3 | old |
| 57 | 217110_s_at | Mucin 4 | MUC4 | 3q29 | AJ242547.1 | 29.5 | Above | 47.5 | new |
| 58 | 223075_s_at | hypothetical protein FLJ12783 | FLJ12783 | 9q34.13-q34.3 | AL136566.1 | 29.5 | Above | 3.9 | new |
| 59 | 229139_at | EST | | 8 | AI202201 | 29.5 | Above | 10.8 | new |
| 60 | 229367_s_at | Hypothetical proteins FLJ22690. | FLJ22690 | 7 | AW130536 | 29.5 | Above | 3.6 | new |
| 61 | 213093_at | FLJ30869 | FLJ30869 | Xq28 | AI471375 | 29.1 | Above | 2.5 | new |
| 62 | 216033_s_at | FYN oncogene related to SRC | FYN | 6 | S74774.1 | 29.1 | Above | 2.7 | new |
| 63 | 202369_s_at | TRAM-like protein | KIAA0057 | 6p21.1-p12 | NM_012288.1 | 28.7 | Above | 3.3 | new |
| 64 | 212592_at | immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides | IGJ | 4q21 | AV733266 | 28.7 | Above | 7.9 | old |
| 65 | 219218_at | hypothetical protein FLJ23058 | FLJ23058 | 17q25.3 | NM_024696.1 | 28.7 | Below | 6.2 | new |
| 66 | 242051_at | EST | | Y | AI695695 | 28.7 | Above | 2.2 | new |
| 67 | 200655_s_at | Calmodulin 1 (phosphorylase kinase, delta) | CALM1 | 14q24-q31 | NM_006888.1 | 28.5 | Above | 1.3 | new |
| 68 | 202794_at | Inositol polyphosphate-1-phosphatase | INPP1 | 2q32 | NM_002194.2 | 28.4 | Above | 1.6 | new |
| 69 | 218348_s_at | HSPC055 protein | HSPC055 | 16p13.3 | NM_014153.1 | 27.7 | Below | 1.1 | new |
| 70 | 205269_at | Lymphocyte cytosolic protein 2 | LCP2 | 5q33.1-qter | AI123251 | 26.9 | Above | 1.6 | new |
| 71 | 238488_at | Ran binding protein 11 | LOC51194 | 5q12.2 | BF511602 | 26.9 | Above | 2.7 | new |
| 72 | 202242_at | Transmembrane 4 superfamily member 2 | TM4SF2 | Xq11.4 | NM_004615.1 | 26.6 | Above | 1.7 | new |
| 73 | 218764_at | Hypothetical protein MGC5363 | MGC5363 | 14q22.1-q22.3 | NM_024064.1 | 26.6 | Above | 1.7 | new |
| 74 | 224811_at | FLJ30652 | FLJ30652 | 3 | BF112093 | 26.6 | Above | 1.5 | new |

| | | | | | | | | | |
|-----|-------------|---|---------------|----------------|-------------|------|-------|------|-----|
| 75 | 225799_at | Hypothetical protein MGC4677 | MGC4677 | 2q12.3 | BF209337 | 26.6 | Above | 2.2 | new |
| 76 | 228297_at | Calponin 3, acidic | CNN3 | 1p22-p21 | AI807004 | 26.6 | Above | 4.7 | old |
| 77 | 203508_at | Tumor necrosis factor receptor superfamily, member 1B | TNFRSF1B | 1p36.3-p36.2 | NM_001066.1 | 26 | Above | 2.6 | old |
| 78 | 208071_s_at | Leukocyte-associated Ig-like receptor 1 | LAIR1 | 19q13.4 | NM_021708.1 | 26 | Above | 2 | old |
| 79 | 209321_s_at | Adenylate cyclase 3. | ADCY3 | 2p24-p22 | AF033861.1 | 26 | Above | 2.1 | old |
| 80 | 226345_at | DKFZp434O1317 | DKFZp434O1317 | 10 | AW270158 | 26 | Below | 1.4 | new |
| 81 | 200863_s_at | RAB11A, member RAS oncogene family | RAB11A | 15q21.3-q22.31 | AI215102 | 25.8 | Above | 1.4 | old |
| 82 | 205270_s_at | Lymphocyte cytosolic protein 2 | LCP2 | 5q33.1-qter | NM_005565.2 | 25.8 | Above | 1.6 | new |
| 83 | 208881_x_at | Isopenentenyl-diphosphate delta isomerase | IDI1 | 10p15.3 | BC005247.1 | 25.8 | Below | 1.7 | new |
| 84 | 212862_at | CDP-diacylglycerol synthase (phosphatidate cytidylyltransferase) 2 | CDS2 | 20p13 | AL568982 | 25.8 | Above | 1.8 | new |
| 85 | 213385_at | Chimerin 2 | CHN2 | 7 | AK026415.1 | 25.8 | Above | 3 | new |
| 86 | 218013_x_at | Dynactin 4 (p62) | DCTN4 | 5q31-q32 | NM_016221.1 | 25.8 | Above | 3.6 | new |
| 87 | 218966_at | Myosin 5C | MYO5C | 15q21 | NM_018728.1 | 25.8 | Above | 1.8 | new |
| 88 | 200742_s_at | Ceroid-lipofuscinosis, neuronal 2, late infantile (Jansky-Bielschowsky disease). A pepstatin-insensitive lysosomal peptidase. | CLN2 | 11p15 | BG231932 | 25 | Above | 1.5 | new |
| 89 | 203217_s_at | Sialyltransferase 9 | SIAT9 | 2p11.2 | NM_003896.1 | 25 | Above | 1.8 | new |
| 90 | 205259_at | Nuclear receptor subfamily 3, group C, member 2 | NR3C2 | 4q31.1 | NM_000901.1 | 25 | Above | 1.9 | new |
| 91 | 220684_at | T-box 21 | TBX21 | 17q21.2 | NM_013351.1 | 25 | Above | 3.3 | new |
| 92 | 225244_at | IMAGE3451454: GRASP protein | IMAGE3451454 | 1q42.13 | AA019893 | 25 | Above | 2 | new |
| 93 | 239519_at | EST | | 10 | AA927670 | 25 | Above | 18.2 | new |
| 94 | 203005_at | Lymphotoxin beta receptor (TNFR superfamily, member 3) | LTBR | 12p13 | NM_002342.1 | 24.3 | Above | 10 | new |
| 95 | 200665_s_at | Secreted protein, acidic, cysteine-rich (osteonectin) | SPARC | 5q31.3-q32 | NM_003118.1 | 24.3 | Above | 9.8 | new |
| 96 | 204004_at | PRKC, apoptosis, WT1, regulator | PAWR | 12q21 | AI336206 | 24.3 | Above | 3 | new |
| 97 | 204576_s_at | KIAA0643 protein | KIAA0643 | 16p12.3 | AA207013 | 24.3 | Above | 2 | new |
| 98 | 214255_at | ATPase, Class V, type 10C | ATP10C | 15q11-q13 | AB011138.1 | 24.3 | Above | 9.9 | new |
| 99 | 216985_s_at | Syntaxin 3A | STX3A | 11q12.3 | AJ002077.1 | 24.3 | Above | 12 | new |
| 100 | 48106_at | FLJ20489 | FLJ20489 | 12p11.1 | H14241 | 24.3 | Above | 2.8 | new |

Table S4. Top 100 chi-square probe sets selected for *E2A-PBX1* in decision tree format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Chi-square value | E2A – PBX1 above/below mean | Fold change | old or new |
|----|-----------------------|--|---------------|-----------------|--------------------------|-------------------------|------------------------------------|--------------------|-------------------|
| 1 | 201579_at | FAT tumor suppressor homolog 1 (Drosophila) | FAT | 4q34-q35 | NM_005245.1 | 88.0 | Above | 9.9 | old |
| 2 | 201695_s_at | nucleoside phosphorylase | NP | 14q13.1 | NM_000270.1 | 88.0 | Above | 3.8 | old |
| 3 | 204674_at | lymphoid-restricted membrane protein | LRMP | 12p12.3 | NM_006152.1 | 88.0 | Above | 5.8 | old |
| 4 | 205253_at | pre-B-cell leukemia transcription factor 1 | PBX1 | 1q23 | NM_002585.1 | 88.0 | Above | 3549.2 | old |
| 5 | 212148_at | pre-B-cell leukemia transcription factor 1, splice variant | PBX1 | 1q23 | BF967998 | 88.0 | Above | 5283.5 | old |
| 6 | 212151_at | pre-B-cell leukemia transcription factor 1, splice variant | PBX1 | 1q23 | BF967998 | 88.0 | Above | 7472.2 | old |
| 7 | 212371_at | DKFZp586C1019 | DKFZp586C1019 | 1 | AL049397.1 | 88.0 | Above | 2.5 | old |
| 8 | 219155_at | retinal degeneration B beta | RDGGB | 17q24.2 | NM_012417.1 | 88.0 | Above | 2.7 | new |
| 9 | 225483_at | hypothetical protein MGC10485 | MGC10485 | 11q25 | AI971602 | 88.0 | Above | 7.7 | new |
| 10 | 227439_at | E2a-Pbx1-associated protein | EB-1 | 12 | AW005572 | 88.0 | Above | 269.8 | new |
| 11 | 227949_at | Q9H4T4 like | H17739 | 20q13.32 | AL357503 | 88.0 | Above | 59.3 | new |
| 12 | 230306_at | hypothetical protein MGC10485 | MGC10485 | 11q25 | AA514326 | 88.0 | Above | 19.2 | new |
| 13 | 231095_at | retinal degeneration B beta | RDGGB | 17q24.2 | AW193811 | 88.0 | Above | 25.6 | new |
| 14 | 203372_s_at | STAT induced STAT inhibitor-2 | SOCS2 | 12q | AB004903.1 | 80.6 | Below | 23.4 | old |
| 15 | 206028_s_at | c-mer proto-oncogene tyrosine kinase | MERTK | 2q14.1 | NM_006343.1 | 80.6 | Above | 23.7 | old |
| 16 | 206181_at | signaling lymphocytic activation molecule | SLAM | 1q22-q23 | NM_003037.1 | 80.6 | Above | 6.3 | old |
| 17 | 208788_at | homolog of yeast long chain polyunsaturated fatty acid elongation enzyme 2 | HELO1 | 6p21.1-p12.1 | AL136939.1 | 80.6 | Above | 2.2 | old |
| 18 | 209760_at | KIAA0922 protein | KIAA0922 | 4q31.23 | AL136932.1 | 80.6 | Above | 2.9 | old |
| 19 | 35974_at | lymphoid-restricted membrane protein | LRMP | 12p12.3 | U10485 | 80.6 | Above | 6.2 | old |
| 20 | 38340_at | huntingtin interacting protein 12 | HIP12 | 12q24 | AB014555 | 80.6 | Above | 3.8 | old |
| 21 | 208644_at | ADP-ribosyltransferase (NAD ⁺ ; poly (ADP-ribose) polymerase) | ADPRT | 1q41-q42 | M32721.1 | 80.2 | Above | 3.0 | old |
| 22 | 212789_at | KIAA0056 protein | KIAA0056 | 11q25 | AI796581 | 80.2 | Above | 3.9 | old |
| 23 | 221113_s_at | wingless-type MMTV integration site family, member 16 | WNT16 | 7q31 | NM_016087.1 | 80.2 | Above | 2547.6 | new |
| 24 | 224022_x_at | wingless-type MMTV integration site family, member 16 | WNT16 | 7q31 | AF169963.1 | 80.2 | Above | 569.1 | new |
| 25 | 231040_at | EST | | 9 | AW512988 | 80.2 | Above | 16.4 | new |

| | | | | | | | | | |
|----|-------------|--|---------------|--------------|-------------|------|-------|-------|-----|
| 26 | 232289_at | FLJ14167 | FLJ14167 | 17 | BF237871 | 80.2 | Above | 144.1 | new |
| 27 | 235666_at | EST | FLJ20489 | 10 | AA903473 | 80.2 | Above | 654.6 | new |
| 28 | 203373_at | STAT induced STAT inhibitor-2 | SOCS2 | 12q | NM_003877.1 | 74.2 | Below | 24.8 | old |
| 29 | 210785_s_at | basement membrane-induced gene | ICB-1 | 1p35.3 | AB035482.1 | 74.2 | Below | 4.1 | old |
| 30 | 224733_at | chemokine-like factor super family 3 | CKLFSF3 | 16q23.1 | AL574900 | 74.2 | Below | 41.7 | new |
| 31 | 225235_at | hypothetical protein MGC14859 | MGC14859 | 5q35.3 | AW007710 | 74.2 | Above | 3.6 | new |
| 32 | 204114_at | nidogen 2 (osteonidogen) | NID2 | 14q21-q22 | NM_007361.1 | 73.1 | Above | 15.1 | old |
| 33 | 211913_s_at | c-mer proto-oncogene tyrosine kinase | MERTK | 2q14.1 | L08961.1 | 72.8 | Above | 37.7 | old |
| 34 | 219551_at | uncharacterized bone marrow protein BM040 | BM040 | 3q21.1 | NM_018456.1 | 72.8 | Above | 3.0 | New |
| 35 | 223693_s_at | hypothetical protein FLJ10324 | FLJ10324 | 7p22 | AL136731.1 | 72.8 | Above | 65.6 | New |
| 36 | 200600_at | moesin | MSN | Xq11.2-q12 | NM_002444.1 | 72.5 | Below | 2.2 | Old |
| 37 | 213909_at | FLJ12280 | FLJ12280 | 3 | AU147799 | 72.5 | Above | 12.5 | New |
| 38 | 221669_s_at | acyl-Coenzyme A dehydrogenase family, member 8 | ACAD8 | 11q25 | BC001964.1 | 72.5 | Above | 2.6 | New |
| 39 | 235911_at | ESTs, Weakly similar to PIHUB6 salivary proline-rich protein precursor PRB1 (large allele) | | 3 | AI885815 | 72.5 | Above | 36.6 | New |
| 40 | 243533_x_at | ESTs | | | H09663 | 72.5 | Above | 23.2 | New |
| 41 | 202615_at | DKFZp686D0521 | DKFZp686D0521 | 9 | BF222895 | 68.6 | Below | 6.2 | Old |
| 42 | 204774_at | ecotropic viral integration site 2A | EVI2A | 17q11.2 | NM_014210.1 | 68.6 | Below | 3.0 | New |
| 43 | 218283_at | synovial sarcoma translocation gene on chromosome 18-like 2 | SS18L2 | 3p21 | NM_016305.1 | 68.6 | Above | 1.6 | New |
| 44 | 209130_at | synaptosomal-associated protein, 23kDa | SNAP23 | 15q14 | BC003686.1 | 67.8 | Below | 1.9 | New |
| 45 | 228580_at | serine protease HTRA3 | HTRA3 | 4p16.1 | AI828007 | 66.6 | Above | 3.8 | New |
| 46 | 202796_at | synaptopodin | KIAA1029 | 5q33.1 | NM_007286.1 | 66.5 | Above | 52.3 | Old |
| 47 | 218640_s_at | phafin 2 | FLJ13187 | 8q21.3 | NM_024613.1 | 66.5 | Above | 3.1 | New |
| 48 | 235099_at | ESTs, Weakly similar to PLLP_HUMAN Plasmolipin [H.sapiens] | | 3 | AW080832 | 66.5 | Above | 6.7 | New |
| 49 | 201889_at | family with sequence similarity 3, member C | FAM3C | 7q22.1-q31.1 | NM_014888.1 | 65.3 | Above | 4.6 | New |
| 50 | 202106_at | golgi autoantigen, golgin subfamily a, 3 | GOLGA3 | 12q24.33 | NM_005895.1 | 65.3 | Above | 3.3 | Old |
| 51 | 202208_s_at | ADP-ribosylation factor-like 7 | ARL7 | 2q37.2 | BC001051.1 | 65.3 | Above | 3.2 | Old |
| 52 | 205173_x_at | CD58 antigen, (lymphocyte function-associated antigen 3) | CD58 | 1p13 | NM_001779.1 | 65.3 | Above | 2.4 | Old |
| 53 | 211744_s_at | CD58 antigen, (lymphocyte function-associated antigen 3) | CD58 | 1p13 | BC005930.1 | 65.3 | Above | 2.5 | Old |
| 54 | 212552_at | hippocalcin-like 1 | HPCAL1 | 2p25.1 | BE617588 | 65.3 | Below | 2.6 | Old |
| 55 | 213358_at | KIAA0802 protein | KIAA0802 | 18p11.21 | AB018345.1 | 65.3 | Above | 12.7 | Old |
| 56 | 222699_s_at | phafin 2 | FLJ13187 | 8q21.3 | BF439250 | 65.3 | Above | 3.5 | New |
| 57 | 225618_at | EST | | 17 | AI769587 | 65.3 | Below | 5.3 | New |
| 58 | 238778_at | DKFZp451L157 | DKFZp451 | 10 | AI244661 | 65.3 | Above | 23.5 | New |

| | | | L157 | | | | | | |
|----|-------------|--|----------------|---------------|-------------|------|-------|--------|-----|
| 59 | 239427_at | ESTs | | 1 | AA131524 | 65.3 | Above | 13.7 | New |
| 60 | 47069_at | Rho GTPase activating protein 8 | ARHGAP8 | 22q13.31 | AA533284 | 65.3 | Above | 3.3 | New |
| 61 | 205769_at | solute carrier family 27 (fatty acid transporter), member 2 | SLC27A2 | 15q21.2 | NM_003645.1 | 65.1 | Above | 56.0 | Old |
| 62 | 210786_s_at | Friend leukemia virus integration 1 | FLI1 | 11q24.1-q24.3 | M93255.1 | 65.1 | Above | 2.2 | Old |
| 63 | 212985_at | DKFZp434E033 | DKFZp434E033 | 4 | BF115739 | 65.1 | Above | 7.1 | New |
| 64 | 227441_s_at | E2a-Pbx1-associated protein | EB-1 | 12 | AW005572 | 65.1 | Above | 1139.4 | New |
| 65 | 234261_at | DKFZp761M10121 | DKFZp761M10121 | 12 | AL137313.1 | 65.1 | Above | 960.8 | New |
| 66 | 244565_at | ESTs | | 10 | AI685824 | 65.1 | Above | 7.6 | New |
| 67 | 202181_at | KIAA0247 gene product | KIAA0247 | 14q24.1 | NM_014734.1 | 63.7 | Above | 1.8 | Old |
| 68 | 202207_at | ADP-ribosylation factor-like 7 | ARL7 | 2q37.2 | NM_005737.2 | 63.7 | Above | 3.2 | Old |
| 69 | 207571_x_at | basement membrane-induced gene | ICB-1 | 1p35.3 | NM_004848.1 | 63.7 | Below | 4.4 | Old |
| 70 | 209558_s_at | huntingtin interacting protein 12 | HIP12 | 12q24 | AB013384.1 | 61.1 | Above | 23.8 | Old |
| 71 | 213005_s_at | KIAA0172 protein | KIAA0172 | 9p24.3 | D79994.1 | 61.1 | Above | 8.3 | Old |
| 72 | 236854_at | cDNA DKFZp667F0617 | DKFZp667F0617 | 20 | AA743694 | 61.1 | Above | 12.6 | New |
| 73 | 226233_at | tubulin-specific chaperone e | TBCE | 1q42.3 | BG112197 | 60.0 | Above | 2.6 | New |
| 74 | 203435_s_at | membrane metallo-endopeptidase (neutral endopeptidase, enkephalinase, CALLA, CD10) | MME | 3q25.1-q25.2 | NM_007287.1 | 59.9 | Below | 2.2 | Old |
| 75 | 202478_at | GS3955 protein | GS3955 | 2p25.1 | NM_021643.1 | 59.3 | Above | 4.0 | Old |
| 76 | 202479_s_at | GS3955 protein | GS3955 | 2p25.1 | BC002637.1 | 59.3 | Above | 3.3 | Old |
| 77 | 203999_at | synaptotagmin I | SYT1 | 12cen-q21 | NM_005639.1 | 59.3 | Above | 3.9 | Old |
| 78 | 212149_at | KIAA0143 protein | KIAA0143 | 8q24.12 | AA805651 | 59.3 | Below | 13.5 | New |
| 79 | 212873_at | minor histocompatibility antigen HA-1 | HA-1 | 19p13.3 | BE349017 | 59.3 | Below | 2.9 | Old |
| 80 | 218346_s_at | p53 regulated PA26 nuclear protein | PA26 | 6q21 | NM_014454.1 | 59.3 | Below | 4.7 | New |
| 81 | 224856_at | FK506 binding protein 5 | FKBP5 | 6p21.3-21.2 | AL122066.1 | 59.3 | Below | 5.5 | Old |
| 82 | 200811_at | cold inducible RNA binding protein | CIRBP | 19p13.3 | NM_001280.1 | 59.1 | Below | 5.8 | Old |
| 83 | 201722_s_at | UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 1 (GalNAc-T1) | GALNT1 | 18q12.1 | NM_020474.2 | 59.1 | Below | 1.8 | New |
| 84 | 223711_s_at | HSPC144 protein | HSPC144 | 11q25 | AF182413.1 | 59.1 | Above | 2.0 | New |
| 85 | 233273_at | cDNA FLJ12010 fis | FLJ12010 | 1 | AU146834 | 59.1 | Above | 30.6 | New |
| 86 | 201460_at | mitogen-activated protein kinase-activated protein kinase 2 | MAPKAPK2 | 1q32 | AI141802 | 57.9 | Above | 2.1 | Old |
| 87 | 202421_at | immunoglobulin superfamily, member 3 | IGSF3 | 1p13 | AB007935.1 | 57.9 | Above | 4.4 | New |
| 88 | 217983_s_at | ribonuclease 6 precursor | RNASE6PL | 6q27 | NM_003730.2 | 57.9 | Below | 3.4 | New |
| 89 | 218087_s_at | sorbin and SH3 domain containing 1 | SORBS1 | 10q23.3-q24.1 | NM_015385.1 | 57.9 | Above | 25.1 | New |

| | | | | | | | | | |
|-----|-------------|---|---------------|---------------|-------------|------|-------|--------|-----|
| 90 | 218491_s_at | HSPC144 protein | HSPC144 | 11q25 | NM_014174.1 | 57.9 | Above | 1.4 | New |
| 91 | 201825_s_at | CGI-49 protein | LOC51097 | 1q44 | AL572542 | 57.8 | Above | 2.2 | Old |
| 92 | 202206_at | ADP-ribosylation factor-like 7 | ARL7 | 2q37.2 | NM_005737.2 | 57.8 | Above | 3.9 | Old |
| 93 | 218683_at | polypyrimidine tract binding protein 2 | PTBP2 | 1p22.11-p21.3 | NM_021190.1 | 57.8 | Above | 1.8 | New |
| 94 | 226590_at | cDNA clone EUROIMAGE 1517766 | | 9 | AA031404 | 57.8 | Above | 3.1 | New |
| 95 | 227440_at | E2a-Pbx1-associated protein | EB-1 | 12 | AW005572 | 57.8 | Above | 1168.9 | New |
| 96 | 229770_at | hypothetical protein FLJ31978 | FLJ31978 | 12q24.33 | AI041543 | 57.8 | Above | 51.8 | New |
| 97 | 40148_at | amyloid beta (A4) precursor protein-binding, family B, member 2 (Fe65-like) | APBB2 | 4p14 | U62325 | 57.8 | Above | 6.2 | Old |
| 98 | 212959_s_at | MGC4170 protein | MGC4170 | 12q23.1 | AK001821.1 | 57.2 | Below | 3.0 | New |
| 99 | 203143_s_at | KIAA0040 gene product | KIAA0040 | 1q24-25 | T79953 | 56.3 | Above | 2.4 | New |
| 100 | 209683_at | hypothetical protein DKFZp566A1524 | DKFZP566A1524 | 2p24.2 | AA243659 | 56.3 | Below | 10.0 | New |

Table S5. Top 100 chi-square probe sets selected for Hyperdiploid >50 in decision tree format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Chi-square value | HD>50 above/below mean | Fold change | old or new |
|----|-----------------------|---|---------------|-----------------|--------------------------|-------------------------|----------------------------------|--------------------|-------------------|
| 1 | 200600_at | Moesin (membrane-organizing extensio spike protein) | MSN | Xq11.2-q12 | NM_002444.1 | 34.0 | Above | 1.9 | old |
| 2 | 200737_at | Phosphoglycerate kinase 1 | PGK1 | Xq13 | NM_000291.1 | 34.0 | Above | 1.8 | old |
| 3 | 200980_s_at | Pyruvate dehydrogenase (lipoamide) alpha 1 | PDHA1 | Xp22.2-p22.1 | NM_000284.1 | 34.0 | Above | 1.7 | new |
| 4 | 201136_at | Proteolipid protein 2 (colonic epithelium-enriched) | PLP2 | Xp11.23 | NM_002668.1 | 34.0 | Above | 3.3 | old |
| 5 | 201807_at | Vacuolar protein sorting 26 (yeast) | VPS26 | 10q21.1 | NM_004896.1 | 34.0 | Above | 1.7 | old |
| 6 | 202214_s_at | Cullin 4B | CUL4B | Xq23 | NM_003588.1 | 34.0 | Above | 1.9 | new |
| 7 | 202557_at | Stress 70 protein chaperone, microsome associated, 60 kD | STCH | 21q11 | AI718418 | 34.0 | Above | 2.0 | new |
| 8 | 202593_s_at | membrane interacting protein of RGS16 | MIR16 | 16p12-p11.2 | NM_016641.1 | 34.0 | Below | 1.6 | new |
| 9 | 203680_at | Protein kinase, cAMP-dependent, regulatory, type II, beta | PRKAR2B | 7q22-q31.1 | NM_002736.1 | 34.0 | Above | 3.3 | new |
| 10 | 204194_at | BTB and CNC homology 1, basic leucine zipper transcription factor 1 | BACH1 | 21q22.11 | NM_001186.1 | 34.0 | Above | 1.8 | new |
| 11 | 205324_s_at | FtsJ homolog 1 (E. coli) | FTSJ1 | Xp11.23 | NM_012280.1 | 34.0 | Above | 2.1 | old |
| 12 | 208598_s_at | Upstream regulatory element binding protein 1 | UREB1 | Xp11.22 | NM_005703.2 | 34.0 | Above | 1.6 | old |
| 13 | 208861_s_at | Alpha thalassemia/mental retardation syndrome X-linked (RAD54 homolog, S. cerevisiae) | ATRX | Xq13.1-q21.1 | U72937.2 | 34.0 | Above | 1.7 | old |

| | | | | | | | | | |
|----|-------------|--|----------|----------------|-------------|------|-------|-----|-----|
| 14 | 211342_x_at | trinucleotide repeat containing 11 (THR-associated protein, 230 kDa subunit) | TNRC11 | Xq13 | BC004354.1 | 34.0 | Above | 1.8 | new |
| 15 | 216071_x_at | Trinucleotide repeat containing 11 | TNRC11 | Xq13 | AF132033 | 34.0 | Above | 1.8 | new |
| 16 | 218573_at | APR-1 protein/melanoma-associated antigen | MAGEH1 | Xp11.22 | NM_014061.1 | 34.0 | Above | 3.0 | new |
| 17 | 219485_s_at | proteasome (prosome, macropain) 26S subunit, non-ATPase, 10 | PSMD10 | Xq22.3 | NM_002814.1 | 34.0 | Above | 2.4 | old |
| 18 | 200655_s_at | Calmodulin 1 (phosphorylase kinase, delta) | CALM1 | 14q24-q31 | NM_006888.1 | 30.1 | Above | 1.7 | old |
| 19 | 200738_s_at | Phosphoglycerate kinase 1 | PGK1 | Xq13 | NM_000291.1 | 30.1 | Above | 1.8 | old |
| 20 | 200944_s_at | High-mobility group (nonhistone chromosomal) protein 14; member of the HMG 14/17 family | HMG14 | 21q22.2 | NM_004965.1 | 30.1 | Above | 1.7 | old |
| 21 | 201092_at | Retinoblastoma binding protein 7 | RBBP7 | Xp22.31 | NM_002893.2 | 30.1 | Above | 1.6 | new |
| 22 | 201100_s_at | Ubiquitin specific protease 9 | USP9X | Xp11.4 | NM_004652.2 | 30.1 | Above | 1.7 | old |
| 23 | 201688_s_at | Tumor protein D52 | TPD52 | 8q21 | BE974098 | 30.1 | Below | 4.1 | new |
| 24 | 201899_s_at | Ubiquitin-conjugating enzyme E2A (RAD6 homolog) | UBE2A | Xq24-q25 | NM_003336.1 | 30.1 | Above | 1.8 | old |
| 25 | 202325_s_at | ATP synthase, H ⁺ transporting, mitochondrial F0 complex, subunit F6 | ATP5J | 21q21.1 | NM_001685.1 | 30.1 | Above | 1.6 | old |
| 26 | 202829_s_at | Synaptobrevin-like 1 | SYBL1 | Xq28 | NM_005638.1 | 30.1 | Above | 1.5 | old |
| 27 | 202854_at | Hypoxanthine phosphoribosyltransferase 1 (Lesch-Nyhan syndrome) | HPRT1 | Xq26.1 | NM_000194.1 | 30.1 | Above | 1.4 | old |
| 28 | 206846_s_at | Histone deacetylase 6 | HDAC6 | Xp11.23 | NM_006044.2 | 30.1 | Above | 1.5 | new |
| 29 | 209370_s_at | SH3-domain binding protein 2 | SH3BP2 | 4p16.3 | AB000462.1 | 30.1 | Above | 3.1 | new |
| 30 | 209565_at | zinc finger protein 183 (RING finger, C3HC4 type) | ZNF183 | Xq25-q26 | BC000832.1 | 30.1 | Above | 2.2 | new |
| 31 | 212846_at | KIAA0179 protein. | KIAA0179 | 21q22.3 | D80001.1 | 30.1 | Above | 2.0 | old |
| 32 | 217356_s_at | Phosphoglycerate kinase | PGK1 | Xq13 | S81916.1 | 30.1 | Above | 1.8 | new |
| 33 | 218163_at | MCT-1 protein | MCT-1 | Xq22-24 | NM_014060.1 | 30.1 | Above | 1.8 | new |
| 34 | 218386_x_at | Ubiquitin specific protease 16; de-ubiquitinates histone H2A; ubiquitous expression. | USP16 | 21q22.11 | NM_006447.1 | 30.1 | Above | 1.7 | new |
| 35 | 218402_s_at | Hermansky-Pudlak syndrome 4 | HPS4 | | NM_022081.1 | 30.1 | Below | 3.4 | new |
| 36 | 218495_at | Ubiquitously-expressed transcript | UXT | Xp11.23-p11.22 | NM_004182.1 | 30.1 | Above | 1.5 | new |
| 37 | 218499_at | Mst3 and SOK1-related kinase/STE20-like kinase; contains a Ser/Thr protein kinase domain | MST4 | Xq26.1 | NM_016542.1 | 30.1 | Above | 2.5 | new |
| 38 | 218757_s_at | Similar to yeast Upf3, variant B | UPF3B | Xq25-q26 | NM_023010.1 | 30.1 | Above | 2.3 | new |
| 39 | 219038_at | Hypothetical protein FLJ11565 | FLJ11565 | Xq22.2 | NM_024657.1 | 30.1 | Above | 6.9 | new |
| 40 | 229967_at | Chemokine-like factor super family 2. | CKLFSF2 | 16q23.1 | AA778552 | 30.1 | Above | 4.3 | new |
| 41 | 242794_at | EST | | 4q31.1 | AI569476 | 30.1 | Above | 3.2 | new |
| 42 | 201132_at | Heterogeneous nuclear ribonucleoprotein H2 (H') | HNRPH2 | Xq22 | NM_019597.1 | 30.0 | Above | 2.0 | old |
| 43 | 201312_s_at | SH3 domain binding glutamic acid-rich protein like | SH3BGRL | Xq13.3 | NM_003022.1 | 30.0 | Above | 1.6 | old |

| | | | | | | | | | |
|----|-------------|--|----------|---------------|-------------|------|-------|------|-----|
| 44 | 201894_s_at | Decorin; glycoprotein that binds to type I collagen fibrils & plays a role in matrix assembly. | DCN | 12q13.2 | NM_001920.1 | 30.0 | Above | 1.5 | new |
| 45 | 201923_at | Peroxiredoxin 4 | PRDX4 | Xp22.13 | NM_006406.1 | 30.0 | Above | 1.9 | old |
| 46 | 202371_at | Hypothetical protein FLJ21174 | FLJ21174 | Xq22.1 | NM_024863.1 | 30.0 | Above | 3.6 | old |
| 47 | 203126_at | Inositol(myo)-1(or 4)-monophosphatase 2 | IMPA2 | 18p11.2 | NM_014214.1 | 30.0 | Above | 4.1 | old |
| 48 | 204219_s_at | proteasome (prosome, macropain) 26S subunit, ATPase, 1 | PSMC1 | 19p13.3 | NM_002802.1 | 30.0 | Above | 1.3 | old |
| 49 | 204835_at | polymerase (DNA directed), alpha | POLA | Xp22.1-p21.3 | NM_016937.1 | 30.0 | Above | 2.0 | old |
| 50 | 212071_s_at | Spectrin, beta, non-erythrocytic 1 | SPTBN1 | 2p21 | BE968833 | 30.0 | Below | 1.7 | new |
| 51 | 212419_at | EST | | 10q22.3 | AL049949.1 | 30.0 | Above | 13.1 | new |
| 52 | 212718_at | Hypothetical protein MGC5370 | MGC5378 | 14q32.2 | BG110231 | 30.0 | Above | 1.5 | new |
| 53 | 213502_x_at | Homo sapiens cDNA FLJ32313 fis, clone PROST2003232, weakly similar to BETA-GLUCURONIDASE PRECURSOR (EC 3.2.1.31) | FLJ32313 | 22q11.23 | X03529 | 30.0 | Below | 1.8 | old |
| 54 | 214051_at | Thymosin, beta | TMSNB | Xq21.33-q22.3 | BF677486 | 30.0 | Above | 3.1 | new |
| 55 | 226039_at | Mannosyl (alpha-1,3)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase | MGAT4A | 2q11.2 | AW006441 | 30.0 | Above | 3.0 | new |
| 56 | 227279_at | hypothetical protein MGC15737 | MGC15737 | Xq22.1 | AA847654 | 30.0 | Above | 5.6 | new |
| 57 | 200642_at | Superoxide dismutase 1, soluble | SOD1 | 21q22.11 | NM_000454.1 | 26.7 | Above | 2.3 | old |
| 58 | 200799_at | Heat shock 70kD protein 1A | HSPA1A | 6p21.3 | NM_005345.3 | 26.7 | Above | 2.7 | old |
| 59 | 200943_at | High-mobility group (nonhistone chromosomal) protein 14; member of the HMG 14/17 family | HMG14 | 21q22.2 | NM_004965.1 | 26.7 | Above | 1.6 | old |
| 60 | 201018_at | Eukaryotic translation initiation factor 1A | EIF1A | Xp22.12 | BE542684 | 26.7 | Above | 1.8 | new |
| 61 | 201311_s_at | SH3 domain binding glutamic acid-rich protein like | SH3BGRL | Xq13.3 | AL515318 | 26.7 | Above | 1.6 | old |
| 62 | 201443_s_at | ATPase, H+ transporting, lysosomal interacting protein 2 | ATP6IP2 | Xq21 | AF248966.1 | 26.7 | Above | 1.9 | old |
| 63 | 201472_at | Von Hippel-Lindau binding protein 1 | VBPI | Xq28 | NM_003372.2 | 26.7 | Above | 1.7 | old |
| 64 | 201689_s_at | Tumor protein D52 | TPD52 | 8q21 | BE974098 | 26.7 | Below | 4.3 | new |
| 65 | 202602_s_at | HIV TAT specific factor 1 | HTATSF1 | Xq26.1-q27.2 | NM_014500.1 | 26.7 | Above | 1.5 | old |
| 66 | 203041_s_at | Lysosomal-associated membrane protein 2 | LAMP2 | Xq24 | J04183.1 | 26.7 | Above | 3.1 | old |
| 67 | 203102_s_at | Mannosyl (alpha-1,6-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase | MGAT2 | 14q21 | NM_002408.2 | 26.7 | Above | 1.6 | new |
| 68 | 203744_at | High-mobility group (nonhistone chromosomal) protein 4 | HMG4 | Xq28 | NM_005342.1 | 26.7 | Above | 1.9 | new |
| 69 | 205518_s_at | Cytidine monophosphate-N-acetylneuraminic acid hydroxylase (CMP-N-acetylneuraminic acid monooxygenase) | CMAH | 6p22-p23 | NM_003570.1 | 26.7 | Below | 2.9 | new |
| 70 | 208683_at | Calpain 2, (m/II) large subunit; calcium- | CAPN2 | 1q41-q42 | M23254.1 | 26.7 | Above | 2.2 | old |

| | | | | | | | | | |
|-----|-------------|---|----------|---------------|-------------|------|-------|------|-----|
| 71 | 209440_at | dependent Cys protease. Phosphoribosyl pyrophosphate synthetase 1; purine biosynthesis. | PRPS1 | Xq21-q27 | BC001605.1 | 26.7 | Above | 1.4 | old |
| 72 | 210786_s_at | Friend leukemia virus integration 1 | FLI1 | 11q24.1-q24.3 | M93255.1 | 26.7 | Below | 2.5 | old |
| 73 | 212070_at | G protein-coupled receptor 56 | GPR56 | 16q13 | AL554008 | 26.7 | Above | 2.4 | new |
| 74 | 213334_x_at | Three prime repair exonuclease 2 | TREX2 | Xq28 | BE676218 | 26.7 | Above | 1.7 | old |
| 75 | 215117_at | Recombination activating gene 2; V(D)J recombinase. | RAG2 | 11p13 | AW058148 | 26.7 | Below | 27.2 | old |
| 76 | 218694_at | ALEX1 protein | ALEX1 | Xq21.33-q22.2 | NM_016608.1 | 26.7 | Above | 2.8 | new |
| 77 | 222741_s_at | hypothetical protein FLJ11101 | FLJ11101 | 6p21.1 | AI761426 | 26.7 | Above | 1.5 | new |
| 78 | 223082_at | SH3-domain kinase binding protein 1 | SH3KBP1 | Xp22.1-p21.3 | AF230904.1 | 26.7 | Above | 2.0 | new |
| 79 | 225105_at | clone MGC:23936 IMAGE:3838595, mRNA, complete cds | | 12q23.3 | BF969397 | 26.7 | Above | 2.1 | new |
| 80 | 225406_at | Twisted gastrulation | TSG | 18p11.3 | AA195009 | 26.7 | Above | 1.9 | new |
| 81 | 225553_at | Homo sapiens cDNA FLJ12874 fis | | 14q22.2 | AL042817 | 26.7 | Above | 1.6 | new |
| 82 | 226199_at | Hypothetical protein MGC23937 | MGC23937 | Xq13.1 | AL563795 | 26.7 | Above | 2.1 | new |
| 83 | 226875_at | Hypothetical protein FLJ32122 | FLJ32122 | Xq24 | AI742838 | 26.7 | Above | 2.3 | new |
| 84 | 232974_at | cDNA FLJ12417 fis | | Xp22.31 | AU148256 | 26.7 | Above | 3.1 | new |
| 85 | 46323_at | Ca ²⁺ -dependent endoplasmic reticulum nucleoside diphosphatase | SHAPY | 17q25.3 | AL120741 | 26.7 | Above | 1.7 | new |
| 86 | 203694_s_at | DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 16 | DDX16 | 6p21.3 | NM_003587.2 | 26.3 | Above | 1.3 | old |
| 87 | 200658_s_at | Prohibitin | PHB | 17q21 | AL560017 | 26.3 | Above | 2.0 | old |
| 88 | 201898_s_at | ubiquitin-conjugating enzyme E2A (RAD6 homolog) | UBE2A | Xq24-q25 | AI126625 | 26.3 | Above | 1.6 | old |
| 89 | 203556_at | KIAA0854 protein | KIAA0854 | 8q24.13 | NM_014943.1 | 26.3 | Below | 1.6 | old |
| 90 | 203745_at | Holocytochrome c synthase (cytochrome c heme- lyase) | HCCS | Xp22.3 | AI801013 | 26.3 | Above | 2.1 | new |
| 91 | 203909_at | Solute carrier family 9 (sodium/hydrogen exchanger), isoform 6 | SLC9A6 | Xq26.3 | NM_006359.1 | 26.3 | Above | 1.9 | old |
| 92 | 204446_s_at | Arachidonate 5-lipoxygenase | ALOX5 | 10q11.2 | NM_000698.1 | 26.3 | Above | 4.2 | old |
| 93 | 205191_at | Retinitis pigmentosa 2 (X-linked recessive) | RP2 | Xp11.4-p11.21 | NM_006915.1 | 26.3 | Above | 2.1 | old |
| 94 | 206874_s_at | Ste20-related serine/threonine kinase | SLK | 10q25.1 | AL138761 | 26.3 | Above | 1.6 | new |
| 95 | 208073_x_at | Tetratricopeptide repeat domain 3 | TTC3 | 21q22.2 | NM_003316.1 | 26.3 | Above | 1.9 | old |
| 96 | 209056_s_at | CDC5 cell division cycle 5-like (<i>S. pombe</i>) | CDC5L | 6p21 | AW268817 | 26.3 | Above | 1.4 | new |
| 97 | 210645_s_at | Tetratricopeptide repeat domain 3 | TTC3 | 21q22.2 | D83077.1 | 26.3 | Above | 2.2 | old |
| 98 | 215773_x_at | ADP-ribosyltransferase (NAD ⁺ ; poly(ADP- ribose) polymerase)-like 2 | ADPRTL2 | 14q11.2-q12 | AJ236912.1 | 26.3 | Above | 1.6 | old |
| 99 | 215884_s_at | Ubiquilin 2 | UBQLN2 | Xp11.23-p11.1 | AK001029.1 | 26.3 | Above | 1.9 | new |
| 100 | 217954_s_at | PHD finger protein 3 | PHF3 | 6 | NM_015153.1 | 26.3 | Above | 1.5 | new |

Table S6. Top 100 chi-square probe sets selected for *MLL* in decision tree format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Chi-square value | MLL above/below mean | Fold change | old or new |
|----|-----------------------|---|---------------|-----------------|--------------------------|-------------------------|-----------------------------|--------------------|-------------------|
| 1 | 202603_at | a disintegrin and metalloproteinase domain 10 | ADAM10 | 15q22 | N51370 | 44.6 | Above | 1.8 | old |
| 2 | 219463_at | chromosome 20 open reading frame 103 | C20orf103 | 20p12 | NM_012261.1 | 44.6 | Above | 24.7 | new |
| 3 | 224772_at | neuron navigator 1 | NAV1 | | AB032977.1 | 44.6 | Below | 3.8 | new |
| 4 | 204069_at | Meis1, myeloid ecotropic viral integration site 1 homolog | MEIS1 | 2p14-p13 | NM_002398.1 | 44.4 | Above | 73.7 | old |
| 5 | 218966_at | myosin 5C | MYO5C | 15q21 | NM_018728.1 | 44.4 | Below | 4.5 | new |
| 6 | 226939_at | cDNA FLJ37247 fis | FLJ37247 | | AI202327 | 44.4 | Above | 6.9 | new |
| 7 | 204446_s_at | arachidonate 5-lipoxygenase | ALOX5 | 10q11.2 | NM_000698.1 | 40.7 | Below | 66.8 | old |
| 8 | 206492_at | fragile histidine triad gene | FHIT | 3p14.2 | NM_002012.1 | 40.7 | Below | 36.6 | old |
| 9 | 212588_at | protein tyrosine phosphatase, receptor type, C | PTPRC | 1q31-q32 | AI809341 | 40.7 | Above | 2.3 | old |
| 10 | 215925_s_at | CD72 antigen | CD72 | 9p11.2 | AF283777.2 | 40.7 | Above | 3.0 | old |
| 11 | 211733_x_at | sterol carrier protein 2 | SCP2 | 1p32 | BC005911.1 | 40.1 | Above | 1.5 | old |
| 12 | 212386_at | cDNA FLJ11918 fis | FLJ11918 | | AK021980.1 | 40.1 | Below | 3.1 | new |
| 13 | 218764_at | Protein Kinase C eta isoform. | PRKCH | 14q22.1-q22.3 | NM_024064.1 | 40.1 | Below | 7.6 | new |
| 14 | 218847_at | IGF-II mRNA-binding protein 2 | IMP-2 | 3q28 | NM_006548.1 | 40.1 | Above | 23.2 | new |
| 15 | 222409_at | coronin, actin binding protein, 1C | CORO1C | 12q24.1 | AL162070.1 | 40.1 | Above | 4.8 | new |
| 16 | 242172_at | ESTs | | | N50406 | 40.1 | Above | 33.6 | new |
| 17 | 201153_s_at | muscleblind-like (Drosophila) | MBNL | 3q25 | NM_021038.1 | 40.0 | Above | 2.1 | old |
| 18 | 210487_at | deoxynucleotidyltransferase, terminal | DNTT | 10q23-q24 | M11722.1 | 40.0 | Below | 2.9 | old |
| 19 | 219686_at | gene for serine/threonine protein kinase | HSA250839 | 4p16.2 | NM_018401.1 | 40.0 | Below | 28.3 | new |
| 20 | 226981_at | Homo sapiens, clone IMAGE:4401491, mRNA | | | AW002079 | 37.4 | Below | 1.0 | new |
| 21 | 203375_s_at | tripeptidyl peptidase II | TPP2 | 13q32-q33 | NM_003291.1 | 37.2 | Above | 1.6 | new |
| 22 | 221676_s_at | coronin, actin binding protein, 1C | CORO1C | 12q24.1 | BC002342.1 | 37.2 | Above | 3.5 | new |
| 23 | 201152_s_at | muscleblind-like (Drosophila) | MBNL | 3q25 | NM_021038.1 | 36.2 | Above | 2.2 | old |
| 24 | 221773_at | ELK3, ETS-domain protein (SRF accessory protein 2) | ELK3 | 12q23 | AW575374 | 36.2 | Below | 8.2 | new |
| 25 | 201162_at | insulin-like growth factor binding protein 7 | IGFBP7 | 4q12 | NM_001553.1 | 36.0 | Above | 4.3 | old |
| 26 | 201163_s_at | insulin-like growth factor binding protein 7 | IGFBP7 | 4q12 | NM_001553.1 | 36.0 | Above | 4.0 | old |
| 27 | 203836_s_at | mitogen-activated protein kinase kinase kinase 5 | MAP3K5 | 6q22.33 | D84476.1 | 36.0 | Above | 13.9 | old |
| 28 | 203837_at | mitogen-activated protein kinase kinase kinase 5 | MAP3K5 | 6q22.33 | NM_005923.2 | 36.0 | Above | 4.2 | old |
| 29 | 213891_s_at | cDNA FLJ11918 fis | FLJ11918 | | AI927067 | 36.0 | Below | 3.2 | new |
| 30 | 214895_s_at | a disintegrin and metalloproteinase domain 10 | ADAM10 | 15q22 | AU135154 | 36.0 | Above | 1.9 | old |

| | | | | | | | | | |
|----|-------------|---|-------------------|--------------|-------------|------|-------|------|-----|
| 31 | 226415_at | KIAA1576 protein | KIAA1576 | 16q22.1 | AA156723 | 36.0 | Above | 40.7 | new |
| 32 | 235879_at | ESTs | | | AI697540 | 36.0 | Above | 3.8 | new |
| 33 | 212387_at | cDNA FLJ11918 fis | FLJ11918 | | AK021980.1 | 35.8 | Below | 3.3 | new |
| 34 | 218988_at | bladder cancer over expressed protein | BLOV1 | 12q15 | NM_018656.1 | 35.8 | Below | 16.3 | new |
| 35 | 228555_at | EST; by BLAT calcium/calmodulin-dependent Protine Kinase type II Delta chain (CAMK GROUP I) | CAMK2D | | AA029441 | 35.8 | Above | 3.1 | new |
| 36 | 202975_s_at | Rho-related BTB domain containing 3 | RHOBTB3 | 5q21.2 | N21138 | 35.3 | Above | 5.5 | old |
| 37 | 201105_at | lectin, galactoside-binding, soluble, 1 (galectin 1) | LGALS1 | 22q13.1 | NM_002305.2 | 34.5 | Above | 14.5 | old |
| 38 | 203434_s_at | membrane metallo-endopeptidase (neutral endopeptidase, enkephalinase, CALLA, CD10) | MME | 3q25.1-q25.2 | AI433463 | 34.1 | Below | 31.2 | old |
| 39 | 212135_s_at | calcium transporting ATPase plasma membrane protein. | ATP2B4 | | AW517686 | 34.1 | Below | 2.4 | new |
| 40 | 212136_at | calcium transporting ATPase plasma membrane protein. | ATP2B4 | | AW517686 | 34.1 | Below | 2.1 | new |
| 41 | 230179_at | cDNA DKFZp547P158 | DKFZp547 P158 | | N52572 | 34.1 | Below | 6.4 | new |
| 42 | 218217_at | likely homolog of rat and mouse retinoid- inducible serine carboxypeptidase | RISC | 17q23.2 | NM_021626.1 | 32.8 | Above | 3.4 | new |
| 43 | 225841_at | hypothetical protein FLJ30525 | FLJ30525 | 1p13.2 | BE502436 | 32.8 | Above | 1.8 | new |
| 44 | 226668_at | hypothetical protein FLJ36175 | FLJ36175 | | W80623 | 32.8 | Above | 2.4 | new |
| 45 | 200989_at | hypoxia-inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor) | HIF1A | 14q21-q24 | NM_001530.1 | 32.2 | Below | 1.8 | old |
| 46 | 201151_s_at | muscleblind-like (Drosophila) | MBNL | 3q25 | NM_021038.1 | 32.2 | Above | 2.6 | old |
| 47 | 201563_at | sorbitol dehydrogenase | SORD | 15q15.3 | L29008.1 | 32.2 | Above | 1.8 | new |
| 48 | 203753_at | transcription factor 4 | TCF4 | 18q21.1 | NM_003199.1 | 32.2 | Below | 2.9 | new |
| 49 | 205668_at | lymphocyte antigen 75 | LY75 | 2q24 | NM_002349.1 | 32.2 | Above | 2.1 | new |
| 50 | 206471_s_at | plexin C1 | PLXNC1 | 12q23.3 | NM_005761.1 | 32.2 | Above | 7.7 | old |
| 51 | 211302_s_at | phosphodiesterase 4B, cAMP-specific | PDE4B | 1p31 | L20966.1 | 32.2 | Below | 3.0 | old |
| 52 | 212012_at | Melanoma associated gene | D2S448 | 2pter-p25.1 | AF200348.1 | 32.2 | Below | 2.4 | old |
| 53 | 212063_at | CD44 antigen | CD44 | 11p13 | BE903880 | 32.2 | Above | 3.1 | new |
| 54 | 213241_at | plexin C1 | PLXNC1 | | AF035307.1 | 32.2 | Above | 2.5 | old |
| 55 | 214651_s_at | homeo box A9 | HOXA9 | 7p15-p14 | U41813.1 | 32.2 | Above | 28.5 | old |
| 56 | 218140_x_at | APMCF1 protein | APMCF1 | 3q22.2 | NM_021203.1 | 32.2 | Above | 1.4 | new |
| 57 | 219988_s_at | hypothetical protein FLJ10597 | FLJ10597 | 1p34.1 | NM_018150.1 | 32.2 | Above | 1.9 | new |
| 58 | 223046_at | egl nine homolog 1 (C. elegans) | EGLN1 | 1q42.1 | NM_022051.1 | 32.2 | Below | 4.2 | new |
| 59 | 224150_s_at | p10-binding protein | BITE | 3q22-q23 | AF289495.1 | 32.2 | Above | 2.1 | new |
| 60 | 224933_s_at | hypothetical protein DKFZp761F0118 | DKFZp761 F0118 | 10q22.1 | AB037801.1 | 32.2 | Above | 1.9 | new |
| 61 | 201078_at | transmembrane 9 superfamily member 2 | TM9SF2 | 13q32.3 | NM_004800.1 | 32.0 | Above | 1.5 | new |

| | | | | | | | | | |
|----|-------------|--|----------|--------------|-------------|------|-------|-------|-----|
| 62 | 205550_s_at | brain and reproductive organ-expressed (TNFRSF1A modulator) | BRE | 2p23.3 | NM_004899.1 | 32.0 | Above | 2.0 | new |
| 63 | 212382_at | cDNA FLJ11918 fis | FLJ11918 | | AK021980.1 | 32.0 | Below | 2.7 | new |
| 64 | 225019_at | calcium/calmodulin-dependent protein kinase (CaM kinase) II delta | CAMK2D | 4q25 | AA777512 | 32.0 | Above | 3.6 | new |
| 65 | 225202_at | Rho-related BTB domain containing 3 | RHOBTB3 | 5q21.2 | BE620739 | 32.0 | Above | 5.5 | old |
| 66 | 228855_at | nudix (nucleoside diphosphate linked moiety X)-type motif 7 | NUDT7 | | AI927964 | 32.0 | Above | 5.6 | new |
| 67 | 231899_at | KIAA1726 protein | KIAA1726 | 11q23.1 | AB051513.1 | 32.0 | Above | 33.0 | new |
| 68 | 52164_at | chromosome 11 open reading frame 24 | C11orf24 | 11q13 | AA065185 | 32.0 | Above | 2.3 | new |
| 69 | 212660_at | KIAA0239 protein | KIAA0239 | 5q31.1 | AI735639 | 31.7 | Below | 1.7 | old |
| 70 | 213513_x_at | actin related protein 2/3 complex, subunit 2, 34kDa | ARPC2 | 2q36.1 | BG034239 | 31.7 | Above | 1.3 | old |
| 71 | 222603_at | hypothetical protein FLJ23309 | FLJ23309 | 9p24 | AL136980 | 31.7 | Above | 3.6 | new |
| 72 | 238558_at | ESTs | | | AI445833 | 31.7 | Above | 3.8 | new |
| 73 | 202391_at | brain abundant, membrane attached signal protein 1 | BASP1 | 5p15.1-p14 | NM_006317.1 | 31.3 | Above | 2.1 | old |
| 74 | 202604_x_at | a disintegrin and metalloproteinase domain 10 | ADAM10 | 15q22 | NM_001110.1 | 31.3 | Above | 1.8 | old |
| 75 | 203435_s_at | membrane metallo-endopeptidase (neutral endopeptidase, enkephalinase, CALLA, CD10) | MME | 3q25.1-q25.2 | NM_007287.1 | 31.3 | Below | 54.8 | old |
| 76 | 204445_s_at | arachidonate 5-lipoxygenase | ALOX5 | 10q11.2 | AI361850 | 31.3 | Below | 687.0 | old |
| 77 | 209705_at | likely ortholog of mouse metal response element binding transcription factor 2 | M96 | 1p22.1 | AF073293.1 | 31.3 | Below | 1.5 | old |
| 78 | 214366_s_at | arachidonate 5-lipoxygenase | ALOX5 | 10q11.2 | AA995910 | 31.3 | Below | 54.7 | old |
| 79 | 215000_s_at | fasciculation and elongation protein zeta 2 (zygin II) | FEZ2 | 2p21 | AL117593.1 | 31.3 | Above | 1.7 | new |
| 80 | 220643_s_at | Fas apoptotic inhibitory molecule | FAIM | 3q23 | NM_018147.1 | 31.3 | Above | 2.9 | new |
| 81 | 226459_at | Homo sapiens gastric cancer-related protein GCYS-20 (gcys-20) mRNA, complete cds | | | AW575754 | 31.3 | Above | 1.6 | new |
| 82 | 238712_at | ESTs | | | BF801735 | 31.3 | Above | 2.7 | new |
| 83 | 229686_at | cDNA FLJ35637 fis | FLJ35637 | | AI436587 | 31.0 | Below | 1.5 | new |
| 84 | 222620_s_at | hypothetical protein similar to mouse Dnajl1 | DNAJL1 | 10p11.23 | BF591419 | 29.8 | Above | 2.4 | new |
| 85 | 224516_s_at | hypothetical protein HSPC195 | HSPC195 | 5q31.3 | BC006428.1 | 29.8 | Above | 2.7 | new |
| 86 | 203217_s_at | sialyltransferase 9 (CMP-NeuAc:lactosylceramide alpha-2,3-sialyltransferase; GM3 synthase) | SIAT9 | 2p11.2 | NM_003896.1 | 28.8 | Below | 2.1 | new |
| 87 | 204030_s_at | schwannomin interacting protein 1 | SCHIP1 | 3q25.32 | NM_014575.1 | 28.8 | Below | 17.6 | old |
| 88 | 209191_at | tubulin beta-5 | TUBB-5 | | BC002654.1 | 28.8 | Above | 6.4 | new |
| 89 | 213541_s_at | v-ets erythroblastosis virus E26 oncogene like (avian) | ERG | 21q22.3 | AI351043 | 28.8 | Below | 2.8 | old |
| 90 | 213773_x_at | Williams Beuren syndrome chromosome region | WBSCR20 | 7q11.23 | AW248552 | 28.8 | Above | 1.3 | new |

| | | | | | | | | | |
|-----|-------------|---|----------|--------------|-------------|------|-------|------|-----|
| | | 20A | A | | | | | | |
| 91 | 219243_at | immunity associated protein 4 | HIMAP4 | 7q35 | NM_018326.1 | 28.8 | Below | 13.4 | new |
| 92 | 219256_s_at | hypothetical protein FLJ20356 | FLJ20356 | 4p16.1 | NM_018986.1 | 28.8 | Below | 2.6 | new |
| 93 | 223358_s_at | phosphodiesterase 7A | PDE7A | 8q13 | AW269834 | 28.8 | Above | 1.5 | new |
| 94 | 224796_at | development and differentiation enhancing factor 1 | DDEF1 | 8q24.1-q24.2 | W03103 | 28.8 | Below | 1.8 | new |
| 95 | 203076_s_at | MAD, mothers against decapentaplegic homolog 2 (Drosophila) | MADH2 | 18q21.1 | U65019.1 | 28.7 | Below | 2.0 | new |
| 96 | 212385_at | cDNA FLJ11918 fis | FLJ11918 | | AK021980.1 | 28.7 | Below | 3.2 | new |
| 97 | 216026_s_at | polymerase (DNA directed), epsilon | POLE | 12q24.3 | AL080203.1 | 28.7 | Below | 3.0 | old |
| 98 | 217118_s_at | KIAA0930 protein | KIAA0930 | 22q13.31 | AK025608.1 | 28.7 | Above | 1.9 | new |
| 99 | 219821_s_at | hypothetical protein FLJ20330 | FLJ20330 | 6pter-p22.1 | NM_018988.1 | 28.7 | Below | 5.5 | new |
| 100 | 201875_s_at | hypothetical protein FLJ21047 | FLJ21047 | 1q23.2 | NM_024569.1 | 28.5 | Above | 2.0 | new |

Table S7. Top 100 chi-square probe sets selected for T-ALL in decision tree format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Chi-square | T-ALL above/below mean | Fold change | old or new |
|----|----------------|--|----------|-------------|-------------------|------------|------------------------|-------------|------------|
| 1 | 201137_s_at | major histocompatibility complex, class II, DP beta 1 | HLA-DPB1 | 6p21.3 | NM_002121.1 | 100.0 | Below | 21.0 | old |
| 2 | 202113_s_at | sorting nexin 2 | SNX2 | 5q23 | AF043453.1 | 100.0 | Below | 4.2 | old |
| 3 | 202114_at | sorting nexin 2 | SNX2 | 5q23 | NM_003100.1 | 100.0 | Below | 4.6 | old |
| 4 | 203675_at | nucleobindin 2 | NUCB2 | 11p15.1-p14 | NM_005013.1 | 100.0 | Above | 3.6 | old |
| 5 | 204670_x_at | major histocompatibility complex, class II, DR beta 3 | HLA-DRB3 | 6p21.3 | NM_002125.1 | 100.0 | Below | 13.4 | old |
| 6 | 205297_s_at | CD79B antigen (immunoglobulin-associated beta) | CD79B | 17q23 | NM_000626.1 | 100.0 | Below | 23.3 | old |
| 7 | 205456_at | CD3E antigen, epsilon polypeptide (TiT3 complex) | CD3E | 11q23 | NM_000733.1 | 100.0 | Above | 20.7 | old |
| 8 | 206398_s_at | CD19 antigen | CD19 | 16p11.2 | NM_001770.1 | 100.0 | Below | 5693.6 | old |
| 9 | 208306_x_at | major histocompatibility complex, class II, DR beta 4 | HLA-DRB4 | 6p21.3 | NM_021983.2 | 100.0 | Below | 8.3 | old |
| 10 | 208894_at | major histocompatibility complex, class II, DR alpha | HLA-DRA | 6p21.3 | M60334.1 | 100.0 | Below | 20.9 | old |
| 11 | 209312_x_at | major histocompatibility complex, class II, DR beta 1 | HLA-DRB1 | 6p21.3 | U65585.1 | 100.0 | Below | 12.6 | old |
| 12 | 209619_at | CD74 antigen (invariant polypeptide of major histocompatibility complex, class II antigen- | CD74 | 5q32 | K01144.1 | 100.0 | Below | 15.1 | old |

associated)

| | | | | | | | | | |
|----|-------------|--|----------|----------------|-------------|-------|-------|--------|-----|
| 13 | 210116_at | SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome) | SH2D1A | Xq25-q26 | AF072930.1 | 100.0 | Above | 150.7 | old |
| 14 | 210982_s_at | major histocompatibility complex, class II, DR alpha | HLA-DRA | 6p21.3 | M60333.1 | 100.0 | Below | 23.4 | old |
| 15 | 211990_at | major histocompatibility complex, class II, DP alpha 1 | HLA-DPA1 | 6p21.3 | M27487.1 | 100.0 | Below | 19.6 | old |
| 16 | 211991_s_at | major histocompatibility complex, class II, DP alpha 1 | HLA-DPA1 | 6p21.3 | M27487.1 | 100.0 | Below | 24.5 | old |
| 17 | 213539_at | CD3D antigen, delta polypeptide (TiT3 complex) | CD3D | 11q23 | NM_000732.1 | 100.0 | Above | 35.7 | old |
| 18 | 214049_x_at | CD7 antigen (p41) | CD7 | 17q25.2-q25.3 | AI829961 | 100.0 | Above | 312.2 | old |
| 19 | 214551_s_at | CD7 antigen (p41) | CD7 | 17q25.2-q25.3 | NM_006137.2 | 100.0 | Above | 228.1 | old |
| 20 | 217147_s_at | T-cell receptor interacting molecule | TRIM | 3q13 | AJ240085.1 | 100.0 | Above | 42.6 | new |
| 21 | 217478_s_at | MHC, class IIa, HLA-DMA | HLA-DMA | | X76775 | 100.0 | Below | 11.9 | old |
| 22 | 221969_at | paired box gene 5 (B-cell lineage specific activator protein) | PAX5 | 9p13 | BF510692 | 100.0 | Below | 3922.0 | new |
| 23 | 227646_at | early B-cell factor | EBF | 5q34 | BG435302 | 100.0 | Below | 85.0 | new |
| 24 | 229487_at | cDNA FLJ39389 fis | FLJ39389 | 5 | W73890 | 100.0 | Below | 7685.7 | new |
| 25 | 229838_at | cDNA FLJ39156 fis | FLJ39156 | | A1377271 | 100.0 | Above | 12.7 | old |
| 26 | 232204_at | early B-cell factor | EBF | 5q34 | AF208502.1 | 100.0 | Below | 7129.1 | new |
| 27 | 203965_at | ubiquitin specific protease 20 | USP20 | 9q34.12-q34.13 | NM_006676.1 | 91.3 | Above | 9.0 | old |
| 28 | 204891_s_at | lymphocyte-specific protein tyrosine kinase | LCK | 1p34.3 | NM_005356.1 | 91.3 | Above | 13.8 | old |
| 29 | 205255_x_at | transcription factor 7 (T-cell specific, HMG-box) | TCF7 | 5q31.1 | NM_003202.1 | 91.3 | Above | 8.4 | old |
| 30 | 207655_s_at | B-cell linker | BLNK | 10q23.2-q23.33 | NM_013314.1 | 91.3 | Below | 103.2 | old |
| 31 | 209771_x_at | CD24 antigen (small cell lung carcinoma cluster 4 antigen) | CD24 | 6q21 | AA761181 | 91.3 | Below | 40.1 | old |
| 32 | 211796_s_at | T cell receptor beta locus | TRB | 7q34 | AF043179.1 | 91.3 | Above | 20.7 | old |
| 33 | 213792_s_at | insulin receptor | INSR | 19p13.3-p13.2 | AA485908 | 91.3 | Below | 8.0 | old |
| 34 | 215193_x_at | major histocompatibility complex, class II, DR beta 3 | HLA-DRB3 | 6p21.3 | AJ297586.1 | 91.3 | Below | 12.1 | old |
| 35 | 216379_x_at | KIAA1919 protein | KIAA1919 | 6q22.1 | AK000168.1 | 91.3 | Below | 44.0 | old |
| 36 | 219191_s_at | bridging integrator 2 | BIN2 | 12q13 | NM_016293.1 | 91.3 | Above | 271.0 | new |
| 37 | 219563_at | hypothetical protein FLJ21276 | FLJ21276 | 14q32.2 | NM_024633.1 | 91.3 | Below | 5.8 | new |
| 38 | 219724_s_at | KIAA0748 gene product | KIAA0748 | 12q12 | NM_014796.1 | 91.3 | Above | 11.6 | new |
| 39 | 221750_at | 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1 (soluble) | HMGCS1 | 5p14-p13 | BG035985 | 91.3 | Above | 3.4 | new |
| 40 | 226157_at | cDNA FLJ39131 fis | FLJ39131 | 3 | AI569747 | 91.3 | Above | 4.4 | old |
| 41 | 226496_at | hypothetical protein FLJ22611 | FLJ22611 | 9p11.1 | BG291039 | 91.3 | Below | 7.6 | new |
| 42 | 266_s_at | CD24 antigen (small cell lung carcinoma cluster 4 | CD24 | 6q21 | L33930 | 91.3 | Below | 69.7 | old |

| | | | | | | | | | | |
|----|-------------|---|----------|-----------------|-------------|------|-------|--------|----------|--|
| | | antigen) | | | | | | | | |
| 43 | 39318_at | T-cell leukemia/lymphoma 1A | TCL1A | 14q32.1 | X82240 | 91.3 | Below | 367.4 | old | |
| 44 | 204214_s_at | RAB32, member RAS oncogene family | RAB32 | 6q24.3 | NM_006834.1 | 90.6 | Above | 127.9 | new | |
| 45 | 204777_s_at | mal, T-cell differentiation protein | MAL | 2cen-q13 | NM_002371.2 | 90.6 | Above | 96.8 | old | |
| 46 | 204890_s_at | lymphocyte-specific protein tyrosine kinase | LCK | 1p34.3 | U07236.1 | 90.6 | Above | 18.6 | old | |
| 47 | 205049_s_at | CD79A antigen (immunoglobulin-associated alpha) | CD79A | 19q13.2 | NM_001783.1 | 90.6 | Below | 11.4 | old | |
| 48 | 205254_x_at | transcription factor 7 (T-cell specific, HMG-box) | TCF7 | 5q31.1 | AW027359 | 90.6 | Above | 352.0 | old | |
| 49 | 205504_at | Bruton agammaglobulinemia tyrosine kinase | BTK | Xq21.33-q22 | NM_000061.1 | 90.6 | Below | 6.6 | old | |
| 50 | 210915_x_at | T cell receptor beta locus | TRB | 7q34 | M15564.1 | 90.6 | Above | 15.9 | old | |
| 51 | 211211_x_at | SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome) | SH2D1A | Xq25-q26 | AF100542.1 | 90.6 | Above | 1963.5 | old | |
| 52 | 213830_at | T cell receptor delta locus | TRD | 14q11.2 | AW007751 | 90.6 | Above | 7411.2 | old | |
| 53 | 216191_s_at | T cell receptor delta locus | TRD | 14q11.2 | X72501.1 | 90.6 | Above | 253.7 | old | |
| 54 | 217143_s_at | T cell receptor delta locus | TRD | 14q11.2 | X06557.1 | 90.6 | Above | 151.9 | old | |
| 55 | 219528_s_at | B-cell CLL/lymphoma 11B (zinc finger protein) | BCL11B | 14q32.31-q32.32 | NM_022898.1 | 90.6 | Above | 11.6 | new | |
| 56 | 220418_at | ubiquitin associated and SH3 domain containing, A | UBASH3A | 21q22.3 | NM_018961.1 | 90.6 | Above | 759.3 | new | |
| 57 | 222895_s_at | B-cell CLL/lymphoma 11B (zinc finger protein) | BCL11B | 14q32.31-q32.32 | AA918317 | 90.6 | Above | 11.7 | new | |
| 58 | 223553_s_at | hypothetical protein FLJ22570 | FLJ22570 | 5q35.3 | BC004564.1 | 90.6 | Below | 6.1 | new | |
| 59 | 225090_at | HRD1 protein | HRD1 | 11q12 | AA844682 | 90.6 | Below | 3.6 | new | |
| 60 | 226459_at | Homo sapiens gastric cancer-related protein GCYS-20 (gcys-20) mRNA, complete cds | | | AW575754 | 90.6 | Below | 10.7 | new | |
| 61 | 228314_at | cDNA FLJ37485 fis | FLJ37485 | | BE877357 | 90.6 | Below | 4.7 | new | |
| 62 | 201384_s_at | membrane component, chromosome 17, surface marker 2 (ovarian carcinoma antigen CA125) | M17S2 | 17q21.1 | NM_005899.1 | 83.8 | Above | 3.3 | new | |
| 63 | 202540_s_at | 3-hydroxy-3-methylglutaryl-Coenzyme A reductase | HMGCR | 5q13.3-q14 | NM_000859.1 | 83.8 | Above | 4.4 | old | |
| 64 | 203198_at | cyclin-dependent kinase 9 (CDC2-related kinase) | CDK9 | 9q34.1 | NM_001261.1 | 83.8 | Below | 4.8 | old | |
| 65 | 203932_at | major histocompatibility complex, class II, DM beta | HLA-DMB | 6p21.3 | NM_002118.1 | 83.8 | Below | 7.9 | old | |
| 66 | 204613_at | phospholipase C, gamma 2 (phosphatidylinositol-specific) | PLCG2 | 16q24.1 | NM_002661.1 | 83.8 | Below | 3.9 | old | |
| 67 | 205267_at | POU domain, class 2, associating factor 1 | POU2AF1 | 11q23.1 | NM_006235.1 | 83.8 | Below | 11.2 | old | |
| 68 | 208650_s_at | CD24 antigen (small cell lung carcinoma cluster 4 antigen) | CD24 | 6q21 | BG327863 | 83.8 | Below | 74.7 | old | |
| 69 | 208651_x_at | CD24 antigen (small cell lung carcinoma cluster 4 antigen) | CD24 | 6q21 | M58664.1 | 83.8 | Below | 52.7 | old | |
| 70 | 209995_s_at | T-cell leukemia/lymphoma 1A | TCL1A | 14q32.1 | BC003574.1 | 83.8 | Below | 20166. | old 2 | |
| 71 | 210038_at | protein kinase C, theta | PRKCQ | 10p15 | AL137145 | 83.8 | Above | 12.7 | old | |
| 72 | 211126_s_at | cysteine and glycine-rich protein 2 | CSRP2 | 12q21.1 | U46006.1 | 83.8 | Below | 18.0 | old | |
| 73 | 220068_at | pre-B lymphocyte gene 3 | VPREB3 | 22q11.23 | NM_013378.1 | 83.8 | Below | 6559.8 | new | |

| | | | | | | | | | |
|-----|-------------|--|---------------|---------------|-------------|------|-------|--------|-----|
| 74 | 226245_at | cDNA DKFZp451C132 | DKFZp451C132 | | U55984 | 83.8 | Above | 8.7 | new |
| 75 | 202615_at | cDNA DKFZp686D0521 | DKFZp686D0521 | | BF222895 | 82.2 | Above | 3.1 | old |
| 76 | 224861_at | cDNA FLJ31057 fis | FLJ31057 | | BF477658 | 82.2 | Above | 3.5 | new |
| 77 | 201194_at | selenoprotein W, 1 | SEPW1 | 19q13.3 | NM_003009.1 | 82.0 | Above | 3.8 | old |
| 78 | 201349_at | solute carrier family 9 (sodium/hydrogen exchanger), isoform 3 regulatory factor 1 | SLC9A3R1 | 17q25.2 | NM_004252.1 | 82.0 | Above | 2.9 | old |
| 79 | 202539_s_at | 3-hydroxy-3-methylglutaryl-Coenzyme A reductase | HMGCR | 5q13.3-q14 | AL518627 | 82.0 | Above | 3.5 | old |
| 80 | 203588_s_at | transcription factor Dp-2 (E2F dimerization partner 2) | TFDP2 | 3q23 | BG034328 | 82.0 | Above | 17.5 | old |
| 81 | 204852_s_at | protein tyrosine phosphatase, non-receptor type 7 | PTPN7 | 1q32.1 | NM_002832.1 | 82.0 | Above | 9.5 | old |
| 82 | 207434_s_at | FXYP domain containing ion transport regulator 2 | FXYP2 | 11q23 | NM_021603.1 | 82.0 | Above | 14.6 | old |
| 83 | 208872_s_at | DNA segment, single copy probe LNS-CAI/LNS-CAII | D5S346 | 5q22-q23 | AA814140 | 82.0 | Below | 2.6 | old |
| 84 | 209200_at | MADS box transcription enhancer factor 2, polypeptide C (myocyte enhancer factor 2C) | MEF2C | 5q14 | N22468 | 82.0 | Below | 7.5 | old |
| 85 | 212795_at | KIAA1033 protein | KIAA1033 | 12q24.11 | AL137753.1 | 82.0 | Below | 2.4 | old |
| 86 | 212827_at | immunoglobulin heavy constant mu | IGHM | 14q32.33 | X17115.1 | 82.0 | Below | 13.1 | old |
| 87 | 213193_x_at | T cell receptor beta locus | TRB | 7q34 | AL559122 | 82.0 | Above | 10.9 | old |
| 88 | 221002_s_at | tetraspanin similar to TM4SF9 | DC-TM4F2 | 10q23.2 | NM_030927.1 | 82.0 | Below | 2.1 | new |
| 89 | 225314_at | hypothetical protein MGC45416 | MGC45416 | 4p12 | BG291649 | 82.0 | Above | 5.5 | new |
| 90 | 227432_s_at | insulin receptor | INSR | 19p13.3-p13.2 | AI215106 | 82.0 | Below | 6.0 | old |
| 91 | 203332_s_at | inositol polyphosphate-5-phosphatase, 145kDa | INPP5D | 2q36-q37 | NM_005541.1 | 81.5 | Below | 2.2 | old |
| 92 | 203589_s_at | transcription factor Dp-2 (E2F dimerization partner 2) | TFDP2 | 3q23 | NM_006286.1 | 81.5 | Above | 35.1 | old |
| 93 | 205674_x_at | FXYP domain containing ion transport regulator 2 | FXYP2 | 11q23 | NM_001680.2 | 81.5 | Above | 12.2 | old |
| 94 | 209881_s_at | Linker for activation of T cells | LAT | 16q13 | AF036905.1 | 81.5 | Above | 1823.4 | old |
| 95 | 211005_at | Linker for activation of T cells | LAT | 16q13 | AF036906.1 | 81.5 | Above | 67.8 | old |
| 96 | 211075_s_at | CD47 | CD47 | | Z25521.1 | 81.5 | Above | 2.1 | old |
| 97 | 211210_x_at | SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome) | SH2D1A | Xq25-q26 | AF100539.1 | 81.5 | Above | 300.2 | old |
| 98 | 213601_at | slit homolog 1 (Drosophila) | SLIT1 | 10q23.3-q24 | AB011537.2 | 81.5 | Above | 1752.1 | old |
| 99 | 213857_s_at | CD47 antigen (Rh-related antigen, integrin-associated signal transducer) | CD47 | 3q13.1-q13.2 | BG230614 | 81.5 | Above | 2.2 | old |
| 100 | 214924_s_at | KIAA1042 protein | KIAA1042 | 3p25.3-p24.1 | AK000754.1 | 81.5 | Below | 2.3 | old |

Table S8. Top 100 chi-square probe sets selected for *TEL-AML1* in decision tree format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Chi-square value | TEL-AML above/below mean | Fold change | old or new |
|----|-----------------------|---|---------------|-----------------|--------------------------|-------------------------|---------------------------------|--------------------|-------------------|
| 1 | 224722_at | KIAA1323 | KIAA1323 | 18q11.1 | W80418 | 75 | Above | 7.6 | new |
| 2 | 227377_at | FLJ12722 | FLJ12722 | 17q21.32 | AK022784.1 | 75 | Above | 2446.3 | new |
| 3 | 237206_at | Homo sapiens cDNA FLJ39434 | FLJ39434 | 17p12 | AI452798 | 75 | Above | 23.7 | new |
| 4 | 241505_at | EST | | | BF513468 | 75 | Above | 13.4 | new |
| 5 | 203184_at | Fibrillin 2 (congenital contractural arachnodactyly) | FBN2 | 5q23.2 | NM_001999.2 | 69.1 | Above | 14.4 | new |
| 6 | 205109_s_at | Rho guanine nucleotide exchange factor (GEF) 4 | ARHGEF4 | 2q22 | NM_015320.1 | 69.1 | Above | 148.1 | old |
| 7 | 210650_s_at | Piccolo (presynaptic cytomatrix protein) | PCLO | 7q21.11 | BC001304.1 | 69.1 | Above | 101.2 | old |
| 8 | 213558_at | Piccolo (presynaptic cytomatrix protein) | PCLO | 7q21.11 | AB011131.1 | 69.1 | Above | 77.5 | old |
| 9 | 220451_s_at | baculoviral IAP repeat-containing 7 (livin) | BIRC7 | 20q13.3 | NM_022161.1 | 69.1 | Above | 25.4 | new |
| 10 | 224720_at | KIAA1323 | KIAA1323 | 18q11.1 | W80418 | 69.1 | Above | 4.3 | new |
| 11 | 235694_at | Homo sapiens, clone IMAGE:4661943, mRNA, partial cds | | 20q13.33 | N49233 | 69.1 | Above | 9.3 | new |
| 12 | 202808_at | Hypothetical protein FLJ20154 | FLJ20154 | 10q24.32 | AK000161.1 | 68.9 | Above | 3.7 | old |
| 13 | 206032_at | Desmocollin 3 | DSC3 | 18q12.1 | AI797281 | 68.9 | Above | 54.1 | new |
| 14 | 206033_s_at | Desmocollin 3 | DSC3 | 18q12.1 | NM_001941.2 | 68.9 | Above | 357.1 | new |
| 15 | 209228_x_at | Putative prostate cancer tumor suppressor gene | N33 | 8p22 | U42349.1 | 68.9 | Above | 20.8 | new |
| 16 | 224725_at | KIAA1323 | KIAA1323 | 18q11.1 | W80418 | 68.9 | Above | 3.6 | new |
| 17 | 203910_at | PTPL1-associated RhoGAP 1 | PARG1 | 1p22.1 | NM_004815.1 | 64 | Above | 7.1 | new |
| 18 | 204849_at | Transcription factor-like 5 (basic helix-loop-helix) | TCFL5 | 20q13.33 | NM_006602.1 | 64 | Above | 8.9 | old |
| 19 | 206231_at | Potassium intermediate/small conductance calcium-activated channel, subfamily N, member 1 | KCNN1 | 19p13.1 | NM_002248.2 | 64 | Above | 72.7 | old |
| 20 | 208056_s_at | Core-binding factor, runt domain, alpha subunit 2; translocated to, 3 | CBFA2T3 | 16q24 | NM_005187.2 | 63 | Above | 2.5 | old |
| 21 | 211222_s_at | Huntingtin-associated protein 1 (neuroan 1) | HAP1 | 17q21.2 | AF040723.1 | 63 | Above | 80.8 | new |
| 22 | 223468_s_at | hypothetical protein from EUROIMAGE 363668 RGM: likely ortholog of chicken repulsive guidance molecule | RGM | 15q26.1 | AL136826.1 | 63 | Above | 10.6 | new |
| 23 | 227266_s_at | FYN-binding protein | FYB | 5p13.1 | BF679849 | 63 | Above | 3.1 | old |

| | | | | | | | | | |
|----|-------------|---|----------|----------|-------------|------|-------|-------|-----|
| 24 | 228158_at | ESTs, Highly similar to A43542 lymphocyte-specific protein 1 - human | | 2p11.1 | AI623211 | 63 | Above | 7.9 | new |
| 25 | 37986_at | Eerythropoietin receptor | EPOR | 19p13.2 | M60459 | 63 | Above | 15.5 | old |
| 26 | 203464_s_at | Epsin 2 | EPN2 | 17p11.1 | NM_014964.1 | 62.9 | Above | 43.3 | new |
| 27 | 213317_at | chloride intracellular channel 5 | CLIC5 | 6p21.1 | AL049313.1 | 62.9 | Above | 99.3 | old |
| 28 | 213423_x_at | Putative prostate cancer tumor suppressor | N33 | 8p22 | AI884858 | 62.9 | Above | 15.7 | new |
| 29 | 226817_at | Desmocollin 2 | DSC2 | 18q12.1 | AU154691 | 62.9 | Above | 48.3 | new |
| 30 | 227862_at | ESTs, Weakly similar to hypothetical protein FLJ22184 | | 1p35.1 | AA037766 | 62.9 | Above | 14.7 | new |
| 31 | 229339_at | EST | | | AI093327 | 62.9 | Above | 31.1 | new |
| 32 | 211795_s_at | FYN binding protein | FYB | 5p13.1 | AF198052.1 | 59.4 | Above | 4.1 | old |
| 33 | 218627_at | Hypothetical protein FLJ11259 | FLJ11259 | 12q23.1 | NM_018370.1 | 57.9 | Above | 4.6 | new |
| 34 | 221748_s_at | Tensin | TNS | 2q35 | AL046979 | 57.9 | Above | 6.6 | new |
| 35 | 200709_at | FK506 binding protein 1A (12kD) | FKBP1A | 20p13 | NM_000801.1 | 57.1 | Above | 1.8 | old |
| 36 | 204615_x_at | Isopentenyl-diphosphate delta isomerase | IDI1 | 10p15.3 | NM_004508.1 | 57.1 | Above | 2.6 | old |
| 37 | 208881_x_at | Isopentenyl-diphosphate delta isomerase | IDI1 | 10p15.3 | BC005247.1 | 57.1 | Above | 2.6 | old |
| 38 | 213301_x_at | Transcriptional intermediary factor 1 | TIF1 | 7q34 | AL538264 | 57.1 | Above | 2.0 | new |
| 39 | 221747_at | Tensin | TNS | 2q35 | AL046979 | 57.1 | Above | 49.2 | new |
| 40 | 224726_at | KIAA1323 | KIAA1323 | 18q11.1 | W80418 | 57.1 | Above | 26.1 | new |
| 41 | 231455_at | ESTs | | 2p25.2 | AA768888 | 57.1 | Above | 7.7 | new |
| 42 | 232750_at | Homo sapiens cDNA FLJ13750 | FLJ13750 | 2q35 | AU158570 | 57.1 | Above | 35.0 | new |
| 43 | 209685_s_at | Protein kinase C, beta 1 | PRKCB1 | 16p11.2 | M13975.1 | 53.6 | Above | 1.9 | old |
| 44 | 204404_at | solute carrier family 12 (sodium/potassium/chloride transporters), member 2 | SLC12A2 | 5q23.3 | NM_001046.1 | 53.4 | Above | 2.0 | new |
| 45 | 239673_at | ESTs | | 4q31.23 | AW080999 | 53.4 | Above | 9.0 | new |
| 46 | 240950_s_at | Homo sapiens cDNA FLJ32658 | FLJ32658 | 19q13.33 | AA400740 | 53.4 | Above | 9.9 | new |
| 47 | 204297_at | Phosphoinositide-3-kinase, class 3 | PIK3C3 | 18q12.3 | NM_002647.1 | 52.5 | Above | 4.5 | old |
| 48 | 206591_at | Recombination activating gene 1 | RAG1 | 11p13 | NM_000448.1 | 52.1 | Above | 5.4 | old |
| 49 | 209962_at | Erythropoietin receptor | EPOR | 19p13.2 | M34986.1 | 52.1 | Above | 17.0 | old |
| 50 | 209963_s_at | Erythropoietin receptor | EPOR | 19p13.2 | M34986.1 | 52.1 | Above | 7.6 | old |
| 51 | 210186_s_at | FK506 binding protein 1A (12kD) | FKBP1A | 20p13 | BC005147.1 | 52.1 | Above | 1.8 | old |
| 52 | 219866_at | Chloride intracellular channel 5 | CLIC5 | 6p21.1 | NM_016929.1 | 52.1 | Above | 60.3 | old |
| 53 | 203474_at | IQ motif containing GTPase activating protein 2 | IQGAP2 | 5q13.2 | NM_006633.1 | 51.6 | Below | 2.8 | old |
| 54 | 210058_at | Mitogen-activated protein kinase 13 | MAPK13 | 6p21.1 | BC000433.1 | 51.6 | Above | 2.3 | new |
| 55 | 211891_s_at | Rho guanine nucleotide exchange factor (GEF) 4 | ARHGEF4 | 2q22 | AB042199.1 | 51.6 | Above | 452.6 | old |
| 56 | 214214_s_at | Complement component 1, q subcomponent binding protein | C1QBP | 17p13.3 | AU151801 | 51.6 | Below | 2.0 | old |
| 57 | 218152_at | High-mobility group 20A | HMG20A | 15q24 | NM_018200.1 | 51.6 | Above | 1.7 | new |
| 58 | 234983_at | ESTs | FLJ21415 | 12q24.22 | BE893995 | 51.6 | Above | 2.4 | new |

| | | | | | | | | | |
|----|-------------|--|-----------|-------------|-------------|------|-------|-------|-----|
| 59 | 240446_at | KIAA1323 | KIAA1323 | 18q11.2 | AI798164 | 51.6 | Above | 102.2 | new |
| 60 | 244107_at | ESTs | | 18q12.1 | AW189097 | 51.6 | Above | 518.9 | new |
| 61 | 205794_s_at | Neuro-oncological ventral antigen 1 | NOVA1 | 14q12 | NM_002515.1 | 51.4 | Above | 40.4 | new |
| 62 | 217628_at | chloride intracellular channel 5 | CLIC5 | 6p21.1 | BF032808 | 51.4 | Above | 87.4 | old |
| 63 | 218804_at | Hypothetical protein FLJ10261 | FLJ10261 | 11q13.3 | NM_018043.1 | 51.4 | Above | 41.6 | new |
| 64 | 230698_at | Homo sapiens mRNA; cDNA DKFZp434H205 | | 7q11.22 | AW072102 | 51.4 | Above | 8.7 | new |
| 65 | 225129_at | cDNA FLJ37548 fis | FLJ37548 | 16q13 | AW170571 | 49.4 | Above | 3.0 | new |
| 66 | 201266_at | Thioredoxin reductase 1 | TXNRD1 | 12q23-q24.1 | NM_003330.1 | 48.2 | Above | 1.7 | old |
| 67 | 203611_at | Telomeric repeat binding factor 2 | TERF2 | 16q22.1 | NM_005652.1 | 48.2 | Above | 5.3 | old |
| 68 | 213017_at | abhydrolase domain containing 3 | ABHD3 | 18q11.1 | AL534702 | 48.2 | Above | 4.0 | new |
| 69 | 236430_at | hypothetical protein MGC23911 | MGC23911 | 16q22.1 | AA708152 | 48.2 | Above | 16.8 | new |
| 70 | 209035_at | Midkine (neurite growth-promoting factor 2). | MDK | 11p11.2 | M69148.1 | 47.7 | Above | 4.6 | old |
| 71 | 209193_at | Pim-1 oncogene | PIM1 | 6p21.2 | M24779.1 | 47.7 | Above | 2.0 | old |
| 72 | 218625_at | Neuritin 1 | NRN1 | 6p24.1 | NM_016588.1 | 47.7 | Above | 5.1 | new |
| 73 | 226038_at | Hypothetical protein FLJ23749 | FLJ23749 | 8p23.1 | BF680438 | 47.7 | Above | 5.2 | new |
| 74 | 232227_at | ESTs, Moderately similar to AF161442_1 HSPC324 | | 9q34.3 | AV736391 | 47.7 | Above | 14.7 | new |
| 75 | 204160_s_at | Ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function) | ENPP4 | 6p12.3 | AW194947 | 46.5 | Above | 7.2 | new |
| 76 | 206233_at | UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 6 | B4GALT6 | 18q11 | AF097159.1 | 46.5 | Above | 2.6 | new |
| 77 | 218813_s_at | SH3-domain GRB2-like endophilin B2 | SH3GLB2 | 9q34.11 | NM_020145.1 | 46.5 | Above | 6.2 | new |
| 78 | 227111_at | Homo sapiens cDNA FLJ31099 fis, clone IMR321000230 | FLJ31099 | 9q33 | BG179317 | 46.5 | Above | 2.7 | new |
| 79 | 202382_s_at | Glucosamine-6-phosphate isomerase | GNPI | 5q21 | NM_005471.1 | 46.2 | Above | 5.6 | new |
| 80 | 202838_at | Fucosidase, alpha-L- 1, tissue | FUCA1 | 1p34 | NM_000147.1 | 46.2 | Above | 4.8 | old |
| 81 | 225731_at | Hypothetical protein KIAA1223 | KIAA1223 | 4q26 | AB033049.1 | 46.2 | Above | 2.8 | new |
| 82 | 225835_at | Homo sapiens cDNA: FLJ21409 fis, clone COL03924 | | 5q23.2 | AK025062.1 | 46.2 | Above | 3.6 | new |
| 83 | 229790_at | Telomeric repeat binding factor 2 | TERF2 | 16q22.1 | AW006832 | 46.2 | Above | 7.4 | old |
| 84 | 230069_at | Hypothetical protein FLJ12876 | FLJ12876 | 5q35.3 | BF593817 | 46.2 | Above | 9.4 | new |
| 85 | 235872_at | ESTs | | | BE408975 | 46.2 | Above | 17.7 | new |
| 86 | 239300_at | EST | | 18q12.3 | AI632214 | 46.2 | Above | 3.0 | new |
| 87 | 241940_at | EST | | 18q11.2 | BF477544 | 46.2 | Above | 2.9 | new |
| 88 | 203370_s_at | Enigma (LIM domain protein) | ENIGMA | 5q35.3 | NM_005451.2 | 45.9 | Above | 8.1 | new |
| 89 | 215149_at | LOC149153: | LOC149153 | 1p36.32 | AF052109.1 | 45.9 | Above | 9.2 | new |
| 90 | 217901_at | Desmoglein 2 desmosomal cadherin | DSG2 | 18q12.1 | BF031829 | 45.9 | Above | 6.7 | new |
| 91 | 235333_at | UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 6 | B4GALT6 | 18q12.1 | BG503479 | 45.9 | Above | 2.0 | new |

| | | | | | | | | | |
|-----|-------------|--|----------|---------|-------------|------|-------|------|-----|
| 92 | 242881_x_at | EST | | | BG285837 | 45.9 | Above | 11.8 | new |
| 93 | 200783_s_at | Stathmin 1/oncoprotein 18 leukemia-associated phosphoprotein | STMN1 | 1p35.1 | NM_005563.2 | 45.8 | Above | 1.5 | old |
| 94 | 201334_s_at | Rho guanine nucleotide exchange factor (GEF) 12 | ARHGEF12 | 11q23.3 | NM_015313.1 | 45.8 | Above | 6.1 | new |
| 95 | 203038_at | Protein tyrosine phosphatase, receptor type, K | PTPRK | 6q22.33 | NM_002844.1 | 45.8 | Above | 9.1 | old |
| 96 | 209735_at | ATP-binding cassette, sub-family G (WHITE), member 2 | ABCG2 | 4q22 | AF098951.2 | 45.8 | Above | 4.5 | new |
| 97 | 212063_at | Unactive progesterone receptor, 23 kD | P23 | 12q12 | BE903880 | 45.8 | Below | 7.4 | new |
| 98 | 212399_s_at | KIAA0121 gene product | KIAA0121 | 3p25.2 | D50911.2 | 45.8 | Above | 1.8 | old |
| 99 | 212438_at | Putative nucleic acid binding protein RY-1 | RY1 | 2p13.1 | BG252325 | 45.2 | Above | 1.7 | old |
| 100 | 214761_at | OLF-1/early B-cell factor associated zinc finger protein | OAZ | 16q12 | AW149417 | 45.2 | Above | 2.1 | old |

Table S9. Top 100 chi-square probe sets selected for *BCR-ABL* in parallel format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Unigene Reference | Chi-square value | BCR-ABL above/below mean | Fold change |
|----|-----------------------|--|---------------|-----------------|--------------------------|--------------------------|-------------------------|---------------------------------|--------------------|
| 1 | 201876_at | paraoxonase 2 | PON2 | 7q21.3 | NM_000305.1 | Hs.169857 | 70.8 | Above | 18.6 |
| 2 | 210830_s_at | paraoxonase 2 | PON2 | 7q21.3 | AF001602.1 | Hs.169857 | 70.8 | Above | 23.4 |
| 3 | 209365_s_at | extracellular matrix protein 1 | ECM1 | 1q21 | U65932.1 | Hs.81071 | 70.4 | Above | 6.0 |
| 4 | 209238_at | syntaxin 3A | STX3A | 11q12.3 | BE966922 | Hs.82240 | 66.5 | Above | 2.2 |
| 5 | 212242_at | tubulin, alpha 1 (testis specific) | TUBA1 | 2q36.2 | AL565074 | Hs.75318 | 64.9 | Above | 3.2 |
| 6 | 222488_s_at | dynactin 4 (p62) | DCTN4 | 5q31-q32 | BE218028 | Hs.180952 | 63.3 | Above | 3.6 |
| 7 | 222762_x_at | LIM domains containing 1 | LIMD1 | 3p21.3 | AU144259 | Hs.48469 | 63.3 | Above | 2.6 |
| 8 | 201310_s_at | P311 protein. Similar to gastrin/cholecystokinin type B receptor | P311 | 5q21.3 | NM_004772.1 | Hs.142827 | 61.3 | Below | 2.2 |
| 9 | 215617_at | Homo sapiens cDNA FLJ11754 fis, clone HEMBA1005588 | | | AU145711 | Hs.301006 | 61.3 | Above | 15.8 |
| 10 | 200951_s_at | cyclin D2 | CCND2 | 12p13 | NM_001759.1 | Hs.75586 | 60.9 | Above | 12.6 |
| 11 | 204430_s_at | solute carrier family 2 (facilitated glucose/fructose transporter), member 5 | SLC2A5 | 1p36.2 | NM_003039.1 | Hs.33084 | 60.9 | Above | 5.1 |
| 12 | 200953_s_at | cyclin D2 | CCND2 | 12p13 | NM_001759.1 | Hs.75586 | 60.7 | Above | 3.6 |
| 13 | 222154_s_at | DKFZP564A2416 protein | DKFZP564A2416 | 2q33.1 | AK002064.1 | Hs.5297 | 60.3 | Above | 12.3 |
| 14 | 204429_s_at | solute carrier family 2 (facilitated glucose/fructose transporter), member 5 | SLC2A5 | 1p36.2 | BE560461 | Hs.33084 | 57.1 | Above | 5.0 |

| | | | | | | | | | |
|----|-------------|---|----------|----------------|-------------|-----------|------|-------|------|
| 15 | 219315_s_at | hypothetical protein FLJ20898 | FLJ20898 | 16p13.12 | NM_024600.1 | Hs.25549 | 57.1 | Above | 5.4 |
| 16 | 241812_at | Homo sapiens cDNA FLJ35598 fis, clone SPLEN2008317 | | | AV648669 | Hs.199438 | 55.9 | Above | 5.2 |
| 17 | 201906_s_at | HYA22 protein | HYA22 | 3p21.3 | NM_005808.1 | Hs.147189 | 54.0 | Above | 43.1 |
| 18 | 202947_s_at | glycophorin C (Gerbich blood group) | GYPC | 2q14-q21 | NM_002101.2 | Hs.81994 | 54.0 | Above | 3.1 |
| 19 | 242579_at | ESTs | | | AA935461 | Hs.161712 | 54.0 | Above | 10.1 |
| 20 | 212203_x_at | interferon induced transmembrane protein 3 (1-8U) | IFITM3 | 11 | BF338947 | Hs.182241 | 52.0 | Above | 2.3 |
| 21 | 240173_at | ESTs | | | AI732969 | Hs.126245 | 52.0 | Above | 10.2 |
| 22 | 212298_at | neuropilin 1 | NRP1 | 10p12 | BE620457 | Hs.69285 | 51.8 | Above | 13.7 |
| 23 | 212592_at | immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides | IGJ | 4q21 | AV733266 | Hs.76325 | 51.8 | Above | 8.0 |
| 24 | 222868_s_at | interleukin 18 binding protein | IL18BP | 11q13 | AI521549 | Hs.325978 | 51.6 | Above | 7.1 |
| 25 | 235988_at | ESTs | | | AA746038 | Hs.125343 | 51.6 | Above | 15.6 |
| 26 | 239273_s_at | matrix metalloproteinase 28 | MMP28 | 17q11-q21.1 | AI927208 | Hs.231958 | 51.6 | Above | 89.8 |
| 27 | 219938_s_at | proline-serine-threonine phosphatase interacting protein 2 | PSTPIP2 | 18q12 | NM_024430.1 | Hs.69149 | 51.6 | Above | 5.2 |
| 28 | 207196_s_at | Nef-associated factor 1 | NAF1 | 5q32-q33.1 | NM_006058.1 | Hs.109281 | 47.2 | Above | 1.4 |
| 29 | 227533_at | ESTs | | | AA732944 | Hs.5415 | 47.1 | Below | 2.7 |
| 30 | 217110_s_at | mucin 4 | MUC4 | 3q29 | AJ242547.1 | Hs.198267 | 45.0 | Above | 47.2 |
| 31 | 220684_at | T-box 21 | TBX21 | 17q21.2 | NM_013351.1 | Hs.272409 | 45.0 | Above | 3.3 |
| 32 | 227182_at | hypothetical protein MGC26847 | MGC26847 | 9q22.2 | AW966474 | Hs.88417 | 45.0 | Above | 7.2 |
| 33 | 223075_s_at | hypothetical protein FLJ12783 | FLJ12783 | 9q34.13-q34.3 | AL136566.1 | Hs.4944 | 45.0 | Above | 3.9 |
| 34 | 202748_at | guanylate binding protein 2, interferon-inducible | GBP2 | 1p22.1 | NM_004120.2 | Hs.171862 | 44.5 | Above | 3.9 |
| 35 | 214657_s_at | multiple endocrine neoplasia I | MEN1 | 11q13 | AU134977 | Hs.240443 | 43.3 | Above | 2.3 |
| 36 | 202771_at | KIAA0233 gene product | KIAA0233 | 16q24.3 | NM_014745.1 | Hs.79077 | 43.3 | Above | 1.9 |
| 37 | 217967_s_at | chromosome 1 open reading frame 24 | C1orf24 | 1q25 | AF288391.1 | Hs.48778 | 43.3 | Above | 3.3 |
| 38 | 220024_s_at | Periaxin | PRX | 19q13.13-q13.2 | NM_020956.1 | Hs.205457 | 43.3 | Above | 8.1 |
| 39 | 203217_s_at | sialyltransferase 9 | SIAT9 | 2p11.2 | NM_003896.1 | Hs.225939 | 43.2 | Above | 1.8 |
| 40 | 238689_at | Homo sapiens cDNA FLJ30646 fis, clone CTONG2004716, weakly similar to Rattus norvegicus mRNA for seven transmembrane receptor | | | BG426455 | Hs.256897 | 43.2 | Above | 10.8 |
| 41 | 200665_s_at | secreted protein, acidic, cysteine-rich (osteonectin) | SPARC | 5q31.3-q32 | NM_003118.1 | Hs.111779 | 42.9 | Above | 10.5 |
| 42 | 200864_s_at | RAB11A, member RAS oncogene family | RAB11A | 15q21.3-q22.31 | NM_004663.1 | Hs.75618 | 42.9 | Above | 1.4 |
| 43 | 213075_at | Homo sapiens, clone IMAGE:3140802, mRNA | | | AL050002.1 | Hs.94795 | 42.9 | Above | 26.4 |
| 44 | 236489_at | ESTs | | | AI282097 | Hs.72307 | 42.9 | Above | 16.6 |
| 45 | 48106_at | hypothetical protein FLJ20489 | FLJ20489 | 12p11.1 | H14241 | Hs.306989 | 42.6 | Above | 2.8 |

| | | | | | | | | | |
|----|-------------|---|--------------|--------------|-------------|-----------|------|-------|-------|
| 46 | 203005_at | lymphotoxin beta receptor (TNFR superfamily, member 3) | LTBR | 12p13 | NM_002342.1 | Hs.1116 | 42.6 | Above | 9.9 |
| 47 | 205467_at | caspase 10, apoptosis-related cysteine protease | CASP10 | 2q33-q34 | NM_001230.1 | Hs.5353 | 42.6 | Above | 3.6 |
| 48 | 220454_s_at | semaphorin 6A | SEMA6A | 5q23.1 | NM_020796.1 | Hs.263395 | 42.6 | Above | 3.3 |
| 49 | 228696_at | prostein protein | LOC85414 | 1q31.1 | AA631143 | Hs.278695 | 42.6 | Above | 95.9 |
| 50 | 235306_at | hypothetical protein DKFZp667I133 | DKFZp667I133 | 7q35 | AI611648 | Hs.95834 | 42.6 | Above | 3.5 |
| 51 | 239272_at | matrix metalloproteinase 28 | MMP28 | 17q11-q21.1 | AI927208 | Hs.231958 | 42.6 | Above | 14.2 |
| 52 | 242572_at | ESTs | | | BF435438 | Hs.269924 | 42.6 | Above | 4.9 |
| 53 | 242677_at | ESTs | | | AI088099 | Hs.173548 | 42.6 | Above | 469.2 |
| 54 | 244597_at | ESTs | | | AA701247 | Hs.269548 | 42.6 | Above | 11.7 |
| 55 | 218013_x_at | dynactin 4 (p62) | DCTN4 | 5q31-q32 | NM_016221.1 | Hs.180952 | 41.5 | Above | 3.6 |
| 56 | 218084_x_at | FXYD domain containing ion transport regulator 5 | FXYD5 | 19q12-q13.1 | NM_014164.2 | Hs.333418 | 41.5 | Above | 1.5 |
| 57 | 221790_s_at | LDL receptor adaptor protein | ARH | 1p36-p35 | AL545035 | Hs.184482 | 41.5 | Above | 3.5 |
| 58 | 201028_s_at | antigen identified by monoclonal antibodies 12E7, F21 and O13 | MIC2 | Xp22.32 | U82164.1 | Hs.177543 | 41.4 | Above | 2.6 |
| 59 | 210105_s_at | FYN oncogene related to SRC, FGR, YES | FYN | 6q21 | M14333.1 | Hs.169370 | 40.7 | Above | 1.8 |
| 60 | 201445_at | calponin 3, acidic | CNN3 | 1p22-p21 | NM_001839.1 | Hs.194662 | 39.8 | Above | 10.7 |
| 61 | 229139_at | Homo sapiens, clone IMAGE:4245141, mRNA | | | AI202201 | Hs.293836 | 39.5 | Above | 10.7 |
| 62 | 229975_at | ESTs | | | AI826437 | Hs.72472 | 39.0 | Above | 9.0 |
| 63 | 205055_at | integrin, alpha E (antigen CD103, human mucosal lymphocyte antigen 1; alpha polypeptide) | ITGAE | 17p13 | NM_002208.3 | Hs.851 | 38.8 | Below | 2.1 |
| 64 | 226190_at | glucose phosphate isomerase | GPI | 19q13.1 | BG029496 | Hs.279789 | 38.8 | Above | 3.5 |
| 65 | 209732_at | C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 2 (activation-induced) | CLECSF2 | 12p13-p12 | BC005254.1 | Hs.85201 | 37.1 | Above | 2.3 |
| 66 | 210487_at | deoxynucleotidyltransferase, terminal | DNTT | 10q23-q24 | M11722.1 | Hs.272537 | 37.1 | Above | 2.0 |
| 67 | 203508_at | tumor necrosis factor receptor superfamily, member 1B | TNFRSF1B | 1p36.3-p36.2 | NM_001066.1 | Hs.256278 | 36.4 | Above | 2.6 |
| 68 | 209321_s_at | adenylate cyclase 3 | ADCY3 | 2p24-p22 | AF033861.1 | Hs.8402 | 36.4 | Above | 2.0 |
| 69 | 212552_at | hippocalcin-like 1 | HPCAL1 | 2p25.1 | BE617588 | Hs.3618 | 36.4 | Above | 2.0 |
| 70 | 219871_at | hypothetical protein FLJ13197 | FLJ13197 | 4p14 | NM_024614.1 | Hs.29725 | 36.4 | Above | 14.4 |
| 71 | 203159_at | glutaminase | GLS | 2q32-q34 | NM_014905.1 | Hs.239189 | 36.4 | Above | 1.5 |
| 72 | 203725_at | growth arrest and DNA-damage-inducible, alpha | GADD45A | 1p31.2-p31.1 | NM_001924.2 | Hs.80409 | 36.4 | Above | 3.1 |
| 73 | 208893_s_at | dual specificity phosphatase 6 | DUSP6 | 12q22-q23 | BC005047.1 | Hs.180383 | 36.4 | Above | 3.6 |
| 74 | 228297_at | calponin 3, acidic | CNN3 | 1p22-p21 | AI807004 | Hs.194662 | 36.4 | Above | 4.9 |
| 75 | 231055_at | ESTs | | | BF432941 | Hs.263462 | 36.4 | Above | 3.4 |
| 76 | 202123_s_at | v-abl Abelson murine leukemia viral oncogene homolog 1 | ABL1 | 9q34.1 | NM_005157.2 | Hs.146355 | 35.2 | Above | 1.8 |

| | | | | | | | | | |
|-----|-------------|--|------------------|---------------|-------------|-----------|------|-------|------|
| 77 | 225108_at | ESTs | | | BF111719 | Hs.356237 | 35.0 | Above | 3.0 |
| 78 | 239519_at | ESTs | | | AA927670 | Hs.131704 | 35.0 | Above | 18.1 |
| 79 | 64064_at | immune associated nucleotide 4 like 1 (mouse) | IAN4L1 | 7q36 | AI435089 | Hs.26194 | 34.1 | Above | 3.4 |
| 80 | 204576_s_at | KIAA0643 protein | KIAA0643 | 16p12.3 | AA207013 | Hs.155995 | 34.1 | Above | 1.9 |
| 81 | 205376_at | inositol polyphosphate-4-phosphatase, type II, 105kDa | INPP4B | 4q31.1 | NM_003866.1 | Hs.153687 | 34.1 | Above | 16.7 |
| 82 | 218693_at | transmembrane 4 superfamily member tetraspan NET-7 | NET-7 | 10q22.3 | NM_012339.1 | Hs.95583 | 34.1 | Above | 21.9 |
| 83 | 219073_s_at | oxysterol binding protein-like 10 | OSBPL10 | 3p22.3 | NM_017784.1 | Hs.321622 | 34.1 | Above | 4.9 |
| 84 | 222126_at | insulin receptor substrate 3-like | IRS3L | 7q22 | AI247494 | Hs.30827 | 34.1 | Above | 7.4 |
| 85 | 216985_s_at | syntaxin 3A | STX3A | 11q12.3 | AJ002077.1 | Hs.82240 | 34.1 | Above | 11.9 |
| 86 | 201809_s_at | endoglin (Osler-Rendu-Weber syndrome 1) | ENG | 9q33-q34.1 | NM_000118.1 | Hs.76753 | 34.1 | Above | 6.0 |
| 87 | 201204_s_at | ribosome binding protein 1 homolog 180kDa | RRBP1 | 20p12 | AI921320 | Hs.98614 | 33.7 | Above | 2.6 |
| 88 | 201743_at | CD14 antigen | CD14 | 5q31.1 | NM_000591.1 | Hs.75627 | 33.7 | Above | 3.6 |
| 89 | 202180_s_at | major vault protein | MVP | 16p13.1-p11.2 | NM_017458.1 | Hs.80680 | 33.7 | Above | 4.3 |
| 90 | 203253_s_at | KIAA0433 protein | KIAA0433 | 5q15 | NM_015216.1 | Hs.26179 | 33.7 | Above | 1.5 |
| 91 | 209030_s_at | immunoglobulin superfamily, member 4 | IGSF4 | 11q23.2 | NM_014333.1 | Hs.70337 | 33.7 | Above | 10.8 |
| 92 | 209829_at | chromosome 6 open reading frame 32 | C6orf32 | 6p22.3-p21.32 | AB002384.1 | Hs.101359 | 33.7 | Above | 2.1 |
| 93 | 212365_at | myosin IB | MYO1B | 2q12-q34 | AK000160.1 | Hs.121576 | 33.7 | Above | 3.0 |
| 94 | 214255_at | ATPase, Class V, type 10C | ATP10C | 15q11-q13 | AB011138.1 | Hs.44697 | 33.7 | Above | 9.8 |
| 95 | 214321_at | nephroblastoma over expressed gene | NOV | 8q24.1 | BF440025 | Hs.235935 | 33.7 | Above | 49.2 |
| 96 | 218086_at | neural proliferation, differentiation and control, 1 | NPDC1 | 9q34.3 | NM_015392.1 | Hs.105547 | 33.7 | Above | 28.8 |
| 97 | 222937_s_at | matrix metalloproteinase 28 | MMP28 | 17q11-q21.1 | AF219624.1 | Hs.231958 | 33.7 | Above | 70.2 |
| 98 | 225113_at | Homo sapiens mRNA; cDNA DKFZp762O2215 (from clone DKFZp762O2215) | | | BF111719 | Hs.331666 | 33.7 | Above | 7.2 |
| 99 | 225244_at | IMAGE3451454: GRASP protein | IMAGE345 1454 | 1q42.13 | AA019893 | Hs.325081 | 33.7 | Above | 2.0 |
| 100 | 228573_at | capillary morphogenesis protein 2 | CMG2 | 4q21.21 | BE673665 | Hs.5897 | 33.7 | Above | 2.2 |

Table S10. Top 100 chi-square probe sets selected for *E2A-PBX1* in parallel format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Unigene Reference | Chi-square value | E2A-PBX Above/below mean | Fold change |
|---|----------------|--|--------|----------|-------------------|-------------------|------------------|--------------------------|-------------|
| 1 | 201695_s_at | nucleoside phosphorylase | NP | 14q13.1 | NM_000270.1 | Hs.75514 | 100.0 | Above | 3.8 |
| 2 | 204674_at | lymphoid-restricted membrane protein | LRMP | 12p12.3 | NM_006152.1 | Hs.40202 | 100.0 | Above | 5.8 |
| 3 | 205253_at | pre-B-cell leukemia transcription factor 1 | PBX1 | 1q23 | NM_002585.1 | Hs.155691 | 100.0 | Above | 3549.2 |

| | | | | | | | | | |
|----|-------------|--|----------|------------|-------------|-----------|-------|-------|--------|
| 4 | 212148_at | pre-B-cell leukemia transcription factor 1, splice variant | PBX1 | | | | 100.0 | Above | 5283.5 |
| 5 | 212151_at | pre-B-cell leukemia transcription factor 1, splice variant | PBX1 | | | | 100.0 | Above | 7472.2 |
| 6 | 212371_at | Homo sapiens mRNA; cDNA DKFZp586C1019 (from clone DKFZp586C1019) | | | AL049397.1 | Hs.12314 | 100.0 | Above | 2.5 |
| 7 | 219155_at | retinal degeneration B beta | RDGBB | 17q24.2 | NM_012417.1 | Hs.333212 | 100.0 | Above | 2.7 |
| 8 | 225483_at | hypothetical protein MGC10485 | MGC10485 | 11q25 | AI971602 | Hs.334684 | 100.0 | Above | 7.7 |
| 9 | 227439_at | E2a-Pbx1-associated protein | EB-1 | 12 | AW005572 | Hs.372732 | 100.0 | Above | 267.7 |
| 10 | 227949_at | Q9H4T4 like | H17739 | 20q13.32 | AL357503 | Hs.288513 | 100.0 | Above | 59.8 |
| 11 | 230306_at | hypothetical protein MGC10485 | MGC10485 | 11q25 | AA514326 | Hs.334684 | 100.0 | Above | 19.1 |
| 12 | 231095_at | ESTs, Moderately similar to retinal degeneration B beta [Homo sapiens] [H.sapiens] | | | AW193811 | Hs.112703 | 100.0 | Above | 26.7 |
| 13 | 35974_at | lymphoid-restricted membrane protein | LRMP | 12p12.3 | U10485 | Hs.40202 | 91.8 | Above | 6.1 |
| 14 | 38340_at | huntingtin interacting protein 12 | HIP12 | 12q24 | AB014555 | Hs.96731 | 91.8 | Above | 3.8 |
| 15 | 206028_s_at | c-mer proto-oncogene tyrosine kinase | MERTK | 2q14.1 | NM_006343.1 | Hs.306178 | 91.8 | Above | 24.2 |
| 16 | 206181_at | signaling lymphocytic activation molecule | SLAM | 1q22-q23 | NM_003037.1 | Hs.32970 | 91.8 | Above | 6.4 |
| 17 | 208644_at | ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase) | ADPRT | 1q41-q42 | M32721.1 | Hs.177766 | 91.3 | Above | 3.0 |
| 18 | 221113_s_at | wingless-type MMTV integration site family, member 16 | WNT16 | 7q31 | NM_016087.1 | Hs.272375 | 91.3 | Above | 2547.6 |
| 19 | 232289_at | Homo sapiens cDNA FLJ14167 fis, clone NT2RP2001214 | | | BF237871 | Hs.200629 | 91.3 | Above | 143.0 |
| 20 | 224022_x_at | wingless-type MMTV integration site family, member 16 | WNT16 | 7q31 | AF169963.1 | Hs.272375 | 91.3 | Above | 565.4 |
| 21 | 235666_at | ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens] [H.sapiens] | | | AA903473 | Hs.153717 | 91.3 | Above | 649.9 |
| 22 | 225235_at | hypothetical protein MGC14859 | MGC14859 | 5q35.3 | AW007710 | Hs.57100 | 84.7 | Above | 3.7 |
| 23 | 224733_at | chemokine-like factor super family 3 | CKLFSF3 | 16q23.1 | AL574900 | Hs.7773 | 84.7 | Below | 41.6 |
| 24 | 211913_s_at | | | | L08961.1 | | 83.1 | Above | 42.3 |
| 25 | 219551_at | uncharacterized bone marrow protein BM040 | BM040 | 3q21.1 | NM_018456.1 | Hs.26892 | 83.1 | Above | 3.0 |
| 26 | 223693_s_at | hypothetical protein FLJ10324 | FLJ10324 | 7p22 | AL136731.1 | Hs.157158 | 83.1 | Above | 65.0 |
| 27 | 200600_at | moesin | MSN | Xq11.2-q12 | NM_002444.1 | Hs.170328 | 82.7 | Below | 2.2 |
| 28 | 213909_at | Homo sapiens cDNA FLJ12280 | | | AU147799 | Hs.288467 | 82.7 | Above | 12.4 |
| 29 | 235911_at | ESTs, Weakly similar to PIHUB6 salivary proline-rich protein precursor PRB1 (large allele) - human [H.sapiens] | | | AI885815 | Hs.184727 | 82.7 | Above | 36.3 |
| 30 | 243533_x_at | ESTs | | | H09663 | Hs.106490 | 82.7 | Above | 24.4 |
| 31 | 204114_at | nidogen 2 (osteonidogen) | NID2 | 14q21-q22 | NM_007361.1 | Hs.82733 | 82.3 | Above | 15.1 |

| | | | | | | | | | |
|----|-------------|--|--------------|---------------|-------------|-----------|------|-------|--------|
| 32 | 202615_at | Homo sapiens mRNA; cDNA DKFZp686D0521 (from clone DKFZp686D0521) | | | BF222895 | Hs.356786 | 78.4 | Below | 6.1 |
| 33 | 212789_at | KIAA0056 protein | KIAA0056 | 11q25 | AI796581 | Hs.13421 | 78.4 | Above | 3.9 |
| 34 | 228580_at | serine protease HTRA3 | HTRA3 | 4p16.1 | AI828007 | Hs.60440 | 76.1 | Above | 3.7 |
| 35 | 202796_at | synaptopodin | KIAA1029 | 5q33.1 | NM_007286.1 | Hs.5307 | 76.1 | Above | 51.9 |
| 36 | 218640_s_at | phafin 2 | FLJ13187 | 8q21.3 | NM_024613.1 | Hs.29724 | 76.1 | Above | 3.1 |
| 37 | 231040_at | ESTs | | | AW512988 | Hs.184780 | 76.1 | Above | 16.3 |
| 38 | 235099_at | ESTs, Weakly similar to PLLP_HUMAN Plasmolipin [H.sapiens] | | | AW080832 | Hs.154986 | 76.1 | Above | 6.7 |
| 39 | 205173_x_at | CD58 antigen, (lymphocyte function-associated antigen 3) | CD58 | 1p13 | NM_001779.1 | Hs.75626 | 74.6 | Above | 2.4 |
| 40 | 211744_s_at | CD58 antigen, (lymphocyte function-associated antigen 3) | CD58 | 1p13 | BC005930.1 | Hs.75626 | 74.6 | Above | 2.4 |
| 41 | 213358_at | KIAA0802 protein | KIAA0802 | 18p11.21 | AB018345.1 | Hs.27657 | 74.6 | Above | 12.6 |
| 42 | 201889_at | family with sequence similarity 3, member C | FAM3C | 7q22.1-q31.1 | NM_014888.1 | Hs.29882 | 74.6 | Above | 4.6 |
| 43 | 222699_s_at | phafin 2 | FLJ13187 | 8q21.3 | BF439250 | Hs.29724 | 74.6 | Above | 3.5 |
| 44 | 238778_at | Homo sapiens mRNA; cDNA DKFZp451L157 (from clone DKFZp451L157) | | | AI244661 | Hs.350684 | 74.6 | Above | 23.3 |
| 45 | 205769_at | solute carrier family 27 (fatty acid transporter), member 2 | SLC27A2 | 15q21.2 | NM_003645.1 | Hs.11729 | 74.4 | Above | 80.6 |
| 46 | 210786_s_at | Friend leukemia virus integration 1 | FLI1 | 11q24.1-q24.3 | M93255.1 | Hs.108043 | 74.4 | Above | 2.2 |
| 47 | 212985_at | cDNA DKFZp434E033 | DKFZp434E033 | | BF115739 | Hs.15740 | 74.4 | Above | 7.0 |
| 48 | 221669_s_at | acyl-Coenzyme A dehydrogenase family, member 8 | ACAD8 | 11q25 | BC001964.1 | Hs.14791 | 74.4 | Above | 2.6 |
| 49 | 227441_s_at | E2a-Pbx1-associated protein | EB-1 | 12 | AW005572 | Hs.372732 | 74.4 | Above | 1139.4 |
| 50 | 234261_at | Homo sapiens mRNA; cDNA DKFZp761M10121 (from clone DKFZp761M10121) | | | AL137313.1 | Hs.306449 | 74.4 | Above | 960.8 |
| 51 | 244565_at | ESTs | | | AI685824 | Hs.171068 | 74.4 | Above | 7.4 |
| 52 | 209760_at | KIAA0922 protein | KIAA0922 | 4q31.23 | AL136932.1 | Hs.37892 | 73.0 | Above | 2.9 |
| 53 | 218283_at | synovial sarcoma translocation gene on chromosome 18-like 2 | SS18L2 | 3p21 | NM_016305.1 | Hs.9774 | 72.8 | Above | 1.6 |
| 54 | 209558_s_at | huntingtin interacting protein 12 | HIP12 | 12q24 | AB013384.1 | Hs.96731 | 70.0 | Above | 23.6 |
| 55 | 213005_s_at | KIAA0172 protein | KIAA0172 | 9p24.3 | D79994.1 | Hs.77546 | 70.0 | Above | 8.3 |
| 56 | 236854_at | Homo sapiens mRNA; cDNA DKFZp667F0617 (from clone DKFZp667F0617) | | | AA743694 | Hs.48984 | 70.0 | Above | 12.5 |
| 57 | 203435_s_at | membrane metallo-endopeptidase (neutral endopeptidase, enkephalinase, CALLA, CD10) | MME | 3q25.1-q25.2 | NM_007287.1 | Hs.1298 | 68.3 | Below | 2.2 |

| | | | | | | | | | |
|----|-------------|---|----------|---------------|-------------|-----------|------|-------|--------|
| 58 | 204774_at | ecotropic viral integration site 2A | EVI2A | 17q11.2 | NM_014210.1 | Hs.70499 | 68.1 | Below | 3.0 |
| 59 | 201579_at | FAT tumor suppressor homolog 1 (Drosophila) | FAT | 4q34-q35 | NM_005245.1 | Hs.166994 | 68.1 | Above | 9.8 |
| 60 | 203999_at | synaptotagmin I | SYT1 | 12cen-q21 | NM_005639.1 | Hs.154679 | 68.1 | Above | 3.9 |
| 61 | 200811_at | cold inducible RNA binding protein | CIRBP | 19p13.3 | NM_001280.1 | Hs.119475 | 67.8 | Below | 5.8 |
| 62 | 202106_at | golgi autoantigen, golgin subfamily a, 3 | GOLGA3 | 12q24.33 | NM_005895.1 | Hs.4953 | 67.8 | Above | 3.2 |
| 63 | 233273_at | Homo sapiens cDNA FLJ12010 fis, clone HEMBB1001635 | | | AU146834 | Hs.296684 | 67.8 | Above | 30.4 |
| 64 | 239427_at | ESTs | | | AA131524 | Hs.374124 | 67.8 | Above | 13.6 |
| 65 | 218087_s_at | sorbin and SH3 domain containing 1 | SORBS1 | 10q23.3-q24.1 | NM_015385.1 | Hs.108924 | 66.3 | Above | 24.9 |
| 66 | 201460_at | mitogen-activated protein kinase-activated protein kinase 2 | MAPKAPK2 | 1q32 | AI141802 | Hs.75074 | 66.3 | Above | 2.1 |
| 67 | 202421_at | immunoglobulin superfamily, member 3 | IGSF3 | 1p13 | AB007935.1 | Hs.81234 | 66.3 | Above | 4.3 |
| 68 | 40148_at | amyloid beta (A4) precursor protein-binding, family B, member 2 (Fe65-like) | APBB2 | 4p14 | U62325 | Hs.324125 | 66.2 | Above | 6.2 |
| 69 | 212165_at | hypothetical protein BC013073 | LOC92703 | 1q31.1 | AL133052.1 | Hs.17481 | 66.2 | Above | 1.9 |
| 70 | 226590_at | Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1517766 | | | AA031404 | Hs.349208 | 66.2 | Above | 3.1 |
| 71 | 227440_at | E2a-Pbx1-associated protein | EB-1 | 12 | AW005572 | Hs.372732 | 66.2 | Above | 1168.9 |
| 72 | 203143_s_at | KIAA0040 gene product | KIAA0040 | 1q24-25 | T79953 | Hs.158282 | 64.7 | Above | 2.4 |
| 73 | 228153_at | Homo sapiens mRNA; cDNA DKFZp451E085 (from clone DKFZp451E085) | | | AI953847 | Hs.294092 | 64.7 | Above | 4.1 |
| 74 | 212959_s_at | MGC4170 protein | MGC4170 | 12q23.1 | AK001821.1 | Hs.7041 | 64.6 | Below | 3.0 |
| 75 | 202181_at | KIAA0247 gene product | KIAA0247 | 14q24.1 | NM_014734.1 | Hs.82426 | 63.7 | Above | 1.8 |
| 76 | 204562_at | interferon regulatory factor 4 | IRF4 | 6p25-p23 | NM_002460.1 | Hs.82132 | 63.7 | Above | 4.9 |
| 77 | 218949_s_at | hypothetical protein FLJ10989 | FLJ10989 | 6q21 | NM_018292.1 | Hs.210778 | 63.7 | Above | 3.2 |
| 78 | 219517_at | hypothetical protein FLJ22637 | FLJ22637 | 15q14 | NM_025165.1 | Hs.296178 | 63.7 | Above | 3.4 |
| 79 | 224856_at | FK506 binding protein 5 | FKBP5 | 6p21.3-21.2 | AL122066.1 | Hs.7557 | 63.7 | Below | 5.5 |
| 80 | 225389_at | BTB domain protein BDPL | BDPL | 14q32 | AW149498 | Hs.7367 | 63.7 | Above | 3.2 |
| 81 | 226392_at | Homo sapiens cDNA: FLJ21652 fis, clone COL08582 | | | AI888503 | Hs.98445 | 63.7 | Below | 3.4 |
| 82 | 219667_s_at | hypothetical protein FLJ20706 | BANK | 4q22.2 | NM_017935.1 | Hs.193736 | 62.9 | Below | 5.6 |
| 83 | 212845_at | KIAA1053 protein | KIAA1053 | 14q22.1 | AB028976.1 | Hs.173571 | 61.9 | Above | 14.6 |
| 84 | 213940_s_at | formin-binding protein 17 | FBP17 | 9q34 | AU145053 | Hs.301763 | 61.9 | Below | 12.3 |
| 85 | 219111_s_at | ATP-dependent RNA helicase | MGC2835 | 12q24.11 | NM_024072.1 | Hs.70582 | 61.9 | Above | 1.9 |
| 86 | 219518_s_at | hypothetical protein FLJ22637 | FLJ22637 | 15q14 | NM_025165.1 | Hs.296178 | 61.9 | Above | 5.7 |
| 87 | 220389_at | hypothetical protein FLJ23514 | FLJ23514 | 11q14.1 | NM_021827.1 | Hs.144913 | 61.9 | Above | 6.5 |
| 88 | 47069_at | Rho GTPase activating protein 8 | ARHGAP8 | 22q13.31 | AA533284 | Hs.102336 | 61.9 | Above | 3.4 |
| 89 | 212774_at | zinc finger protein 238 | ZNF238 | 1q44-qter | AJ223321 | Hs.69997 | 61.9 | Above | 1.8 |
| 90 | 201443_s_at | ATPase, H ⁺ transporting, lysosomal interacting protein 2 | ATP6IP2 | Xq21 | AF248966.1 | Hs.183434 | 60.9 | Below | 1.9 |

| | | | | | | | | | |
|-----|-------------|---|---------|---------------|-------------|-----------|------|-------|------|
| 91 | 204836_at | glycine dehydrogenase (decarboxylating; glycine decarboxylase, glycine cleavage system protein P) | GLDC | 9p22 | NM_000170.1 | Hs.380791 | 60.1 | Above | 3.1 |
| 92 | 217732_s_at | integral membrane protein 2B | ITM2B | 13q14.3 | AF092128.1 | Hs.239625 | 60.1 | Below | 2.0 |
| 93 | 201029_s_at | antigen identified by monoclonal antibodies 12E7, F21 and O13 | MIC2 | Xp22.32 | NM_002414.1 | Hs.177543 | 59.8 | Below | 5.7 |
| 94 | 204174_at | arachidonate 5-lipoxygenase-activating protein | ALOX5AP | 13q12 | NM_001629.1 | Hs.100194 | 59.8 | Below | 48.1 |
| 95 | 206255_at | B lymphoid tyrosine kinase | BLK | 8p23-p22 | NM_001715.1 | Hs.2243 | 59.8 | Above | 6.0 |
| 96 | 212873_at | minor histocompatibility antigen HA-1 | HA-1 | 19p13.3 | BE349017 | Hs.196914 | 59.8 | Below | 2.9 |
| 97 | 211963_s_at | actin related protein 2/3 complex, subunit 5, 16kDa | ARPC5 | 1q24.3 | AL516350 | Hs.82425 | 59.6 | Above | 1.7 |
| 98 | 218683_at | polypyrimidine tract binding protein 2 | PTBP2 | 1p22.11-p21.3 | NM_021190.1 | Hs.34956 | 59.6 | Above | 1.8 |
| 99 | 221234_s_at | BTB and CNC homology 1, basic leucine zipper transcription factor 2 | BACH2 | 6q15 | NM_021813.1 | Hs.88414 | 59.6 | Above | 2.7 |
| 100 | 240718_at | ESTs, Highly similar to I38656 lymphoid-restricted membrane protein - human [H.sapiens] | | | AW303384 | Hs.124922 | 59.6 | Above | 11.9 |

Table S11. Top 100 chi-square probe sets selected for Hyperdiploid >50 in parallel format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Unigene Reference | Chi-square value | HD>50 above/below mean | Fold change |
|----|-----------------------|---|---------------|-----------------|--------------------------|--------------------------|-------------------------|----------------------------------|--------------------|
| 1 | 216071_x_at | trinucleotide repeat containing 11 | TNRC11 | Xq13 | AF132033 | Hs.211607 | 91.8 | Above | 1.8 |
| 2 | 218757_s_at | similar to yeast Upf3, variant B | UPF3B | Xq25-q26 | NM_023010.1 | Hs.103832 | 91.3 | Above | 2.3 |
| 3 | 208861_s_at | alpha thalassemia/mental retardation syndrome X-linked (RAD54 homolog, S. cerevisiae) | ATRAX | Xq13.1-q21.1 | U72937.2 | Hs.96264 | 91.3 | Above | 1.7 |
| 4 | 200980_s_at | pyruvate dehydrogenase (lipoamide) alpha 1 | PDHA1 | Xp22.2-p22.1 | NM_000284.1 | Hs.1023 | 83.1 | Above | 1.7 |
| 5 | 202371_at | hypothetical protein FLJ21174 | FLJ21174 | Xq22.1 | NM_024863.1 | Hs.194329 | 83.1 | Above | 3.6 |
| 6 | 205324_s_at | FtsJ homolog 1 (E. coli) | FTSJ1 | Xp11.23 | NM_012280.1 | Hs.23170 | 83.1 | Above | 2.2 |
| 7 | 201899_s_at | ubiquitin-conjugating enzyme E2A (RAD6 homolog) | UBE2A | Xq24-q25 | NM_003336.1 | Hs.80612 | 82.7 | Above | 1.8 |
| 8 | 212846_at | KIAA0179 protein | KIAA0179 | 21q22.3 | D80001.1 | Hs.152629 | 82.7 | Above | 2.0 |
| 9 | 218499_at | Mst3 and SOK1-related kinase | MST4 | Xq26.1 | NM_016542.1 | Hs.23643 | 82.7 | Above | 2.5 |
| 10 | 221689_s_at | Down syndrome critical region gene 5 | DSCR5 | 21q22.2 | AB035745.1 | Hs.66493 | 82.7 | Above | 2.2 |
| 11 | 219038_at | hypothetical protein FLJ11565 | FLJ11565 | Xq22.2 | NM_024657.1 | Hs.61763 | 78.8 | Above | 6.8 |
| 12 | 208598_s_at | Upstream regulatory element binding protein 1 | UREB1 | Xp11.22 | NM_005703.2 | | 78.4 | Above | 1.6 |

| | | | | | | | | | |
|----|-------------|---|-----------|---------------|-------------|-----------|------|-------|------|
| 13 | 201100_s_at | ubiquitin specific protease 9, X chromosome (fat facets-like Drosophila) | USP9X | Xp11.4 | NM_004652.2 | Hs.77578 | 76.1 | Above | 1.7 |
| 14 | 201132_at | heterogeneous nuclear ribonucleoprotein H2 (H') | HNRPH2 | Xq22 | NM_019597.1 | Hs.278857 | 74.6 | Above | 2.0 |
| 15 | 201898_s_at | ubiquitin-conjugating enzyme E2A (RAD6 homolog) | UBE2A | Xq24-q25 | AI126625 | Hs.80612 | 74.6 | Above | 1.6 |
| 16 | 213289_at | Homo sapiens cDNA FLJ39590 fis, clone SKNMC1000079 | | | BE221922 | Hs.356269 | 74.6 | Above | 3.8 |
| 17 | 219485_s_at | proteasome (prosome, macropain) 26S subunit, non-ATPase, 10 | PSMD10 | Xq22.3 | NM_002814.1 | Hs.7756 | 74.6 | Above | 2.4 |
| 18 | 218878_s_at | sirtuin silent mating type information regulation 2 homolog 1 (S. cerevisiae) | SIRT1 | 10q22.2 | NM_012238.3 | Hs.31176 | 74.4 | Above | 1.4 |
| 19 | 228005_at | ESTs, Weakly similar to T45117 hU1-70K protein (286 AA) [imported] - human (fragment) [H.sapiens] | | | BE677308 | Hs.96716 | 74.4 | Above | 3.8 |
| 20 | 201136_at | proteolipid protein 2 (colonic epithelium-enriched) | PLP2 | Xp11.23 | NM_002668.1 | Hs.77422 | 73.0 | Above | 3.4 |
| 21 | 202214_s_at | cullin 4B | CUL4B | Xq23 | NM_003588.1 | Hs.155976 | 73.0 | Above | 1.9 |
| 22 | 242794_at | ESTs | | | AI569476 | Hs.177135 | 70.0 | Above | 3.2 |
| 23 | 211342_x_at | trinucleotide repeat containing 11 (THR-associated protein, 230 kDa subunit) | TNRC11 | Xq13 | BC004354.1 | | 68.1 | Above | 1.8 |
| 24 | 212419_at | Homo sapiens, similar to Y43E12A.2.p, clone MGC:33537 IMAGE:4821347, mRNA, complete cds | | | AL049949.1 | Hs.28264 | 68.1 | Above | 13.0 |
| 25 | 207785_s_at | H-2K binding factor-2 | LOC51580 | 9 | NM_015874.1 | Hs.347340 | 67.8 | Above | 1.8 |
| 26 | 212460_at | hypothetical protein MGC24447 | MGC24447 | 14 | BE738425 | Hs.353161 | 67.8 | Above | 2.2 |
| 27 | 203745_at | holocytochrome c synthase (cytochrome c heme-lyase) | HCCS | Xp22.3 | AI801013 | Hs.211571 | 67.8 | Above | 2.0 |
| 28 | 209565_at | zinc finger protein 183 (RING finger, C3HC4 type) | ZNF183 | Xq25-q26 | BC000832.1 | Hs.64794 | 67.8 | Above | 2.2 |
| 29 | 201099_at | ubiquitin specific protease 9, X chromosome (fat facets-like Drosophila) | USP9X | Xp11.4 | AA824386 | Hs.77578 | 66.3 | Above | 1.6 |
| 30 | 203974_at | DNA segment, numerous copies, expressed probes (GS1 gene) | DXF68S1E | Xp22.32 | NM_012080.1 | Hs.78991 | 66.3 | Above | 4.0 |
| 31 | 215884_s_at | ubiquilin 2 | UBQLN2 | Xp11.23-p11.1 | AK001029.1 | Hs.4552 | 66.2 | Above | 1.9 |
| 32 | 200642_at | superoxide dismutase 1, soluble | SOD1 | 21q22.11 | NM_000454.1 | Hs.75428 | 64.7 | Above | 2.3 |
| 33 | 200738_s_at | phosphoglycerate kinase 1 | PGK1 | Xq13 | NM_000291.1 | Hs.78771 | 64.7 | Above | 1.8 |
| 34 | 201312_s_at | SH3 domain binding glutamic acid-rich protein like | SH3BGRL | Xq13.3 | NM_003022.1 | Hs.14368 | 64.7 | Above | 1.6 |
| 35 | 226875_at | hypothetical protein FLJ32122 | FLJ32122 | Xq24 | AI742838 | Hs.107513 | 64.6 | Above | 2.3 |
| 36 | 228057_at | similar to Smhs1 protein | LOC115265 | 4q22.1 | AA528140 | Hs.107515 | 63.7 | Above | 6.7 |

| | | | | | | | | | |
|----|-------------|--|----------|---------------|-------------|-----------|------|-------|------|
| 37 | 232974_at | Homo sapiens cDNA FLJ12417 fis, clone MAMMA1003039 | | | AU148256 | Hs.226469 | 63.7 | Above | 3.1 |
| 38 | 232034_at | Homo sapiens mRNA; cDNA DKFZp564N0763 (from clone DKFZp564N0763) | | | AL117607.1 | Hs.175563 | 62.1 | Above | 6.1 |
| 39 | 200944_s_at | high-mobility group nucleosome binding domain 1 | HMGN1 | 21q22.2 | NM_004965.1 | Hs.251064 | 61.9 | Above | 1.7 |
| 40 | 201443_s_at | ATPase, H ⁺ transporting, lysosomal interacting protein 2 | ATP6IP2 | Xq21 | AF248966.1 | Hs.183434 | 61.9 | Above | 1.9 |
| 41 | 226333_at | interleukin 6 receptor | IL6R | 1q21 | AV700030 | Hs.193400 | 61.9 | Above | 8.7 |
| 42 | 203585_at | zinc finger protein 185 (LIM domain) | ZNF185 | Xq28 | NM_007150.1 | Hs.16622 | 60.9 | Above | 10.7 |
| 43 | 206846_s_at | histone deacetylase 6 | HDAC6 | Xp11.23 | NM_006044.2 | Hs.6764 | 60.1 | Above | 1.5 |
| 44 | 216095_x_at | myotubularin related protein 1 | MTMR1 | Xq28 | AF057354.1 | Hs.372428 | 60.1 | Above | 1.8 |
| 45 | 226039_at | Mannosyl (alpha-1,3)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase | MGAT4A | 2q11.2 | AW006441 | Hs.24210 | 60.1 | Above | 3.0 |
| 46 | 227210_at | Homo sapiens cDNA FLJ32568 fis, clone SPLEN2000098 | | | T65020 | Hs.12699 | 60.1 | Above | 3.6 |
| 47 | 236172_at | ESTs | | | AW206817 | Hs.158186 | 60.1 | Above | 2.4 |
| 48 | 214735_at | phosphoinositide-binding protein PIP3-E | PIP3-E | 6q25.2 | AW166711 | Hs.185140 | 59.8 | Above | 6.4 |
| 49 | 206147_x_at | sex comb on midleg-like 2 (Drosophila) | SCML2 | Xp22 | NM_006089.1 | Hs.171558 | 59.6 | Above | 2.2 |
| 50 | 209022_at | stromal antigen 2 | STAG2 | Xq25 | AK026678.1 | Hs.8217 | 59.6 | Above | 1.7 |
| 51 | 213282_at | Homo sapiens cDNA FLJ39590 fis, clone SKNMC1000079 | | | BE221922 | Hs.356269 | 59.6 | Above | 2.2 |
| 52 | 218582_at | hypothetical protein FLJ20445 | FLJ20445 | 10q23.33 | NM_017824.1 | Hs.343748 | 59.6 | Above | 1.6 |
| 53 | 218694_at | ALEX1 protein | ALEX1 | Xq21.33-q22.2 | NM_016608.1 | Hs.9728 | 59.6 | Above | 2.8 |
| 54 | 226785_at | Homo sapiens cDNA FLJ36516 fis, clone TRACH2001898 | | | BF475862 | Hs.88252 | 59.6 | Above | 1.9 |
| 55 | 200655_s_at | calmodulin 1 (phosphorylase kinase, delta) | CALM1 | 14q24-q31 | NM_006888.1 | Hs.177656 | 58.7 | Above | 1.7 |
| 56 | 210176_at | toll-like receptor 1 | TLR1 | 4p14 | AL050262.1 | Hs.2474 | 58.2 | Above | 2.1 |
| 57 | 202711_at | ephrin-B1 | EFNB1 | Xq12 | NM_004429.1 | Hs.144700 | 58.2 | Above | 14.9 |
| 58 | 204446_s_at | arachidonate 5-lipoxygenase | ALOX5 | 10q11.2 | NM_000698.1 | Hs.89499 | 58.2 | Above | 4.2 |
| 59 | 219297_at | similar to rab11-binding protein | FLJ11116 | Xq24 | NM_019045.1 | Hs.98510 | 58.2 | Above | 2.4 |
| 60 | 201923_at | peroxiredoxin 4 | PRDX4 | Xp22.13 | NM_006406.1 | Hs.83383 | 58.2 | Above | 1.9 |
| 61 | 212929_s_at | KIAA0592 protein | KIAA0592 | 10q11.21 | W68158 | Hs.13273 | 58.2 | Above | 5.0 |
| 62 | 204045_at | transcription elongation factor A (SII)-like 1 | TCEAL1 | Xq22.1 | NM_004780.1 | Hs.95243 | 56.8 | Above | 2.1 |
| 63 | 208117_s_at | hypothetical protein FLJ12525 | FLJ12525 | Xq12-q13 | NM_031206.1 | Hs.321618 | 56.8 | Above | 2.2 |
| 64 | 209620_s_at | ATP-binding cassette, sub-family B (MDR/TAP), member 7 | ABCB7 | Xq12-q13 | AB005289.1 | Hs.125856 | 56.8 | Above | 2.1 |
| 65 | 209679_s_at | hypothetical protein from clone 643 | LOC57228 | 12q11 | BC003379.1 | Hs.206501 | 56.8 | Above | 5.6 |
| 66 | 217954_s_at | PHD finger protein 3 | PHF3 | | NM_015153.1 | Hs.78893 | 56.8 | Above | 1.5 |
| 67 | 218414_s_at | LIS1-interacting protein NUDE1, rat homolog | NUDE1 | 16p13.11 | NM_017668.1 | Hs.263925 | 56.8 | Above | 2.0 |

| | | | | | | | | | |
|----|-------------|---|-----------|--------------|-------------|-----------|------|-------|-----|
| 68 | 226037_s_at | TAF9-like RNA polymerase II, TATA box binding protein (TBP)-associated factor, 31kDa | TAF9L | Xq13.1-q21.1 | AL049589 | Hs.171723 | 56.8 | Above | 1.9 |
| 69 | 227279_at | hypothetical protein MGC15737 | MGC15737 | Xq22.1 | AA847654 | Hs.39122 | 56.8 | Above | 5.6 |
| 70 | 200600_at | Moesin (membrane organizing extension spike protein) | MSN | Xq11.2-q12 | NM_002444.1 | Hs.170328 | 56.2 | Above | 1.9 |
| 71 | 200701_at | Niemann-Pick disease, type C2 | NPC2 | 14q24.3 | NM_006432.1 | Hs.119529 | 55.9 | Above | 1.7 |
| 72 | 200737_at | phosphoglycerate kinase 1 | PGK1 | Xq13 | NM_000291.1 | Hs.78771 | 55.9 | Above | 1.8 |
| 73 | 201642_at | interferon gamma receptor 2 (interferon gamma transducer 1) | IFNGR2 | 21q22.11 | NM_005534.1 | Hs.177559 | 55.9 | Above | 2.5 |
| 74 | 214527_s_at | polyglutamine binding protein 1 | PQBP1 | Xp11.23 | AB041836.1 | Hs.30570 | 55.9 | Above | 1.7 |
| 75 | 223082_at | SH3-domain kinase binding protein 1 | SH3KBP1 | Xp22.1-p21.3 | AF230904.1 | Hs.153260 | 55.9 | Above | 2.0 |
| 76 | 225182_at | Homo sapiens EST from clone 491476, full insert | | | AL355685.1 | Hs.9042 | 55.9 | Above | 2.1 |
| 77 | 226760_at | Homo sapiens cDNA FLJ32174 fis, clone PLACE6001064 | | | BF666325 | Hs.297007 | 55.9 | Above | 2.2 |
| 78 | 204572_s_at | protein (peptidyl-prolyl cis/trans isomerase) NIMA-interacting, 4 (parvulin) | PIN4 | Xq13 | NM_006223.1 | Hs.11774 | 54.8 | Above | 2.9 |
| 79 | 201092_at | retinoblastoma binding protein 7 | RBBP7 | Xp22.31 | NM_002893.2 | Hs.31314 | 54.0 | Above | 1.6 |
| 80 | 204690_at | syntaxin 8 | STX8 | 17p12 | NM_004853.1 | Hs.380938 | 54.0 | Above | 1.4 |
| 81 | 206016_at | JM1 protein | JM1 | Xp11.23 | NM_014008.1 | Hs.26333 | 54.0 | Above | 2.3 |
| 82 | 218021_at | peroxisomal short-chain alcohol dehydrogenase | humNRDR | 14q11.2 | NM_021004.1 | Hs.6318 | 54.0 | Above | 4.0 |
| 83 | 221808_at | RAB9A, member RAS oncogene family | RAB9A | Xp22.2 | NM_004251.1 | Hs.330994 | 54.0 | Above | 1.8 |
| 84 | 223497_at | KIAA1411 protein | KIAA1411 | 6q12-q13 | AL136820.1 | Hs.107287 | 54.0 | Above | 3.8 |
| 85 | 226335_at | Homo sapiens cDNA FLJ12807 fis, clone NT2RP2002316 | | | BG498334 | Hs.188361 | 54.0 | Above | 2.0 |
| 86 | 201311_s_at | SH3 domain binding glutamic acid-rich protein like | SH3BGRL | Xq13.3 | AL515318 | Hs.14368 | 53.0 | Above | 1.6 |
| 87 | 202829_s_at | synaptobrevin-like 1 | SYBL1 | Xq28 | NM_005638.1 | Hs.24167 | 53.0 | Above | 1.5 |
| 88 | 209370_s_at | SH3-domain binding protein 2 | SH3BP2 | 4p16.3 | AB000462.1 | Hs.167679 | 53.0 | Above | 3.2 |
| 89 | 46323_at | Ca ²⁺ -dependent endoplasmic reticulum nucleoside diphosphatase | SHAPY | 17q25.3 | AL120741 | Hs.8859 | 52.3 | Above | 1.7 |
| 90 | 200658_s_at | prohibitin | PHB | 17q21 | AL560017 | Hs.75323 | 52.3 | Above | 2.0 |
| 91 | 202974_at | membrane protein, palmitoylated 1, 55kDa | MPP1 | Xq28 | NM_002436.2 | Hs.1861 | 52.3 | Above | 2.2 |
| 92 | 221188_s_at | cell death-inducing DFFA-like effector b | CIDEB | 14q11.2 | NM_014430.1 | Hs.288835 | 52.3 | Above | 2.4 |
| 93 | 203776_at | T54 protein | T54 | Xp11.23 | NM_015698.1 | Hs.100391 | 52.3 | Above | 2.0 |
| 94 | 211974_x_at | CD59 antigen p18-20 (antigen identified by monoclonal antibodies 16.3A5, EJ16, EJ30, EL32 and G344) | CD59 | 11p13 | AL513759 | Hs.278573 | 52.3 | Above | 1.7 |
| 95 | 213000_at | nuclear matrix protein NXP-2 | NXP-2 | 21q22.13 | AP000693 | Hs.70359 | 52.3 | Above | 1.9 |
| 96 | 225361_x_at | similar to hypothetical protein MGC17347 | LOC159090 | Xq26.2 | AI341165 | Hs.353200 | 52.3 | Above | 1.8 |
| 97 | 226905_at | Homo sapiens, clone IMAGE:4564684, mRNA, partial cds | | | BG036514 | Hs.345588 | 52.3 | Above | 4.0 |

partial cds

| | | | | | | | | | |
|-----|-----------|--|----------|--------|------------|-----------|------|-------|-----|
| 98 | 223294_at | hypothetical protein LOC51260 | LOC51260 | Xq13.1 | BC001220.1 | Hs.128764 | 52.3 | Above | 1.8 |
| 99 | 226810_at | Homo sapiens mRNA; cDNA DKFZp761M0111 (from clone DKFZp761M0111) | | | BE500942 | Hs.13299 | 52.2 | Above | 2.5 |
| 100 | 238469_at | hypothetical protein FLJ21079 | FLJ21079 | 6q14.1 | BE620374 | Hs.16512 | 51.8 | Above | 2.2 |

Table S12. Top 100 chi-square probe sets selected for *MLL* in parallel format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Unigene Reference | Chi-square value | MLL above/below mean | Fold change |
|----|----------------|--|-----------|---------------|-------------------|-------------------|------------------|----------------------|-------------|
| 1 | 226939_at | Homo sapiens cDNA FLJ37247 fis, clone BRAMY2006397 | | | AI202327 | Hs.44833 | 92.2 | Above | 7 |
| 2 | 219463_at | chromosome 20 open reading frame 103 | C20orf103 | 20p12 | NM_012261.1 | Hs.22920 | 86.2 | Above | 24 |
| 3 | 204069_at | Meis1, myeloid ecotropic viral integration site 1 homolog | MEIS1 | 2p14-p13 | NM_002398.1 | Hs.170177 | 84.9 | Above | 73 |
| 4 | 203837_at | mitogen-activated protein kinase kinase kinase 5 | MAP3K5 | 6q22.33 | NM_005923.2 | Hs.151988 | 77.3 | Above | 4 |
| 5 | 226415_at | KIAA1576 protein | KIAA1576 | 16q22.1 | AA156723 | Hs.22975 | 77.3 | Above | 40 |
| 6 | 222409_at | coronin, actin binding protein, 1C | CORO1C | 12q24.1 | AL162070.1 | Hs.17377 | 71.1 | Above | 5 |
| 7 | 242172_at | ESTs | | | N50406 | Hs.191349 | 71.1 | Above | 33 |
| 8 | 235879_at | ESTs | | | AI697540 | Hs.310286 | 70.0 | Above | 4 |
| 9 | 231899_at | KIAA1726 protein | KIAA1726 | 11q23.1 | AB051513.1 | Hs.164719 | 65.6 | Above | 33 |
| 10 | 212386_at | Homo sapiens cDNA FLJ11918 fis, clone HEMBB1000272 | | | AK021980.1 | Hs.289068 | 64.5 | Below | 3 |
| 11 | 201153_s_at | muscleblind-like (Drosophila) | MBNL | 3q25 | NM_021038.1 | Hs.28578 | 63.4 | Above | 2 |
| 12 | 218217_at | likely homolog of rat and mouse retinoid-inducible serine carboxypeptidase | RISC | 17q23.2 | NM_021626.1 | Hs.106747 | 63.4 | Above | 3 |
| 13 | 201875_s_at | hypothetical protein FLJ21047 | FLJ21047 | 1q23.2 | NM_024569.1 | Hs.14891 | 63.0 | Above | 2 |
| 14 | 205717_x_at | protocadherin gamma subfamily C, 3 | PCDHGC3 | 5q31 | NM_002588.1 | Hs.284180 | 63.0 | Above | 5 |
| 15 | 205821_at | DNA segment on chromosome 12 (unique) 2489 expressed sequence | D12S2489E | 12p13.2-p12.3 | NM_007360.1 | Hs.74085 | 63.0 | Above | 9 |
| 16 | 205899_at | cyclin A1 | CCNA1 | 13q12.3-q13 | NM_003914.1 | Hs.79378 | 63.0 | Above | 104 |
| 17 | 209079_x_at | protocadherin gamma subfamily A, 1 | PCDHGA1 | 5q31 | AF152318.1 | Hs.333404 | 63.0 | Above | 5 |
| 18 | 210660_at | leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 1 | LILRA1 | 19q13.4 | AF025529.1 | Hs.166156 | 63.0 | Above | 532 |
| 19 | 211066_x_at | protocadherin gamma subfamily C, 3 | PCDHGC3 | 5q31 | BC006439.1 | Hs.284180 | 63.0 | Above | 5 |

| | | | | | | | | | |
|----|-------------|--|-----------|---------------|-------------|-----------|------|-------|----|
| 20 | 215836_s_at | protocadherin gamma subfamily B, 7 | PCDHGB7 | 5q31 | AK026188.1 | Hs.256783 | 63.0 | Above | 4 |
| 21 | 212135_s_at | calcium transporting ATPase plasma membrane protein. | ATP2B4 | | AW517686 | Hs.356388 | 62.6 | Below | 2 |
| 22 | 226668_at | hypothetical protein FLJ36175 | FLJ36175 | | W80623 | Hs.20848 | 62.4 | Above | 2 |
| 23 | 201152_s_at | muscleblind-like (Drosophila) | MBNL | 3q25 | NM_021038.1 | Hs.28578 | 60.7 | Above | 2 |
| 24 | 218847_at | IGF-II mRNA-binding protein 2 | IMP-2 | 3q28 | NM_006548.1 | Hs.30299 | 60.7 | Above | 23 |
| 25 | 238712_at | ESTs | | | BF801735 | Hs.181574 | 60.7 | Above | 3 |
| 26 | 234032_at | | | | AF119847.1 | | 60.1 | Above | 4 |
| 27 | 218764_at | hypothetical protein MGC5363 | MGC5363 | 14q22.1-q22.3 | NM_024064.1 | Hs.1880 | 59.3 | Below | 8 |
| 28 | 203345_s_at | likely ortholog of mouse metal response element binding transcription factor 2 | M96 | 1p22.1 | AI566096 | Hs.31016 | 58.5 | Below | 2 |
| 29 | 203836_s_at | mitogen-activated protein kinase kinase kinase 5 | MAP3K5 | 6q22.33 | D84476.1 | Hs.151988 | 58.5 | Above | 14 |
| 30 | 227999_at | hypothetical protein BC011630 | LOC170394 | 10q26.3 | AI290476 | Hs.157728 | 58.5 | Above | 17 |
| 31 | 235479_at | ESTs, Moderately similar to KIAA0940 protein [Homo sapiens] [H.sapiens] | | | AI948598 | Hs.156469 | 58.5 | Above | 4 |
| 32 | 220643_s_at | Fas apoptotic inhibitory molecule | FAIM | 3q23 | NM_018147.1 | Hs.173438 | 58.2 | Above | 3 |
| 33 | 201151_s_at | muscleblind-like (Drosophila) | MBNL | 3q25 | NM_021038.1 | Hs.28578 | 56.8 | Above | 3 |
| 34 | 205668_at | lymphocyte antigen 75 | LY75 | 2q24 | NM_002349.1 | Hs.153563 | 56.8 | Above | 2 |
| 35 | 52164_at | chromosome 11 open reading frame 24 | C11orf24 | 11q13 | AA065185 | Hs.303025 | 56.4 | Above | 2 |
| 36 | 56256_at | CGI-40 protein | LOC51092 | 11q23.3 | AA150165 | Hs.33724 | 56.4 | Above | 2 |
| 37 | 228855_at | nudix (nucleoside diphosphate linked moiety X)-type motif 7 | NUDT7 | | AI927964 | Hs.115920 | 56.4 | Above | 6 |
| 38 | 235173_at | muscleblind-like (Drosophila) | MBNL | 3q25 | AA093668 | Hs.28578 | 56.0 | Above | 41 |
| 39 | 241391_at | ESTs, Moderately similar to hypothetical protein FLJ20378 [Homo sapiens] [H.sapiens] | | | AA654772 | Hs.186564 | 56.0 | Above | 3 |
| 40 | 201105_at | lectin, galactoside-binding, soluble, 1 (galectin 1) | LGALS1 | 22q13.1 | NM_002305.2 | Hs.227751 | 55.9 | Above | 15 |
| 41 | 202975_s_at | Rho-related BTB domain containing 3 | RHOBTB3 | 5q21.2 | N21138 | Hs.10432 | 55.9 | Above | 6 |
| 42 | 236558_at | ESTs | | | AA699809 | Hs.189900 | 55.9 | Above | 18 |
| 43 | 225563_at | Homo sapiens cDNA FLJ33988 fis, clone DFNES2006346, weakly similar to PAB-DEPENDENT POLY(A)-SPECIFIC RIBONUCLEASE SUBUNIT PAN3 (EC 3.1.13.4) | | | AI970788 | Hs.190153 | 55.7 | Above | 3 |
| 44 | 221286_s_at | proapoptotic caspase adaptor protein | FLJ32987 | 5q23-5q31 | NM_016459.1 | Hs.122492 | 54.1 | Below | 4 |
| 45 | 203375_s_at | tripeptidyl peptidase II | TPP2 | 13q32-q33 | NM_003291.1 | Hs.1117 | 52.9 | Above | 2 |
| 46 | 215000_s_at | fasciculation and elongation protein zeta 2 (zygin II) | FEZ2 | 2p21 | AL117593.1 | Hs.103419 | 52.9 | Above | 2 |
| 47 | 209191_at | tubulin beta-5 | TUBB-5 | | BC002654.1 | Hs.274398 | 52.9 | Above | 6 |
| 48 | 225202_at | Rho-related BTB domain containing 3 | RHOBTB3 | 5q21.2 | BE620739 | Hs.10432 | 52.5 | Above | 6 |
| 49 | 225232_at | phosphatidylinositol-3 phosphate 3-phosphatase | 3PAP | 5p13.2 | AA524700 | Hs.93872 | 51.6 | Above | 2 |

| | | | | | | | | | | |
|----|-------------|--|---------------|-------------|-------------|-----------|------|-------|----|--|
| | | adaptor subunit | | | | | | | | |
| 50 | 227611_at | hypothetical protein FLJ25005 | FLJ25005 | 15q26.3 | AA442856 | Hs.181426 | 51.6 | Above | 3 | |
| 51 | 212136_at | calcium transporting ATPase plasma membrane protein. | ATP2B4 | | AW517686 | Hs.356388 | 50.2 | Below | 2 | |
| 52 | 215925_s_at | CD72 antigen | CD72 | 9p11.2 | AF283777.2 | Hs.116481 | 50.2 | Above | 3 | |
| 53 | 221676_s_at | coronin, actin binding protein, 1C | CORO1C | 12q24.1 | BC002342.1 | Hs.17377 | 50.2 | Above | 4 | |
| 54 | 212660_at | KIAA0239 protein | KIAA0239 | 5q31.1 | AI735639 | Hs.9729 | 49.6 | Below | 2 | |
| 55 | 234987_at | SAM domain and HD domain 1 | SAMHD1 | 20pter-q12 | AV715309 | Hs.23889 | 49.4 | Above | 4 | |
| 56 | 205550_s_at | brain and reproductive organ-expressed (TNFRSF1A modulator) | BRE | 2p23.3 | NM_004899.1 | Hs.80426 | 49.3 | Above | 2 | |
| 57 | 210993_s_at | MAD, mothers against decapentaplegic homolog 1 (Drosophila) | MADH1 | 4q28 | U54826.1 | Hs.79067 | 49.3 | Below | 96 | |
| 58 | 205078_at | phosphatidylinositol glycan, class F | PIGF | 2p21-p16 | NM_002643.1 | Hs.348397 | 49.3 | Above | 2 | |
| 59 | 212174_at | adenylate kinase 2 | AK2 | 1p34 | AK023758.1 | Hs.171811 | 49.3 | Above | 2 | |
| 60 | 241681_at | ESTs | | | AW296451 | Hs.24605 | 49.3 | Above | 2 | |
| 61 | 203820_s_at | IGF-II mRNA-binding protein 3 | KOC1 | 7p11 | NM_006547.1 | Hs.79440 | 49.0 | Above | 3 | |
| 62 | 209168_at | Homo sapiens cDNA FLJ38338 fis, clone FCBBF3027104, highly similar to Mus musculus proteolipid M6B isoform alpha-beta-TMD-omega (M6B) mRNA | | | AF016004.1 | Hs.379090 | 49.0 | Above | 4 | |
| 63 | 213566_at | ribonuclease, RNase A family, k6 | RNASE6 | 14q11.1 | NM_005615.1 | Hs.23262 | 49.0 | Above | 6 | |
| 64 | 224622_at | KIAA1322 protein | KIAA1322 | 4p16.2 | AB037743.1 | Hs.72242 | 49.0 | Above | 2 | |
| 65 | 236921_at | ESTs | | | BE504716 | Hs.372166 | 49.0 | Above | 3 | |
| 66 | 239278_at | ESTs, Weakly similar to JC5238 galactosylceramide-like protein, GCP - human [H.sapiens] | | | AI471969 | Hs.182606 | 49.0 | Above | 3 | |
| 67 | 226459_at | Homo sapiens gastric cancer-related protein GCYS-20 (gcys-20) mRNA, complete cds | | | AW575754 | Hs.86437 | 47.7 | Above | 2 | |
| 68 | 239369_at | Homo sapiens cDNA FLJ33328 fis, clone BRACE1000051, weakly similar to KERATIN, ULTRA HIGH-SULFUR MATRIX PROTEIN | | | AW966156 | Hs.323991 | 47.3 | Above | 5 | |
| 69 | 218581_at | hypothetical protein FLJ12816 | FLJ12816 | 14q11.1 | NM_022060.1 | Hs.9175 | 47.1 | Above | 23 | |
| 70 | 214651_s_at | homeo box A9 | HOX A9 | 7p15-p14 | U41813.1 | Hs.380229 | 47.1 | Above | 28 | |
| 71 | 224933_s_at | hypothetical protein DKFZp761F0118 | DKFZp761F0118 | 10q22.1 | AB037801.1 | Hs.6685 | 47.1 | Above | 2 | |
| 72 | 239393_at | ESTs | | | AW510927 | Hs.371883 | 46.5 | Above | 6 | |
| 73 | 209112_at | cyclin-dependent kinase inhibitor 1B (p27, Kip1) | CDKN1B | 12p13.1-p12 | BC001971.1 | Hs.238990 | 46.0 | Above | 2 | |
| 74 | 217707_x_at | SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2 | SMARCA2 | 9p22.3 | AI535683 | Hs.198296 | 46.0 | Above | 2 | |

| | | | | | | | | | |
|-----|-------------|--|----------|---------------|-------------|-----------|------|-------|----|
| 75 | 221830_at | Homo sapiens cDNA FLJ37267 fis, clone BRAMY2011297 | | | AI302106 | Hs.301746 | 46.0 | Above | 2 |
| 76 | 202761_s_at | synaptic nuclei expressed gene 2 | SYNE-2 | 14q23.1-q23.2 | NM_015180.1 | Hs.57749 | 45.9 | Below | 4 |
| 77 | 202976_s_at | Rho-related BTB domain containing 3 | RHOBTB3 | 5q21.2 | NM_014899.1 | Hs.10432 | 45.9 | Above | 12 |
| 78 | 213241_at | plexin C1 | PLXNC1 | | AF035307.1 | Hs.184697 | 45.9 | Above | 3 |
| 79 | 230332_at | ESTs | | | AA872187 | Hs.33665 | 45.4 | Above | 3 |
| 80 | 209905_at | hypothetical protein MGC1934 | MGC1934 | 7p21.1 | AI246769 | Hs.127428 | 44.9 | Above | 31 |
| 81 | 226789_at | ESTs, Weakly similar to A46506 leukocyte activation antigen M6 - human [H.sapiens] | | | W84421 | Hs.356113 | 44.9 | Above | 2 |
| 82 | 204019_s_at | likely ortholog of mouse Sh3 domain YSC-like 1 | SH3YL1 | 2p25.2 | NM_015677.1 | Hs.25213 | 44.0 | Above | 2 |
| 83 | 201874_at | hypothetical protein FLJ21047 | FLJ21047 | 1q23.2 | BF978611 | Hs.14891 | 43.9 | Above | 2 |
| 84 | 238558_at | ESTs | | | AI445833 | Hs.282887 | 43.9 | Above | 4 |
| 85 | 235964_x_at | SAM domain and HD domain 1 | SAMHD1 | 20pter-q12 | AA603344 | Hs.23889 | 43.6 | Above | 3 |
| 86 | 219099_at | chromosome 12 open reading frame 5 | C12orf5 | 12p13.3 | NM_020375.1 | Hs.24792 | 43.3 | Above | 3 |
| 87 | 201064_s_at | poly(A) binding protein, cytoplasmic 4 (inducible form) | PABPC4 | 1p32-p36 | NM_003819.2 | Hs.169900 | 43.3 | Above | 2 |
| 88 | 206471_s_at | plexin C1 | PLXNC1 | 12q23.3 | NM_005761.1 | Hs.286229 | 43.3 | Above | 9 |
| 89 | 208717_at | oxidase (cytochrome c) assembly 1-like | OXA1L | 14q11.2 | BC001669.1 | Hs.151134 | 43.3 | Above | 2 |
| 90 | 210594_x_at | myelin protein zero-like 1 | MPZL1 | 1q23.2 | AF239756.1 | Hs.287832 | 43.3 | Above | 5 |
| 91 | 221675_s_at | choline phosphotransferase 1 | CHPT1 | 12q | AF195624.1 | Hs.171889 | 43.3 | Above | 2 |
| 92 | 225810_at | hypothetical protein FLJ20313 | FLJ20313 | 15q12 | AL572015 | Hs.126721 | 43.3 | Above | 2 |
| 93 | 242260_at | ESTs, Moderately similar to cytokine receptor-like factor 2; cytokine receptor CRL2 precursor [Homo sapiens] [H.sapiens] | | | BG283790 | Hs.214238 | 43.3 | Above | 10 |
| 94 | 224763_at | ESTs | | | AL137450.1 | Hs.374308 | 43.3 | Above | 2 |
| 95 | 227534_at | ESTs, Highly similar to RIKEN cDNA 1110018J18 [Mus musculus] [M.musculus] | | | AL535879 | Hs.42311 | 43.3 | Above | 35 |
| 96 | 235529_x_at | SAM domain and HD domain 1 | SAMHD1 | 20pter-q12 | BF437747 | Hs.23889 | 43.3 | Above | 4 |
| 97 | 212588_at | protein tyrosine phosphatase, receptor type, C | PTPRC | 1q31-q32 | AI809341 | Hs.170121 | 43.2 | Above | 2 |
| 98 | 233931_at | Homo sapiens cDNA FLJ11919 fis, clone HEMBB1000274 | | | AK021981.1 | Hs.193525 | 43.2 | Above | 22 |
| 99 | 240666_at | ESTs | | | AI732568 | Hs.191453 | 43.2 | Above | 4 |
| 100 | 209167_at | glycoprotein M6B | GPM6B | Xp22.2 | AF016004.1 | Hs.5422 | 43.0 | Above | 10 |

Table S13. Top 100 chi-square probe sets selected for T-ALL in parallel format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Unigene Reference | Chi-square value | T-ALL Above/below mean | Fold change |
|----|-----------------------|---|---------------|-----------------|--------------------------|--------------------------|-------------------------|-------------------------------|--------------------|
| 1 | 201137_s_at | major histocompatibility complex, class II, DP beta 1 | HLA-DPB1 | 6p21.3 | NM_002121.1 | Hs.814 | 100.0 | Below | 24.2 |
| 2 | 202113_s_at | sorting nexin 2 | SNX2 | 5q23 | AF043453.1 | Hs.11183 | 100.0 | Below | 4.2 |
| 3 | 202114_at | sorting nexin 2 | SNX2 | 5q23 | NM_003100.1 | Hs.11183 | 100.0 | Below | 4.7 |
| 4 | 203675_at | nucleobindin 2 | NUCB2 | 11p15.1-p14 | NM_005013.1 | Hs.3164 | 100.0 | Above | 3.7 |
| 5 | 204670_x_at | major histocompatibility complex, class II, DR beta 3 | HLA-DRB3 | 6p21.3 | NM_002125.1 | Hs.308026 | 100.0 | Below | 15.3 |
| 6 | 205297_s_at | CD79B antigen (immunoglobulin-associated beta) | CD79B | 17q23 | NM_000626.1 | Hs.89575 | 100.0 | Below | 21.8 |
| 7 | 205456_at | CD3E antigen, epsilon polypeptide (TiT3 complex) | CD3E | 11q23 | NM_000733.1 | Hs.3003 | 100.0 | Above | 20.8 |
| 8 | 206398_s_at | CD19 antigen | CD19 | 16p11.2 | NM_001770.1 | Hs.96023 | 100.0 | Below | 5693.6 |
| 9 | 208306_x_at | major histocompatibility complex, class II, DR beta 4 | HLA-DRB4 | 6p21.3 | NM_021983.2 | Hs.318720 | 100.0 | Below | 9.8 |
| 10 | 208894_at | major histocompatibility complex, class II, DR alpha | HLA-DRA | 6p21.3 | M60334.1 | Hs.76807 | 100.0 | Below | 27.5 |
| 11 | 209312_x_at | major histocompatibility complex, class II, DR beta 1 | HLA-DRB1 | 6p21.3 | U65585.1 | Hs.375570 | 100.0 | Below | 14.6 |
| 12 | 209619_at | CD74 antigen (invariant polypeptide of major histocompatibility complex, class II antigen-associated) | CD74 | 5q32 | K01144.1 | Hs.84298 | 100.0 | Below | 18.9 |
| 13 | 210116_at | SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome) | SH2D1A | Xq25-q26 | AF072930.1 | Hs.151544 | 100.0 | Above | 159.0 |
| 14 | 210982_s_at | major histocompatibility complex, class II, DR alpha | HLA-DRA | 6p21.3 | M60333.1 | Hs.76807 | 100.0 | Below | 32.5 |
| 15 | 211990_at | major histocompatibility complex, class II, DP alpha 1 | HLA-DPA1 | 6p21.3 | M27487.1 | Hs.914 | 100.0 | Below | 29.4 |
| 16 | 211991_s_at | major histocompatibility complex, class II, DP alpha 1 | HLA-DPA1 | 6p21.3 | M27487.1 | Hs.914 | 100.0 | Below | 35.9 |
| 17 | 213539_at | CD3D antigen, delta polypeptide (TiT3 complex) | CD3D | 11q23 | NM_000732.1 | Hs.95327 | 100.0 | Above | 37.0 |
| 18 | 214049_x_at | CD7 antigen (p41) | CD7 | 17q25.2-q25.3 | A1829961 | Hs.36972 | 100.0 | Above | 327.3 |
| 19 | 214551_s_at | CD7 antigen (p41) | CD7 | 17q25.2-q25.3 | NM_006137.2 | Hs.36972 | 100.0 | Above | 241.8 |
| 20 | 217147_s_at | T-cell receptor interacting molecule | TRIM | 3q13 | AJ240085.1 | Hs.138701 | 100.0 | Above | 43.6 |
| 21 | 217478_s_at | MHC, class IIa, HLA-DMA | HLA-DMA | | X76775 | | 100.0 | Below | 14.1 |
| 22 | 221969_at | paired box gene 5 (B-cell lineage specific activator protein) | PAX5 | 9p13 | BF510692 | Hs.22030 | 100.0 | Below | 3922.0 |

| | | | | | | | | | |
|----|-------------|--|----------|-----------------|-------------|-----------|-------|-------|---------|
| 23 | 227646_at | early B-cell factor | EBF | 5q34 | BG435302 | Hs.32425 | 100.0 | Below | 79.2 |
| 24 | 229487_at | Homo sapiens cDNA FLJ39389 | | | W73890 | Hs.120785 | 100.0 | Below | 7685.7 |
| 25 | 229838_at | Homo sapiens cDNA FLJ39156 | | | AI377271 | Hs.376140 | 100.0 | Above | 13.4 |
| 26 | 232204_at | early B-cell factor | EBF | 5q34 | AF208502.1 | Hs.32425 | 100.0 | Below | 7112.4 |
| 27 | 266_s_at | CD24 antigen (small cell lung carcinoma cluster 4 antigen) | CD24 | 6q21 | L33930 | Hs.286124 | 91.3 | Below | 135.5 |
| 28 | 39318_at | T-cell leukemia/lymphoma 1A | TCL1A | 14q32.1 | X82240 | Hs.2484 | 91.3 | Below | 18114.7 |
| 29 | 203965_at | ubiquitin specific protease 20 | USP20 | 9q34.12-q34.13 | NM_006676.1 | Hs.5452 | 91.3 | Above | 9.6 |
| 30 | 204891_s_at | lymphocyte-specific protein tyrosine kinase | LCK | 1p34.3 | NM_005356.1 | Hs.1765 | 91.3 | Above | 14.3 |
| 31 | 205255_x_at | transcription factor 7 (T-cell specific, HMG-box) | TCF7 | 5q31.1 | NM_003202.1 | Hs.169294 | 91.3 | Above | 8.5 |
| 32 | 207655_s_at | B-cell linker | BLNK | 10q23.2-q23.33 | NM_013314.1 | Hs.167746 | 91.3 | Below | 133.1 |
| 33 | 209771_x_at | CD24 antigen (small cell lung carcinoma cluster 4 antigen) | CD24 | 6q21 | AA761181 | Hs.286124 | 91.3 | Below | 51.5 |
| 34 | 211796_s_at | T cell receptor beta locus | TRB | 7q34 | AF043179.1 | Hs.303157 | 91.3 | Above | 21.5 |
| 35 | 213792_s_at | insulin receptor | INSR | 19p13.3-p13.2 | AA485908 | Hs.89695 | 91.3 | Below | 8.6 |
| 36 | 215193_x_at | major histocompatibility complex, class II, DR beta 3 | HLA-DRB3 | 6p21.3 | AJ297586.1 | Hs.308026 | 91.3 | Below | 16.7 |
| 37 | 216379_x_at | KIAA1919 protein | KIAA1919 | 6q22.1 | AK000168.1 | Hs.381004 | 91.3 | Below | 56.0 |
| 38 | 219191_s_at | bridging integrator 2 | BIN2 | 12q13 | NM_016293.1 | Hs.14770 | 91.3 | Above | 264.1 |
| 39 | 219563_at | hypothetical protein FLJ21276 | FLJ21276 | 14q32.2 | NM_024633.1 | Hs.41502 | 91.3 | Below | 5.8 |
| 40 | 219724_s_at | KIAA0748 gene product | KIAA0748 | 12q12 | NM_014796.1 | Hs.33187 | 91.3 | Above | 11.4 |
| 41 | 221750_at | 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1 (soluble) | HMGCS1 | 5p14-p13 | BG035985 | Hs.77910 | 91.3 | Above | 3.5 |
| 42 | 226157_at | Homo sapiens cDNA FLJ39131 | | | AI569747 | Hs.379018 | 91.3 | Above | 4.6 |
| 43 | 226496_at | hypothetical protein FLJ22611 | FLJ22611 | 9p11.1 | BG291039 | Hs.27774 | 91.3 | Below | 7.7 |
| 44 | 204214_s_at | RAB32, member RAS oncogene family | RAB32 | 6q24.3 | NM_006834.1 | Hs.32217 | 90.6 | Above | 137.0 |
| 45 | 204777_s_at | mal, T-cell differentiation protein | MAL | 2cen-q13 | NM_002371.2 | Hs.80395 | 90.6 | Above | 95.2 |
| 46 | 204890_s_at | lymphocyte-specific protein tyrosine kinase | LCK | 1p34.3 | U07236.1 | Hs.1765 | 90.6 | Above | 19.4 |
| 47 | 205049_s_at | CD79A antigen (immunoglobulin-associated alpha) | CD79A | 19q13.2 | NM_001783.1 | Hs.79630 | 90.6 | Below | 11.6 |
| 48 | 205254_x_at | transcription factor 7 (T-cell specific, HMG-box) | TCF7 | 5q31.1 | AW027359 | Hs.169294 | 90.6 | Above | 377.1 |
| 49 | 205504_at | Bruton agammaglobulinemia tyrosine kinase | BTK | Xq21.33-q22 | NM_000061.1 | Hs.159494 | 90.6 | Below | 7.3 |
| 50 | 210915_x_at | T cell receptor beta locus | TRB | 7q34 | M15564.1 | Hs.303157 | 90.6 | Above | 16.2 |
| 51 | 211211_x_at | SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome) | SH2D1A | Xq25-q26 | AF100542.1 | Hs.151544 | 90.6 | Above | 2103.7 |
| 52 | 213830_at | T cell receptor delta locus | TRD | 14q11.2 | AW007751 | Hs.2014 | 90.6 | Above | 7379.7 |
| 53 | 216191_s_at | T cell receptor delta locus | TRD | 14q11.2 | X72501.1 | Hs.2014 | 90.6 | Above | 253.1 |
| 54 | 217143_s_at | T cell receptor delta locus | TRD | 14q11.2 | X06557.1 | Hs.2014 | 90.6 | Above | 150.1 |
| 55 | 219528_s_at | B-cell CLL/lymphoma 11B (zinc finger protein) | BCL11B | 14q32.31-q32.32 | NM_022898.1 | Hs.57987 | 90.6 | Above | 11.9 |

| | | | | | | | | | |
|----|-------------|---|----------|-----------------|-------------|-----------|------|-------|---------|
| 56 | 220418_at | ubiquitin associated and SH3 domain containing, A | UBASH3A | 21q22.3 | NM_018961.1 | Hs.183924 | 90.6 | Above | 766.4 |
| 57 | 222895_s_at | B-cell CLL/lymphoma 11B (zinc finger protein) | BCL11B | 14q32.31-q32.32 | AA918317 | Hs.57987 | 90.6 | Above | 12.1 |
| 58 | 223553_s_at | hypothetical protein FLJ22570 | FLJ22570 | 5q35.3 | BC004564.1 | Hs.122559 | 90.6 | Below | 7.0 |
| 59 | 225090_at | HRD1 protein | HRD1 | 11q12 | AA844682 | Hs.334819 | 90.6 | Below | 3.8 |
| 60 | 226459_at | Homo sapiens gastric cancer-related protein GCYS-20 (gcys-20) mRNA, complete cds | | | AW575754 | Hs.86437 | 90.6 | Below | 11.9 |
| 61 | 228314_at | Homo sapiens cDNA FLJ37485 | | | BE877357 | Hs.24181 | 90.6 | Below | 5.1 |
| 62 | 201384_s_at | membrane component, chromosome 17, surface marker 2 (ovarian carcinoma antigen CA125) | M17S2 | 17q21.1 | NM_005899.1 | Hs.277721 | 83.8 | Above | 3.3 |
| 63 | 202540_s_at | 3-hydroxy-3-methylglutaryl-Coenzyme A reductase | HMGCR | 5q13.3-q14 | NM_000859.1 | Hs.11899 | 83.8 | Above | 4.6 |
| 64 | 203198_at | cyclin-dependent kinase 9 (CDC2-related kinase) | CDK9 | 9q34.1 | NM_001261.1 | Hs.150423 | 83.8 | Below | 5.2 |
| 65 | 203932_at | major histocompatibility complex, class II, DM beta | HLA-DMB | 6p21.3 | NM_002118.1 | Hs.1162 | 83.8 | Below | 8.2 |
| 66 | 204613_at | phospholipase C, gamma 2 (phosphatidylinositol-specific) | PLCG2 | 16q24.1 | NM_002661.1 | Hs.75648 | 83.8 | Below | 4.0 |
| 67 | 205267_at | POU domain, class 2, associating factor 1 | POU2AF1 | 11q23.1 | NM_006235.1 | Hs.2407 | 83.8 | Below | 11.7 |
| 68 | 208650_s_at | CD24 antigen (small cell lung carcinoma cluster 4 antigen) | CD24 | 6q21 | BG327863 | Hs.286124 | 83.8 | Below | 110.8 |
| 69 | 208651_x_at | CD24 antigen (small cell lung carcinoma cluster 4 antigen) | CD24 | 6q21 | M58664.1 | Hs.286124 | 83.8 | Below | 49.3 |
| 70 | 209995_s_at | T-cell leukemia/lymphoma 1A | TCL1A | 14q32.1 | BC003574.1 | Hs.2484 | 83.8 | Below | 20166.2 |
| 71 | 210038_at | protein kinase C, theta | PRKCQ | 10p15 | AL137145 | Hs.211593 | 83.8 | Above | 12.7 |
| 72 | 211126_s_at | cysteine and glycine-rich protein 2 | CSRP2 | 12q21.1 | U46006.1 | Hs.10526 | 83.8 | Below | 16.8 |
| 73 | 220068_at | pre-B lymphocyte gene 3 | VPREB3 | 22q11.23 | NM_013378.1 | Hs.136713 | 83.8 | Below | 6559.8 |
| 74 | 226245_at | Homo sapiens mRNA; cDNA DKFZp451C132 | | | U55984 | Hs.356531 | 83.8 | Above | 9.1 |
| 75 | 202615_at | Homo sapiens mRNA; cDNA DKFZp686D0521 | | | BF222895 | Hs.356786 | 82.2 | Above | 3.1 |
| 76 | 224861_at | Homo sapiens cDNA FLJ31057 | | | BF477658 | Hs.296261 | 82.2 | Above | 3.5 |
| 77 | 201194_at | selenoprotein W, 1 | SEPW1 | 19q13.3 | NM_003009.1 | Hs.14231 | 82.0 | Above | 4.0 |
| 78 | 201349_at | solute carrier family 9 (sodium/hydrogen exchanger), isoform 3 regulatory factor 1 | SLC9A3R1 | 17q25.2 | NM_004252.1 | Hs.184276 | 82.0 | Above | 2.7 |
| 79 | 202539_s_at | 3-hydroxy-3-methylglutaryl-Coenzyme A reductase | HMGCR | 5q13.3-q14 | AL518627 | Hs.11899 | 82.0 | Above | 3.5 |
| 80 | 203588_s_at | transcription factor Dp-2 (E2F dimerization partner 2) | TFDP2 | 3q23 | BG034328 | Hs.19131 | 82.0 | Above | 18.5 |
| 81 | 204852_s_at | protein tyrosine phosphatase, non-receptor type 7 | PTPN7 | 1q32.1 | NM_002832.1 | Hs.35 | 82.0 | Above | 9.3 |
| 82 | 207434_s_at | FXVD domain containing ion transport regulator 2 | FXVD2 | 11q23 | NM_021603.1 | Hs.19520 | 82.0 | Above | 15.7 |

| | | | | | | | | | |
|-----|-------------|--|----------|---------------|-------------|-----------|------|-------|--------|
| 83 | 208872_s_at | DNA segment, single copy probe LNS-CAI/LNS-CAII (deleted in polyposis) | D5S346 | 5q22-q23 | AA814140 | Hs.178112 | 82.0 | Below | 2.7 |
| 84 | 209200_at | MADS box transcription enhancer factor 2, polypeptide C (myocyte enhancer factor 2C) | MEF2C | 5q14 | N22468 | Hs.78995 | 82.0 | Below | 9.1 |
| 85 | 212795_at | KIAA1033 protein | KIAA1033 | 12q24.11 | AL137753.1 | Hs.12144 | 82.0 | Below | 2.5 |
| 86 | 212827_at | immunoglobulin heavy constant mu | IGHM | 14q32.33 | X17115.1 | Hs.153261 | 82.0 | Below | 16.1 |
| 87 | 213193_x_at | T cell receptor beta locus | TRB | 7q34 | AL559122 | Hs.303157 | 82.0 | Above | 11.1 |
| 88 | 221002_s_at | tetraspanin similar to TM4SF9 | DC-TM4F2 | 10q23.2 | NM_030927.1 | Hs.101395 | 82.0 | Below | 2.1 |
| 89 | 225314_at | hypothetical protein MGC45416 | MGC45416 | 4p12 | BG291649 | Hs.95835 | 82.0 | Above | 5.7 |
| 90 | 227432_s_at | insulin receptor | INSR | 19p13.3-p13.2 | AI215106 | Hs.89695 | 82.0 | Below | 6.9 |
| 91 | 203332_s_at | inositol polyphosphate-5-phosphatase, 145kDa | INPP5D | 2q36-q37 | NM_005541.1 | Hs.155939 | 81.5 | Below | 2.4 |
| 92 | 203589_s_at | transcription factor Dp-2 (E2F dimerization partner 2) | TFDP2 | 3q23 | NM_006286.1 | Hs.19131 | 81.5 | Above | 37.6 |
| 93 | 205674_x_at | FXVD domain containing ion transport regulator 2 | FXVD2 | 11q23 | NM_001680.2 | Hs.19520 | 81.5 | Above | 13.0 |
| 94 | 209881_s_at | linker for activation of T cells | LAT | 16q13 | AF036905.1 | Hs.83496 | 81.5 | Above | 1953.6 |
| 95 | 211005_at | linker for activation of T cells | LAT | 16q13 | AF036906.1 | Hs.83496 | 81.5 | Above | 70.4 |
| 96 | 211075_s_at | CD47 | | | Z25521.1 | | 81.5 | Above | 2.1 |
| 97 | 211210_x_at | SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome) | SH2D1A | Xq25-q26 | AF100539.1 | Hs.151544 | 81.5 | Above | 315.1 |
| 98 | 213601_at | slit homolog 1 (Drosophila) | SLIT1 | 10q23.3-q24 | AB011537.2 | Hs.133466 | 81.5 | Above | 1877.2 |
| 99 | 213857_s_at | CD47 antigen (Rh-related antigen, integrin-associated signal transducer) | CD47 | 3q13.1-q13.2 | BG230614 | Hs.82685 | 81.5 | Above | 2.2 |
| 100 | 214924_s_at | KIAA1042 protein | KIAA1042 | 3p25.3-p24.1 | AK000754.1 | Hs.6705 | 81.5 | Below | 2.4 |

Table S14. Top 100 chi-square probe sets selected for *TEL-AML1* in parallel format

| | U133 probe set | Gene description | Symbol | Location | GenBank Reference | Unigene Reference | Chi-square value | Above/below | Fold change |
|---|----------------|--|----------|---------------|-------------------|-------------------|------------------|-------------|-------------|
| 1 | 224722_at | KIAA1323 | KIAA1323 | 18q11.1 | W80418 | Hs.34892 | 100.0 | Above | 7.7 |
| 2 | 227377_at | FLJ12722 | FLJ12722 | | AK022784.1 | Hs.24739 | 100.0 | Above | 2446.3 |
| 3 | 237206_at | Homo sapiens cDNA FLJ39434 | FLJ39434 | | AI452798 | Hs.42128 | 100.0 | Above | 23.5 |
| 4 | 241505_at | ESTs | | | BF513468 | Hs.196779 | 100.0 | Above | 13.2 |
| 5 | 205109_s_at | Rho guanine nucleotide exchange factor (GEF) 4 | ARHGEF4 | 2q22 | NM_015320.1 | Hs.6066 | 92.6 | Above | 147.0 |
| 6 | 210650_s_at | piccolo (presynaptic cytomatrix protein) | PCLO | 7q11.23-q21.3 | BC001304.1 | Hs.12376 | 92.6 | Above | 100.4 |
| 7 | 213558_at | piccolo (presynaptic cytomatrix protein) | PCLO | 7q11.23-q21.3 | AB011131.1 | Hs.12376 | 92.6 | Above | 76.8 |
| 8 | 224720_at | KIAA1323 protein | KIAA1323 | 18q11.1 | W80418 | Hs.34892 | 92.6 | Above | 4.3 |

| | | | | | | | | | |
|----|-------------|---|----------|---------------|-------------|-----------|------|-------|-------|
| 9 | 235694_at | Homo sapiens, clone IMAGE:4661943, mRNA, partial cds | | | N49233 | Hs.46914 | 92.6 | Above | 9.2 |
| 10 | 202808_at | hypothetical protein FLJ20154 | FLJ20154 | 10q24.32 | AK000161.1 | Hs.10346 | 92.2 | Above | 3.7 |
| 11 | 206032_at | desmocollin 3 | DSC3 | 18q12.1 | AI797281 | Hs.41690 | 92.2 | Above | 53.6 |
| 12 | 206033_s_at | desmocollin 3 | DSC3 | 18q12.1 | NM_001941.2 | Hs.41690 | 92.2 | Above | 354.6 |
| 13 | 224725_at | KIAA1323 protein | KIAA1323 | 18q11.1 | W80418 | Hs.34892 | 92.2 | Above | 3.6 |
| 14 | 204849_at | transcription factor-like 5 (basic helix-loop-helix) | TCFL5 | 20q13.3-qter | NM_006602.1 | Hs.30696 | 86.2 | Above | 8.8 |
| 15 | 206231_at | potassium intermediate/small conductance calcium-activated channel, subfamily N, member 1 | KCNN1 | 19p13.1 | NM_002248.2 | Hs.158173 | 86.2 | Above | 72.1 |
| 16 | 220451_s_at | baculoviral IAP repeat-containing 7 (livin) | BIRC7 | 20q13.3 | NM_022161.1 | Hs.256126 | 86.2 | Above | 25.2 |
| 17 | 211222_s_at | huntingtin-associated protein 1 (neuroan 1) | HAP1 | 17q21.2-q21.3 | AF040723.1 | Hs.158300 | 84.9 | Above | 80.2 |
| 18 | 228158_at | ESTs, Highly similar to A43542 lymphocyte-specific protein 1 - human | | | AI623211 | Hs.374584 | 84.9 | Above | 8.2 |
| 19 | 203464_s_at | epsin 2 | EPN2 | 17p11.1 | NM_014964.1 | Hs.7407 | 84.7 | Above | 42.9 |
| 20 | 213317_at | chloride intracellular channel 5 | CLIC5 | 6p21.1 | AL049313.1 | Hs.21103 | 84.7 | Above | 98.5 |
| 21 | 226817_at | Desmocollin 2 | DSC2 | 18q12.1 | AU154691 | Hs.95612 | 84.7 | Above | 47.9 |
| 22 | 229339_at | EST | | | AI093327 | Hs.42128 | 84.7 | Above | 30.8 |
| 23 | 203184_at | fibryllin 2 (congenital contractural arachnodactyly) | FBN2 | 5q23-q31 | NM_001999.2 | Hs.79432 | 80.4 | Above | 14.3 |
| 24 | 218627_at | hypothetical protein FLJ11259 | FLJ11259 | 12q23.1 | NM_018370.1 | Hs.184465 | 78.5 | Above | 4.6 |
| 25 | 221748_s_at | Tensin | TNS | 2q35 | AL046979 | Hs.9973 | 78.5 | Above | 6.5 |
| 26 | 37986_at | erythropoietin receptor | EPOR | 19p13.3-p13.2 | M60459 | Hs.127826 | 78.5 | Above | 16.8 |
| 27 | 204615_x_at | isopentenyl-diphosphate delta isomerase | IDI1 | 10p15.3 | NM_004508.1 | Hs.76038 | 77.3 | Above | 2.6 |
| 28 | 208881_x_at | isopentenyl-diphosphate delta isomerase | IDI1 | 10p15.3 | BC005247.1 | Hs.76038 | 77.3 | Above | 2.6 |
| 29 | 221747_at | Tensin | TNS | 2q35 | AL046979 | Hs.9973 | 77.3 | Above | 48.7 |
| 30 | 232750_at | Homo sapiens cDNA FLJ13750 | | | AU158570 | Hs.298099 | 77.3 | Above | 34.7 |
| 31 | 227862_at | ESTs, Weakly similar to hypothetical protein FLJ22184 | | | AA037766 | Hs.125073 | 77.1 | Above | 14.3 |
| 32 | 203910_at | PTPL1-associated RhoGAP 1 | PARG1 | 1p22.1 | NM_004815.1 | Hs.70983 | 75.2 | Above | 7.1 |
| 33 | 239673_at | ESTs | | | AW080999 | Hs.163924 | 72.9 | Above | 8.9 |
| 34 | 240950_s_at | hypothetical protein FLJ32658 | FLJ32658 | 19q13.33 | AA400740 | Hs.177990 | 72.9 | Above | 10.3 |
| 35 | 208056_s_at | core-binding factor, runt domain, alpha subunit 2; translocated to, 3 | CBFA2T3 | 16q24 | NM_005187.2 | Hs.110099 | 71.8 | Above | 2.5 |
| 36 | 206591_at | recombination activating gene 1 | RAG1 | 11p13 | NM_000448.1 | Hs.73958 | 71.1 | Above | 5.4 |
| 37 | 209963_s_at | erythropoietin receptor | EPOR | 19p13.3-p13.2 | M34986.1 | Hs.127826 | 71.1 | Above | 8.0 |
| 38 | 219866_at | chloride intracellular channel 5 | CLIC5 | 6p12.1-21.1 | NM_016929.1 | Hs.283021 | 71.1 | Above | 59.8 |
| 39 | 224726_at | KIAA1323 protein | KIAA1323 | 18q11.1 | W80418 | Hs.34892 | 71.1 | Above | 25.8 |
| 40 | 204297_at | phosphoinositide-3-kinase, class 3 | PIK3C3 | 18q12.3 | NM_002647.1 | Hs.32971 | 70.6 | Above | 4.5 |
| 41 | 217628_at | chloride intracellular channel 5 | CLIC5 | 6p21.1 | BF032808 | Hs.21103 | 70.0 | Above | 86.8 |

| | | | | | | | | | |
|----|-------------|--|----------|---------------|-------------|-----------|------|-------|-------|
| 42 | 218804_at | hypothetical protein FLJ10261 | FLJ10261 | 11q13.3 | NM_018043.1 | Hs.26176 | 70.0 | Above | 41.3 |
| 43 | 211891_s_at | Rho guanine nucleotide exchange factor (GEF) 4 | ARHGEF4 | 2q22 | AB042199.1 | Hs.6066 | 70.0 | Above | 452.6 |
| 44 | 240446_at | KIAA1323 | KIAA1323 | | AI798164 | Hs.140903 | 70.0 | Above | 101.4 |
| 45 | 244107_at | ESTs | | | AW189097 | Hs.370922 | 70.0 | Above | 518.9 |
| 46 | 209228_x_at | Putative prostate cancer tumor suppressor | N33 | 8p22 | U42349.1 | Hs.71119 | 67.9 | Above | 20.6 |
| 47 | 225129_at | Homo sapiens cDNA FLJ37548 fis, clone BRCAN2027970, moderately similar to Homo sapiens copine I mRNA | | | AW170571 | Hs.376310 | 67.9 | Above | 2.9 |
| 48 | 203611_at | telomeric repeat binding factor 2 | TERF2 | 16q22.1 | NM_005652.1 | Hs.100030 | 66.4 | Above | 5.3 |
| 49 | 236430_at | hypothetical protein MGC23911 | MGC23911 | 16q23.1 | AA708152 | Hs.119840 | 66.4 | Above | 16.6 |
| 50 | 200709_at | FK506 binding protein 1A (12kDa) | FKBP1A | 20p13 | NM_000801.1 | Hs.380080 | 65.6 | Above | 1.8 |
| 51 | 209035_at | midkine (neurite growth-promoting factor 2) | MDK | 11p11.2 | M69148.1 | Hs.82045 | 65.6 | Above | 4.5 |
| 52 | 226038_at | hypothetical protein FLJ23749 | FLJ23749 | 8p23.1 | BF680438 | Hs.180178 | 65.6 | Above | 5.2 |
| 53 | 232227_at | ESTs, Moderately similar to AF161442_1 HSPC324 | | | AV736391 | Hs.191591 | 65.6 | Above | 14.7 |
| 54 | 205794_s_at | neuro-oncological ventral antigen 1 | NOVA1 | 14q | NM_002515.1 | Hs.214 | 63.8 | Above | 40.1 |
| 55 | 227111_at | Homo sapiens cDNA FLJ31099 fis, clone IMR321000230 | | | BG179317 | Hs.177633 | 63.8 | Above | 2.7 |
| 56 | 230698_at | Homo sapiens mRNA; cDNA DKFZp434H205 | | | AW072102 | Hs.101375 | 63.8 | Above | 8.7 |
| 57 | 203038_at | protein tyrosine phosphatase, receptor type, K | PTPRK | 6q22.2-23.1 | NM_002844.1 | Hs.79005 | 63.4 | Above | 9.0 |
| 58 | 202838_at | fucosidase, alpha-L- 1, tissue | FUCA1 | 1p34 | NM_000147.1 | Hs.576 | 63.0 | Above | 4.8 |
| 59 | 225835_at | Homo sapiens cDNA: FLJ21409 fis, clone COL03924 | | | AK025062.1 | Hs.172129 | 63.0 | Above | 3.6 |
| 60 | 229790_at | telomeric repeat binding factor 2 | TERF2 | 16q22.1 | AW006832 | Hs.100030 | 63.0 | Above | 7.3 |
| 61 | 235872_at | ESTs | | | BE408975 | Hs.63335 | 63.0 | Above | 17.5 |
| 62 | 239300_at | ESTs | | | AI632214 | Hs.193041 | 63.0 | Above | 3.0 |
| 63 | 241940_at | ESTs | | | BF477544 | Hs.188270 | 63.0 | Above | 2.9 |
| 64 | 203370_s_at | enigma (LIM domain protein) | ENIGMA | 5q35.3 | NM_005451.2 | Hs.102948 | 62.9 | Above | 8.1 |
| 65 | 218152_at | high-mobility group 20A | HMG20A | 15q24 | NM_018200.1 | Hs.69594 | 62.9 | Above | 1.7 |
| 66 | 218813_s_at | SH3-domain GRB2-like endophilin B2 | SH3GLB2 | 9q34 | NM_020145.1 | Hs.30002 | 62.9 | Above | 6.1 |
| 67 | 210058_at | mitogen-activated protein kinase 13 | MAPK13 | 6p21.1 | BC000433.1 | Hs.178695 | 62.6 | Above | 2.3 |
| 68 | 218856_at | tumor necrosis factor receptor superfamily, member 21 | TNFRSF21 | 6p21.1-12.2 | NM_016629.1 | Hs.159651 | 62.6 | Above | 13.1 |
| 69 | 234983_at | ESTs | | | BE893995 | Hs.31447 | 62.6 | Above | 2.4 |
| 70 | 209193_at | pim-1 oncogene | PIM1 | 6p21.2 | M24779.1 | Hs.81170 | 60.7 | Above | 2.0 |
| 71 | 209962_at | erythropoietin receptor | EPOR | 19p13.3-p13.2 | M34986.1 | Hs.127826 | 60.7 | Above | 18.9 |
| 72 | 218625_at | neuritin 1 | NRN1 | 6p24.1 | NM_016588.1 | Hs.103291 | 60.7 | Above | 5.0 |
| 73 | 223468_s_at | hypothetical protein from EUROIMAGE 363668 RGM: likely ortholog of chicken repulsive guidance molecule | RGM | 15q26.1 | AL136826.1 | Hs.271277 | 60.7 | Above | 10.5 |

| | | | | | | | | | |
|-----|-------------|---|----------|--------------|-------------|-----------|------|-------|------|
| 74 | 229356_x_at | hypothetical protein KIAA1259 | KIAA1259 | 15q13.3 | AK002176.1 | Hs.40193 | 60.7 | Above | 1.7 |
| 75 | 204404_at | solute carrier family 12 (sodium/potassium/chloride transporters), member 2 | SLC12A2 | 5q23.3 | NM_001046.1 | Hs.110736 | 59.3 | Above | 2.0 |
| 76 | 212063_at | CD44 antigen (homing function and Indian blood group system) | CD44 | 11p13 | BE903880 | Hs.169610 | 59.3 | Below | 7.4 |
| 77 | 221773_at | ELK3, ETS-domain protein (SRF accessory protein 2) | ELK3 | 12q23 | AW575374 | Hs.288555 | 59.1 | Above | 3.1 |
| 78 | 208886_at | H1 histone family, member 0 | H1F0 | 22q13.1 | BC000145.1 | Hs.226117 | 58.5 | Above | 4.2 |
| 79 | 212399_s_at | KIAA0121 gene product | KIAA0121 | 3p25.2 | D50911.2 | Hs.155584 | 58.5 | Above | 1.8 |
| 80 | 219954_s_at | cytosolic beta-glucosidase | GLUC | 4p15.31 | NM_020973.1 | Hs.146182 | 58.5 | Above | 37.3 |
| 81 | 220692_at | HSPC047 protein | HSPC047 | 7q11.22 | NM_014147.1 | Hs.278943 | 58.5 | Above | 16.8 |
| 82 | 225357_s_at | hypothetical protein KIAA1259 | KIAA1259 | 15q13.3 | AI659419 | Hs.40193 | 58.5 | Above | 4.9 |
| 83 | 206316_s_at | kinetochore associated 1 | KNTC1 | 12q24.31 | NM_014708.1 | Hs.115778 | 58.2 | Above | 2.7 |
| 84 | 207957_s_at | protein kinase C, beta 1 | PRKCB1 | 16p11.2 | NM_002738.1 | Hs.77202 | 56.8 | Above | 2.1 |
| 85 | 213122_at | KIAA1750 protein | KIAA1750 | 8q22.1 | AI096375 | Hs.173094 | 56.8 | Above | 8.9 |
| 86 | 226885_at | ESTs | | | AI743880 | Hs.12876 | 56.8 | Above | 10.7 |
| 87 | 218742_at | protein related to Narf | HPRN | 16p13.3 | NM_022493.1 | Hs.22158 | 56.6 | Above | 4.2 |
| 88 | 206574_s_at | protein tyrosine phosphatase type IVA, member 3 | PTP4A3 | | NM_007079.1 | Hs.43666 | 56.4 | Above | 6.9 |
| 89 | 209199_s_at | MADS box transcription enhancer factor 2, polypeptide C (myocyte enhancer factor 2C) | MEF2C | 5q14 | N22468 | Hs.78995 | 56.4 | Above | 1.8 |
| 90 | 212154_at | syndecan 2 (heparan sulfate proteoglycan 1, cell surface-associated, fibroglycan) | SDC2 | 8q22-q23 | J04621.1 | Hs.1501 | 56.4 | Above | 6.6 |
| 91 | 213423_x_at | Putative prostate cancer tumor suppressor | N33 | 8p22 | AI884858 | Hs.71119 | 56.4 | Above | 15.6 |
| 92 | 228831_s_at | hypothetical protein FLJ00058 | FLJ00058 | 19p13.3 | AL039870 | Hs.6750 | 56.4 | Above | 2.0 |
| 93 | 208910_s_at | complement component 1, q subcomponent binding protein | C1QBP | 17p13.3 | L04636.1 | Hs.78614 | 56.0 | Below | 1.9 |
| 94 | 217873_at | MO25 protein | MO25 | 2q36.1 | NM_016289.1 | Hs.6406 | 56.0 | Above | 1.4 |
| 95 | 201976_s_at | myosin X | MYO10 | 5p15.1-p14.3 | NM_012334.1 | Hs.61638 | 55.9 | Above | 12.5 |
| 96 | 209685_s_at | protein kinase C, beta 1 | PRKCB1 | 16p11.2 | M13975.1 | Hs.77202 | 55.9 | Above | 1.9 |
| 97 | 213017_at | abhydrolase domain containing 3 | ABHD3 | 18q11.1 | AL534702 | Hs.13377 | 55.9 | Above | 4.0 |
| 98 | 213261_at | KIAA0342 gene product | KIAA0342 | 3p22.1 | AA035414 | Hs.16950 | 55.9 | Above | 1.6 |
| 99 | 224881_at | Homo sapiens mRNA; cDNA DKFZp434B1620 (from clone DKFZp434B1620) | | | AV724827 | Hs.43112 | 55.9 | Above | 1.7 |
| 100 | 225731_at | KIAA1223 | KIAA1223 | 4q26 | AB033049.1 | Hs.28783 | 55.9 | Above | 2.8 |

IV: Diagnostic Accuracy

The class discriminating genes were identified as described above, and then used in an ANN-based supervised learning algorithm. As previously discussed (section II, Supplemental Information), class assignment was based on a differential diagnostic tree format and required that the node value for assignment exceeded a statistically defined confidence level. The results of this analysis are shown in Table 4 of the paper and are included below for the reader's convenience.

| Subgroup | Training Set ^a | | Test Set ^b | |
|--------------------------|--------------------------------|----------------------------|--------------------------|--------------------------|
| | Apparent Accuracy ^c | True Accuracy ^d | Sensitivity ^e | Specificity ^f |
| <i>T-ALL</i> | 100% | 100% | 100% | 100% |
| <i>E2A-PBX1</i> | 100% | 100% | 100% | 100% |
| <i>TEL-AML1</i> | 98% | 100% | 100% | 100% |
| <i>BCR-ABL</i> | 100% | 95% | 75% | 100% |
| <i>MLL</i> rearrangement | 100% | 100% | 100% | 100% |
| Hyperdiploid >50 | 100% | 100% | 100% | 100% |

^a training set consisted of 100 cases with distribution: [T-ALL 12, *E2A-PBX1* 13, *TEL-AML1* 15, *BCR-ABL* 11, *MLL* 15, HD>50 13, other 21]

^b blinded test set consisted of 32 cases [T-ALL 2, *E2A-PBX1* 5, *TEL-AML1* 5, *BCR-ABL* 4, *MLL* 5, HD>50 4, other 7]

^c apparent accuracy determined by 3-fold cross-validation

^d true accuracy determined by class prediction on the blinded test set.

^e Sensitivity = (the number of positive cases predicted)/(the number of true positives).

^f Specificity = (the number of negative cases predicted)/(the number of true negatives).

To control for over-fitting of the data, we performed 10 additional rounds of analysis. For each round, new training and test sets were developed and discriminating probe sets reselected exclusively using the new training sets. The top 20 and 50 probe sets were then used in an ANN-based supervised learning algorithm, and their true accuracy assessed on the new test sets. This resulted in an average accuracy of class assignment of 97% (range 93.8%-100%) using 20 probes per class. Shown in Tables S15 and S16 are the results from these analyses. The numbers listed under the individual leukemia subtypes represent the number of misclassified cases in the training and test sets. The overall accuracies are listed on the right.

Table S15. Training and Test Set Permutation Results - Errors per group using 20 probe sets

| | T-ALL | | E2A-PBX1 | | TEL-AML1 | | BCR-ABL | | MLL | | Hyperdip>50 | | Overall Accuracy | |
|-----------|----------------|------|-----------------|------|-----------------|------|----------------|------|------------|------|-----------------------|------|-------------------------|------|
| | Training | Test | Training | Test | Training | Test | Training | Test | Training | Test | Training | Test | Training | Test |
| 1 | 0 ¹ | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 100 | 93.8 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 99 | 93.8 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 98 | 100 |
| 4 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 98 | 96.9 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 100 | 96.9 |
| 6 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 99 | 100 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 99 | 96.9 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 99 | 93.8 |

¹The number of misclassified cases obtained when diagnosing the indicated leukemia subtype

Table S16. Training and test set permutation results – errors per group using 50 probe sets

| | T-ALL | | E2A-PBX1 | | TEL-AML1 | | BCR-ABL | | MLL | | Hyperdip>50 | | Overall Accuracy | |
|-----------|----------------|------|-----------------|------|-----------------|------|----------------|------|------------|------|-----------------------|------|-------------------------|------|
| | Training | Test | Training | Test | Training | Test | Training | Test | Training | Test | Training | Test | Training | Test |
| 1 | 0 ¹ | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 99 | 100 |
| 2 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 99 | 90.6 |
| 3 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 97 | 96.9 |
| 4 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 97 | 96.9 |
| 5 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 2 | 1 | 0 | 98 | 90.6 |
| 6 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 98 | 100 |
| 7 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 99 | 100 |
| 8 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 99 | 100 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 100 | 96.9 |
| 10 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 97 | 93.8 |

¹The number of misclassified cases obtained when diagnosing the indicated leukemia subtype

Comparison of supervised learning algorithms

The performance of other supervised learning algorithms was compared to ANN. Using the original training and test sets, chi-squared was used to select the desired number of probes sets, and then the selected probes were used to build a model using ANN, SVM, and *k*-NN. ANN was performed with one hidden layer consisting of 4 nodes and the backpropagation epoch number was 5000. For the other algorithms, the linear SVM kernel was used and the *k*-NN parameter was 3. The comparison of the results is shown in Table S17 below. The comparison was performed using the top 20 and 50 probe sets, as well as the top 20 and 50 genes. The numbers correspond to the number of errors made in either the training or test set by class for each metric. Overall, ANN and SVM performed fairly comparably while *k*-NN gave slightly poorer results.

Table S17. Comparison of supervised learning algorithms

| | ANN | | SVM | | <i>k</i> -NN | |
|------------------|----------------|------|----------|------|--------------|------|
| | Training | Test | Training | Test | Training | Test |
| top 20 probes | | | | | | |
| T-ALL | 0 ¹ | 0 | 0 | 0 | 0 | 0 |
| E2A-PBX1 | 0 | 0 | 0 | 0 | 0 | 0 |
| TEL-AML1 | 0 | 0 | 0 | 0 | 0 | 0 |
| BCR-ABL | 1 | 2 | 1 | 2 | 2 | 1 |
| MLL | 0 | 0 | 0 | 0 | 0 | 1 |
| Hyperdiploid >50 | 0 | 0 | 0 | 0 | 0 | 0 |
| top 50 probes | | | | | | |
| T-ALL | 0 | 0 | 0 | 0 | 0 | 0 |
| E2A-PBX1 | 0 | 0 | 0 | 0 | 0 | 0 |
| TEL-AML1 | 1 | 0 | 0 | 0 | 0 | 0 |
| BCR-ABL | 0 | 1 | 1 | 1 | 2 | 1 |
| MLL | 0 | 0 | 0 | 0 | 0 | 1 |
| Hyperdiploid >50 | 0 | 0 | 0 | 0 | 0 | 0 |
| top 20 genes | | | | | | |
| T-ALL | 0 | 0 | 0 | 0 | 0 | 0 |
| E2A-PBX1 | 0 | 0 | 0 | 0 | 0 | 0 |
| TEL-AML1 | 1 | 0 | 0 | 0 | 0 | 0 |
| BCR-ABL | 1 | 2 | 1 | 2 | 1 | 1 |
| MLL | 0 | 0 | 0 | 0 | 0 | 1 |
| Hyperdiploid >50 | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | |
|------------------|---|---|---|---|---|---|
| top 50 genes | | | | | | |
| T-ALL | 0 | 0 | 0 | 0 | 0 | 0 |
| E2A-PBX1 | 0 | 0 | 0 | 0 | 0 | 0 |
| TEL-AML1 | 1 | 0 | 0 | 0 | 0 | 0 |
| BCR-ABL | 0 | 1 | 1 | 1 | 3 | 1 |
| MLL | 0 | 0 | 0 | 0 | 0 | 0 |
| Hyperdiploid >50 | 0 | 0 | 0 | 1 | 0 | 0 |

¹ The number of misclassified cases

V: Comparison of Expression Profiles and Real-time PCR (Taqman)

An evaluation of the discriminating probe sets selected from the U133A and U133B microarrays revealed that 20% of the identified class discriminators correspond to genes that were represented on the U95Av2 microarray, but had not been selected as class discriminators using the latter platform. A possible explanation for this observation is that the improved oligonucleotide design of the U133 microarrays results in an increase in both sensitivity/specificity, and thus allows the identification of class discriminators that had previously on the U95Av2 microarray fallen below statistical significance. To confirm this possibility, real-time RT-PCR (Taqman; Perkin-Elmer/Applied Biosystems, Foster City, CA) assays were developed for eleven genes that met these criteria. The probes selected were against discriminating genes for *BCR-ABL* (Neuropillin1, 212298_at; and ATP10C, 214255_at), *E2A-PBX1* (FLJ12280, 213909_at; and FAM3C, 201889_at), Hyperdiploid > 50 chromosomes (EST, 212419_at), T-ALL (RAB32, 204214_s_at; and TRIM, 217147_s_at), and *TEL-AML1* (Desmocollin 3, 206033_s_at; Epsin 2, 203464_s_at; NOVA1, 205794_s_at; and HAP1, 211222_s_at). Taqman primers were developed for each of these genes using the target sequences against which the microarray oligonucleotide probes were designed. Amplified product was quantitated using SYBR green and the resulting values normalized using the commercially available Taqman GAPDH primer and probe sequences.

RNA from representative diagnostic bone marrow samples of each class (four per class) were diluted to 35 ng/μl then treated for 15 minutes with 1.0 unit of DNase I (Invitrogen, Carlsbad, California) using the Invitrogen protocol to remove any contaminating DNA. Samples were reverse transcribed following the Taqman RT protocol using 1X RT mix (Perkin-Elmer/Applied Biosystems) and MultiScribe (Applied Biosystems). Samples were incubated at 25°C for 10 minutes, 48°C for 30 minutes, 95°C for 5 minutes, then

placed on ice for 2 minutes. Real-time PCR was performed on a PE Applied Biosystems 7700 prism using oligonucleotide primers designed using Primer Express. The following primers were used: (1) *Desmocollin 3*, 5'- GCA ACC AGT ATC ACT TCC CTG TT - 3' (forward primer), and 5'- GAA TTA CAA ATT CGG GCA TAC ATG -3' (reverse primer); (2) *Epsin 2*, 5'- GGA GGT GCA ATG GGA TGG -3' (forward), and 5'- GCC CTG CAC CGT CTG AAG -3' (reverse); (3) *Nova 1*, 5'- CAT CCC AGC TGC TCC TTT CT -3' (forward), and 5'- AGG CTG GAC GAA ATT CAG ACA -3' (reverse); (4) *HAPI*, 5'- AGC GCC TTC CCA ACA ATG T -3' (forward), and 5'- ACA AGA ACC AGG GTG GCT ACC -3' (reverse); (5) *Neuropilin 1*, 5'- AAA GAC ATT TGT TGG GAG TCA CAT T -3' (forward), and 5'- CAT GGT GAT CAA TAT TTT CCT GGA A -3' (reverse); (6) *ATP10C*, 5'- TTG AGT GCA TCC CAG CAT TC -3' (forward), and 5'- CCA GGG ATG TAA TAC CTA CAA GGC -3' (reverse); (7) *FLJ12280*, 5'- TTG CCT TGC TTA GAG AAT TAC TGC -3' (forward), and 5'- ACC CTG AGG GCC TAG AAA TCT G -3'; (8) *FAM3C*, 5'- TGT GTA TGA AGT ATC TCA AAC TGG AAC AT -3' (forward), and 5'- ATG CAT CTA AAC TTT AGG TTC GAA ATT -3' (reverse); (9) *FAB32*, 5'- CCC TGC TGT CCT CTT GGC T -3' (forward), and 5'- AAT TGG TCC ACC TGG GAA GG -3' (reverse); (10) *TRIM*, 5'- TTA CAA AGT ATT TTT CCC AAA GAT AGC TT -3' (forward), and 5'- TTT CTC CAA GTG ACT ATC TCT GGC TAG -3' (reverse); and (11) 212419: EST, 5'- GAG AAG GCT CCG ACG TCT CC -3', and 5'- CCA GGA ATA AGG CAC ACA ACG -3' (reverse).

The PCR reactions and amplification protocols were set up as previously described.³ The reverse transcribed cDNA was amplified in 1X SYBR green master mix containing a Taq Polymerase with 100 nM each of a forward and reverse primer. The GAPDH control for each cDNA was set up according to the Applied Biosystems protocol. Amplification proceeded as outlined in the “Quickstart

Guide: ABI Prism™ Dissociation Curve Software” instructions. Data collection was performed by linking a dissociation curve profile on the end of the PCR run. Dilution controls were set up for each gene of interest as well as for GAPDH, and standard curves were generated at the end of the run. The expression levels were quantitated for each gene, and then normalized to the quantitated GAPDH. The normalized real-time result for the four samples of each class were average together and the mean for that class was compared to the mean intensity value for the same four samples from the microchip array. This comparison is shown in Table S18 below.

Although the selected probe sets were from the lower end of the ranked list of discriminating genes for each class, their expression was significantly higher by Taqman measurements in the identical class that was identified to have high expression by microarray analysis. The high degree of correlation between the microarray data and that obtained using Taqman methodology suggests that the U133 microarray platform was correctly identifying these genes as class discriminators.

Table S18 Comparison of Expression profiles and Real-time PCR (Taqman)

| | | <i>TEL- AML1</i> | <i>BCR- ABL</i> | <i>E2A- PBX1</i> | T-ALL | Hyperdip>50 | <i>MLL</i> |
|----------------------|--------|----------------------|---------------------|----------------------|-------|-------------|------------|
| 206033_s_at | Affy | 1676.9 | 1 | 1 | 1 | 1 | 1 |
| Desmocollin 3 | RQ-PCR | 474.5 | 0 | 50.3 | 0 | 0 | 0 |
| 203464_s_at | Affy | 320.3 | 1 | 1 | 1 | 1 | 1 |
| Epsin 2 | RQ-PCR | 881.6 | 6.4 | 4.2 | 3.5 | 29.7 | 9.5 |
| 205794_s_at | Affy | 594.5 | 90.7 | 1 | 1 | 1 | 1 |
| NOVA1 | RQ-PCR | 785.9 | 10.9 | 0 | 0.5 | 0 | 0 |

| | | | | | | | |
|--------------------|--------|--------|---------|---------|---------|--------|--------|
| 211222_s_at | Affy | 1010.3 | 1 | 1 | 1 | 1 | 1 |
| HAP1 | RQ-PCR | 382.4 | 1.6 | 4.5 | 0.06 | 2.1 | 0.4 |
| 212298_at | Affy | 133.6 | 906.5 | 1 | 1 | 134.1 | 68.3 |
| Neuropilin1 | RQ-PCR | 94.1 | 8784.3 | 56.8 | 9.8 | 435.5 | 175.1 |
| 214255_at | Affy | 215 | 999.6 | 1 | 1 | 94.9 | 139.1 |
| ATP10C | RQ-PCR | 56.3 | 70406.4 | 31948.8 | 30961.9 | 3387.2 | 1209.3 |
| 213909_at | Affy | 1 | 71.2 | 1410.1 | 138.2 | 116.9 | 113.9 |
| FLJ12280 | RQ-PCR | 2.98 | 0.02 | 61.4 | 0.1 | 0.02 | 0.02 |
| 201889_at | Affy | 823.7 | 1034 | 3370 | 435.7 | 562.5 | 728.6 |
| FAM3C | RQ-PCR | 589.3 | 2228.4 | 10631.4 | 984.6 | 294.8 | 1258.9 |
| 204214_s_at | Affy | 1 | 1 | 1 | 1229.2 | 1 | 1 |
| RAB32 | RQ-PCR | 81.3 | 6.5 | 6.3 | 130.1 | 1 | 1.9 |
| 217147_s_at | Affy | 1 | 52.6 | 1 | 2531.3 | 1 | 1 |
| TRIM | RQ-PCR | 344.2 | 5.7 | 15.4 | 18.5 | 8.5 | 4 |
| 212419_at | Affy | 220 | 1 | 1 | 1 | 663.3 | 1 |
| EST | RQ-PCR | 252.5 | 66.5 | 83.2 | 26.2 | 789.2 | 179.2 |

VI: References

1. Yeoh EJ, Ross ME, Shurtleff SA, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell*. 2002;1:133-143.
2. Khan J, Wei JS, Ringner M, Saal LH, Ladanyi M, Westermann F, Berthold F, Schwab M, Antonescu CR, Peterson C, Meltzer PS. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat Med* 2001;7:673-679.
3. Vandesompele, J, DePaepe A, and Speleman F. Elimination of primer-dimer artifacts and genomic coamplification using a two-step SYBR green 1 real-time RT-PCR. *Analytical Biochemistry* 2002;303: 95-98.