Supplement Information for:

ACUTE LYMPHOBLASTIC LEUKEMIA WITH TEL-AML1 FUSION HAS LOWER EXPRESSION OF GENES INVOLVED IN PURINE METABOLISM AND LOWER DE NOVO PURINE SYNTHESIS

Gianluigi Zaza\textsuperscript{a,f}, Wenjian Yang\textsuperscript{a}, Leo Kager\textsuperscript{a}, Meyling Cheok\textsuperscript{a}, James Downing\textsuperscript{b}, Ching-Hon Pui\textsuperscript{b,c,e}, Cheng Cheng\textsuperscript{d}, Mary V. Relling\textsuperscript{a,e}, William E. Evans\textsuperscript{a,e}
Table of Contents

1 Supplemental methods and data

1.1 RNA quality control and source
1.2 Real Time RT-PCR

2 Supplemental Results

2.1 Patient characteristics
2.2 Purine metabolism pathway (KEGG) utilized for the microarray analysis
2.3 Clustering of patients based on gene expression pattern
2.4 Gene expression analysis with a case with split TEL gene

Figure M1: Real-time (TaqMan) RT-PCR standard curve
Figure M2: Real-time (Taq Man) RT-PCR vs. Affymetrix Gene Chip
Figure M3: Protein level by chemioluminescence vs. Affymetrix Gene Chip

Table A: Patient characteristics for ‘Training set’
Table B: Patient characteristics for ‘Test set’
Table C: Discriminanting genes selected TEL-AML1 vs. non-TEL-AML1 expression

Figure A: Purine metabolism graphical pathway diagram (KEGG pathway map)
Figure B: Correlation between percent of ALL cells in S-phase of the cell cycle and rate of DNPS
Figure C: “Unsupervised” hierarchical clustering and principal component analysis
Figure D: Hierarchical clustering and principal component analysis using 39 patients (with one additional case)
1 Supplemental methods and data

1.1 RNA quality control and source

Total RNA integrity was assessed by electrophoresis using the Agilent Bioanalyzer (Agilent, Palo Alto, California). Mononuclear (leukemia) cells were isolated from each bone marrow aspirate by centrifugation over Ficoll. The median percent blasts (post-Ficoll) was 97% in the diagnostic bone marrow samples.

1.2 Real time RT-PCR

Gene expression was also determined by real time RT-PCR assays (reverse transcription polymerase chain reaction) in four patient samples, for four genes that were discriminating between TEL-AML1 vs. non-TEL-AML1. Reactions were performed using TaqMan Universal PCR Master Mix Kit and the 7900-sequence-detection system (Applied Biosystem, Foster City, California).

One µg of total RNA was treated with Dnase I and reverse transcribed using Superscript II Rnase H- reverse transcriptase and oligo dT primers (invitrogen, Carlsbad, California). Additionally, controls that contained either no template or no reverse transcriptase, were included as negative controls in each run. Aliquots (0.5 µl) of RT reaction mixture (20 µl) were used for quantification of PAICS, IMPDH2, FHIT, NME1 and RNase P gene expression. The housekeeping gene RNase P (Applied Biosystems, Foster City, California) was used for normalization.
The total volume of PCR reaction was 50 µl of RT-product, 400 nM each of the forward and reverse primers, 250 nM of probe, and 1X master mix. The following thermal cycling parameters were used: two minutes at 50°C (activation of UNG enzyme to remove the carry-over PCR products), ten minutes at 95°C to activate AmpliTaq Gold DNA polymerase, 15 seconds at 95°C to denaturate and one minute at 60°C for annealing and extension, for a total of 45 cycles.

To estimate the amount of each of the two mRNAs in the four patient samples, we used linear regression analysis based on a standard curve representing six serial dilutions of cDNA made from the Nalm6 human leukemia cells (American Type Culture Collection, Rockville, Maryland). In the standard curve, we plotted fluorescent signal intensities against the number of PCR cycles on a semi-logarithmic scale.
Figure M1: Real-time (TaqMan) RT-PCR standard curve

Real-time RT-PCR standard curves for PAICS, IMPDH2, FHIT, NME1 and RNase P

All unknown samples were analyzed in duplicate in parallel with a standardization series from Nalm6 cDNA. The relative quantification values of genes for each sample were calculated based on their C_T value and the corresponding standard curves.

Each normalized and log transformed real-time RT-PCR gene expression level (IMPDH2, PAICS, FHIT and NME1) was compared to the log transformed signal of gene expression from the Affymetrix MAS 5.0 output. The correlation between real-time RT-PCR and Affymetrix GeneChip® was 94% for IMPDH2, 97% for PAICS, 89% for NME1 and 87% for FHIT, thereby confirming expression levels determined by the gene expression array.
Figure M2: Real-time (TaqMan) RT-PCR vs. Affymetrix GeneChip®

Real-time (TaqMan) RT-PCR results are plotted versus expression levels determined by Affymetrix GeneChip® results for IMPDH2, PAICS, FHIT and NME1 in four patients.

**IMPDH2 gene expression real-time PCR vs. GeneChip**

\[ y = 1.4846x - 6.6899 \]

\[ R^2 = 0.9456 \]

**PAICS gene expression real-time PCR vs. GeneChip**

\[ y = 2.236x - 9.3583 \]

\[ R^2 = 0.9785 \]

**NME1 gene expression real-time PCR vs. GeneChip**

\[ y = 1.8026x - 8.0946 \]

\[ R^2 = 0.8958 \]

**FHIT gene expression real-time PCR vs. GeneChip**

\[ y = 1.6966x - 5.9076 \]

\[ R^2 = 0.8722 \]
Figure M3: Protein level by chemioluminescence vs. Affymerix GeneChip®

Relative protein level are plotted versus expression levels determined by Affymetrix GeneChip® results for *FHIT* and *NME1* in five patients.
2 Supplemental results

2.1 Patient characteristics

Table A: Patient characteristics for ‘Training set’

The training set comprised 38 pediatric patients with newly diagnosed acute lymphoblastic leukemia.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Age at diagnosis</th>
<th>sex</th>
<th>race</th>
<th>WBC count</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=38)</td>
<td>(Median in years)</td>
<td></td>
<td></td>
<td>(X 10^9/L (median))</td>
</tr>
<tr>
<td>TEL-AML1</td>
<td>14</td>
<td>4.1</td>
<td>Male: 6</td>
<td>White: 8</td>
<td>15.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Female: 8</td>
<td>Black: 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hispanic: 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Other: 3</td>
<td></td>
</tr>
<tr>
<td>non-TEL-AML1</td>
<td>24</td>
<td>6.4</td>
<td>Male: 12</td>
<td>White: 16</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Female: 12</td>
<td>Black: 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hispanic: 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Other: 3</td>
<td></td>
</tr>
</tbody>
</table>

\[ P = 0.28 \] \[ P = 0.74 \] \[ P = 0.72 \] \[ P = 0.84 \]

\[^{1}\]P value determined by Fisher's exact test except age by Wilcoxon's rank sum test.
Table B: Patient characteristics for ‘Test set’

The independent “test-set” comprised 17 patients with newly diagnosed acute lymphoblastic leukemia.

<table>
<thead>
<tr>
<th>Group</th>
<th>n (n=17)</th>
<th>Age at diagnosis (Median in years)</th>
<th>sex</th>
<th>race</th>
<th>WBC count (X 10^9/L (median))</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEL-AML1</td>
<td>7</td>
<td>3.5</td>
<td>Male: 6</td>
<td>White: 6</td>
<td>8.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Female: 1</td>
<td>Hispanic: 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Other: 1</td>
<td></td>
</tr>
<tr>
<td>non-TEL-AML1</td>
<td>10</td>
<td>8.4</td>
<td>Male: 5</td>
<td>White: 8</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Female: 5</td>
<td>Hispanic: 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Other: 1</td>
<td></td>
</tr>
<tr>
<td>P value¹</td>
<td></td>
<td>P = 0.22</td>
<td>P=0.3</td>
<td>P=1</td>
<td>P =0.66</td>
</tr>
</tbody>
</table>

¹P value determined by Fisher’s exact test except age by Wilcoxon’s rank sum test.
2.2 Purine metabolism pathway (KEGG) utilized for microarray analysis

For the microarray analysis, we used 129 probe sets (81 genes) present in the Kyoto Encyclopedia of Genes and Genomes (KEGG).

Figure A: Purine metabolism graphical pathway diagram (KEGG pathway map). A rectangle represents an enzyme. The enzymes whose genes were present on the microarray and therefore included in the analysis are shown by colored (green shaded) rectangles.
Figure B: Correlation between percent of ALL cells in S-phase of the cell cycle and rate of DNPS
2.3 Clustering of patients based on gene expression patterns

Figure C: “Unsupervised” hierarchical clustering and principal component analysis (PCA)

129 purine metabolism probe sets were used for the “unsupervised” hierarchical clustering. Patients with *TEL-AML1* fusion tended to cluster together.
2.4 Gene expression analysis with a case with split TEL gene

Hierarchical clustering using one additional case that was not included in our analysis of TEL-AML1, because of an equivocal molecular diagnosis. Statistical analysis identified 18 probe sets (16 genes) that discriminated TEL-AML1 vs. non-TEL-AML1.

Hierarchical clustering using these 18 probe sets discriminated the patients in two groups, with TEL-AML1 ALL and the one case with a split TEL gene ALL (blue) clustering together.

Figure D: Hierarchical clustering and principal component analysis using 39 patients, with TEL-AML1 ALL (green, n=14), non-TEL-AML1 (red, n=24) and one with a split TEL gene (blue).