Acute lymphoblastic leukemia (ALL) is the most common malignant disease in children (1). Although event-free survival rates range between 70% and 75%, relapse of ALL remains the fourth most frequent diagnosis in childhood cancer affecting 25% of ALL patients (2). At relapse, overall cure rates of ~40% remain unsatisfactory and survival rates are particularly poor in certain subgroups (3). Treatment of relapse usually implies intensified polychemotherapy containing high-dose elements. Although in most children, second remissions can be induced high-risk patients require further treatment intensification by stem cell transplantation. Stem cell transplantation provides a better relapse-free survival rate than chemotherapy alone but is also associated with higher treatment-related morbidity and mortality rates (2). Thus, alternative treatment strategies are needed to improve the treatment and outcome of patients with relapsed ALL.

In the treatment of ALL relapse, intensity of chemotherapy and stem cell transplantation are indicated by well-established prognostic factors (summarized in ref. 2), with the most evident being the time to relapse. Early recurrence of disease with respect to frontline therapy affects ~60% of patients and is associated with a high rate of nonresponse to treatment, shorter duration of second complete remission, and a low event-free survival rate (4). Furthermore, site and immunophenotype of relapse are highly significant independent prognostic factors. Children with isolated bone marrow relapse have a poorer outcome than the 20% of patients with an extramedullary involvement (5). T-lineage immunophenotype seen in ~15% of relapse patients generally implies an inferior prognosis than B-cell precursor immunophenotype (6). In the current relapse trial of the German Berlin-Frankfurt-Münster
Gene expression profiling.

To investigate molecular determinants of the major prognostic factors at ALL relapse, we did gene expression profiling of 60 prospectively collected samples of ALL relapse patients enrolled in the current relapse trial of the BFM study group ALL-REZ BFM 2002. We specifically aimed at comparing prognostic groups according to “time of relapse,” “site of relapse,” and “MRD” to ask whether commonly denominating gene expression patterns cosegregate with these clinically distinct risk groups.

Materials and Methods

Patients and samples. To ensure homogeneous sample quality, we isolated RNA from freshly purified mononuclear cells instead of cryopreserved leukemic specimen. Therefore, we prospectively collected samples in the course of the German relapse trial ALL-REZ BFM 2002, approved by the Institutional Review Board. Written consent was obtained from parents or guardians. Diagnostic bone marrow and peripheral blood aspirates of patients with first relapse of ALL were selected to contain >60% leukemic blasts based on morphologic evaluation of smear preparations before further enrichment by Ficoll density gradient separation of mononuclear cells. On this basis, patients with isolated extramedullary relapse presenting with absent or <5% bone marrow blasts generally had to be excluded from the microarray analysis (~15% of patients). Bone marrow–derived samples were preferred over peripheral blood samples. In addition, samples were selected based on availability and high quality of RNAs.

Gene expression profiling. Ficoll density gradient purified mononuclear cells were immediately lysed in RLT buffer (Qiagen, Hilden, Germany) after preparation. RNA was isolated using the Qiagen RNeasy Mini kit, including Qiashredder and DNase on column digest. RNA quality was assessed using an Agilent 2100 Bioanalyzer (Agilent, Palo Alto, CA). Biotin-labeled cRNA (BioArray High-Yield RNA Transcription kit, Enzo Diagnostics, Farmingdale, NY) was hybridized to the Affymetrix HG-U133A microarray (Affymetrix, Santa Clara, CA). The streptavidin-phycocerythrin, biotinylated anti-streptavidin-stained arrays were scanned by the GeneChip System confocal scanner (Hewlett-Packard, Santa Clara, CA). Gene expression and clinical data as well as Affymetrix cel files have been deposited in the public Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo). The accession number assigned is GSE4693.

Microarray data preprocessing. Gene expression profiles were background-corrected and normalized on probe level using the variance stabilization method by Huber et al. (16). Normalized probe intensities were summarized into gene expression levels using the additive model described in Irizarry et al. (17) fitted by the median polish method (18). These preprocessing steps were done using implementations in Bioconductor (19). In the analysis that includes both our data and 104 previously published initial ALL samples (10), we applied the two steps described above for the joint set of raw expression data, thus renormalizing the previously published data. This ensures comparability of the two data sets.

Statistical analyses of microarray data. To rank genes according to expression differences between two predefined groups, we used a regularized t score (20). Significance cutoffs were computed using 10,000 random permutations of class labels. Quantile-quantile plots (QQ-plots) were produced to display significant differences between the distribution of original scores and the distribution of scores expected by chance. These differences are indicated by deviations of the line from the main diagonal in the plot. However, small deviations also occur due to random fluctuation. To distinguish genuine signals from random fluctuations, we calculated 200 “random QQ-plots,” comparing a single random class assignment to the complete set of random scores obtained from all permutations. We added auxiliary lines parallel to the main diagonal to the QQ-plots, such that 95% of the random QQ-plots remain closer to the main diagonal. We termed these lines confidence lines. If the line of a QQ-plot did not cross this confidence interval, we concluded that no significant differential expression of genes exists. Otherwise, we computed empirical Ps for each gene by determining the fraction of random scores above the original score and estimated false discovery rates according to Storey and Tibshirani (21). The complete procedure has been implemented in the Bioconductor compliant R package twilight (22). Lists of genes were analyzed for overrepresentation of gene ontology (GO) terms (23) using Fisher’s exact test.

In addition to association analysis, we also used predictive analysis approaches. The questions here are whether the expression profile of a leukemic sample contain enough information to predict and whether it belongs to a certain ALL subtype or to predict its fraction of S and G2-M phase cells. In both cases, we strictly distinguished between a training set on which the predictive models were learned and an independent test set on which the predictive model was evaluated. Which genes and how many of them should be included into the model was determined solely based on the training sets. For classification of microarray expression profiles into discrete ALL subtypes, we used the nearest shrunken centroid methods by Tibshirani et al. (24). The diagnostic signatures of initial ALL published by Ross et al. (10) were recalculated on the jointly normalized data. For predicting the fraction of S and G2-M phase cells in samples based on gene expression profiles, we used least angle regression (25).

Determination of cell cycle distributions by DNA flow cytometry. To determine the cell cycle distribution, frozen or freshly prepared mononuclear cells of leukemic samples (0.5×10^7–1×10^7) were washed with PBS, fixed with ice-cold ethanol (70%), and stored at 4°C until use. Before flow cytometry, cells were washed with PBS and resuspended in a volume of 200 μL. Cells were stained by adding 15 μL of 1.5 mg/mL propidium iodide (Sigma-Aldrich, Munich, Germany) and RNA was digested by adding RNase A (Sigma-Aldrich) to a final concentration of 500 μg/mL. The cell cycle distribution was measured.
Results

Subclass prediction as a quality measure of ALL relapse microarray data. Leukemic samples were obtained prospectively over a period of 18 months from children diagnosed with first relapse of ALL, enrolled in the German relapse trial ALL-REZ BFM 2002. Fifty-seven bone marrow–derived samples and 3 peripheral blood–derived samples with a mean blast count of 84% (minimum, 60%; maximum, 98% before enrichment) were selected and hybridized to Affymetrix U133A GeneChips. Detailed clinical and genetic information is given in Supplementary Table S1. Comparing the patients either included in (n = 60) or excluded from (n = 65) the microarray analysis in the same period, we observed no selection bias in terms of sex, age, immunophenotype, and frontline therapy. Due to the general exclusion of isolated extramedullary relapse (standard-risk group, S1), a significant selection in favor of isolated bone marrow relapse and, accordingly, high-risk group S4 was seen. In addition, a significant selection bias in favor of higher peripheral blast counts (PBC) at relapse diagnosis (P < 0.001) was observed. When patients with extramedullary relapse were also excluded from the not analyzed cohort, the bias was much reduced, albeit still significant (P = 0.035).

After GeneChip hybridization, we first set out to gain a measure for the quality of the microarray analysis. Therefore, we asked whether those gene expression signatures that predict genetic and immunologic subtypes in initial ALL (10, 12) could also faithfully classify relapse samples based on the present analysis. To this end, published data from 104 initial ALL patients (10) were used as training set to recalculate the diagnostic signatures of the ALL subtypes shown in Table 1. These signatures were then applied to all 28 cases of the present study belonging to these subtypes (independent test set). All except one sample were classified correctly (96% accuracy; Table 1). Reevaluation of the incorrectly classified sample suggested hybridization artifacts as the underlying cause of misclassification. Thus, in principle, meaningful results can be extrapolated from the present data set and microarray-based ALL subtype classification is stable across different stages of ALL.

Identification of genes differentially expressed between prognostic groups of ALL relapse patients. To identify molecular determinants underlying the major prognostic factors at ALL relapse, we first grouped patients according to each of these factors. Specifically, we split patient samples in a poor and a good prognosis group using the prognostic factors time of relapse, site of relapse, and MRD, respectively (Fig. 1). We then tested these groups for differences in gene expression. With respect to time of relapse, we compared 14 patients with very early relapse (within 18 months after initial diagnosis) in the poor prognosis group to 28 patients with late relapse (>6 months after cessation of frontline treatment) in the good prognosis group. This selection was undertaken to facilitate a direct comparison of the two most diverse prognostic groups. In addition, we compared 10 patients with early relapse (>18 months after initial diagnosis but <6 months after cessation of frontline treatment) having a better prognosis than very early relapsed children but still a poorer prognosis than late relapsed patients to the good prognosis group. About the site of relapse, we compared 45 patients with isolated bone marrow relapse to 7 patients with combined bone marrow relapse. Finally, according to MRD, 17 molecular poor responders with high MRD values (>1 residual leukemic cell in 1,000 cells after two induction cycles) were compared with 12 molecular good responders with low MRD values (<1 residual leukemic cell in 1,000 cells after two induction cycles). For these analyses, we focused on the 52 patients with B-lineage ALL to avoid a distortion of results solely based on gene expression differences between ALL immunophenotypes. Pronounced differences in gene expression profiles between B- and T-lineage ALL have been described before and were confirmed by our data, supporting the notion that these ALL subtypes represent distinct biological and clinical entities (10, 12).

The result of each comparison was visualized by a QQ-plot comparing the expression differences observed between two prognostic groups to differences expected by chance (Fig. 2). The main diagonal in these plots, therefore, indicates the situation that expression differences between the two groups compared do not exceed random differences. In contrast, deviations of the line from this main diagonal reflect significant differential gene expression between the two groups compared. However, a comparison of patient groups that were randomly chosen and hence presumed to display no meaningful differences in gene expression shows, albeit mild, deviations from

<table>
<thead>
<tr>
<th>Predicted*</th>
<th>BCR-ABL</th>
<th>E2A-PBX1</th>
<th>TEL-AML1</th>
<th>MLL rearranged</th>
<th>Hyperdiploid</th>
<th>T-cell ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>True†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCR-ABL</td>
<td>2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>E2A-PBX1</td>
<td>—</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>TEL-AML1</td>
<td>1</td>
<td>—</td>
<td>6</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>MLL rearranged</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Hyperdiploid</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>7</td>
<td>—</td>
</tr>
<tr>
<td>T-cell ALL</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>8</td>
</tr>
</tbody>
</table>

*Columns: predicted subtype of sample.

†Rows: true subtype of sample.
The expression of S and G2-M phase genes correlates with cell cycle distribution pointing to an increased proliferative capacity of very early relapse blasts. From the 14 very early and 28 late B-cell precursor relapse patients, only 11 and 22 samples, respectively, were available for DNA flow cytometry. Thus, cell cycle profiles could be determined for a total of 33 samples. Evaluation of these data sets revealed a significant increase of S and G2-M phase cells in very early compared with late relapse (median, 10.9% versus 5.03%; exact Ps < 0.0001, Mann-Whitney U test; Fig. 4A). We next aimed at correlating the increase in S and G2-M phases with the increase in cell cycle gene expression at the level of individual patients. We revised the list of differentially expressed genes to update the gene classification in those cases where the GO annotation lagged behind recent information in the literature. In doing so, we identified ~27% (20 of 74) of the up-regulated genes to be related to the cell cycle and cell division. As shown in Fig. 4B, an elevated proportion of S and G2-M phase cells correlates well with an increased expression level of cell cycle genes in individual patients. Only few exceptions exist, for instance, late relapse sample patient 55 shows an increased S and G2-M phase fraction but no up-regulation of cell cycle gene expression. We also included in this analysis four very early and two late T-lineage relapse samples to investigate whether they conform to the above findings for B-lineage relapse. Indeed, the six T-lineage samples support the correlation too,
indicating that an enhanced expression of cell cycle genes and an increase of the $S$ and $G_2$-M phase fraction also discriminate very early and late T-lineage relapse.

To investigate whether the expression data contains sufficient information to allow a prediction of the percentage of $S$ and $G_2$-M phase cells in individual samples, we randomly split our data into a training set ($n = 26$) and an independent test set ($n = 12$). Using exclusively the training set, we learned a predictive regression model based on 6 $S$ phase and 18 $G_2$-M phase probe sets selected from the complete list of 83 genes. The resulting prediction signature included 19 of the 24 probe sets ($5 S$ phase and $14 G_2$-M related). Its performance on the training set and the independent test set is shown in Fig. 4C. The prediction accuracy is high on both training and test data and the correlation between predicted and measured proliferation ratios is significant ($P < 0.001$). Thus, these data are consistent with the results shown in Fig. 4B and further show that the expression of the identified $S$ and $G_2$-M phase genes strongly correlates with the $S$ and $G_2$-M phase fraction of blasts at the level of individual patients. To address, whether the amount of peripheral blasts at diagnosis is influenced by the amount of $S$ and $G_2$-M phase cells, we correlated these variables (Supplementary Fig. S2A). The PBC has been used as prognostic marker in former ALL-REZ BFM trials and patients analyzed in this gene expression study show higher PBC than excluded patients. However, no positive correlation between the PBC and the fraction of $S$ and $G_2$-M phase cells was observed and accordingly no regression model to predict the PBC from gene

![Fig. 2. Quantification of differential gene expression between prognostic groups of patients with B-cell precursor ALL relapse. The degree of differential gene expression between two prognostic groups compared was visualized by QQ-plots, in which observed expression differences are plotted against differences expected by chance. Solid black diagonal line, expression differences between the two groups compared do not exceed random differences; dashed lines (confidence lines), 95% of comparisons of randomly chosen patient groups (example given in A) stay within these boundaries. B, time to relapse: 14 very early versus 28 later elapsed patients. C, site of relapse: 7 patients with combined bone marrow relapse versus 45 with isolated bone marrow relapse. D, MRD status after two induction cycles within the intermediate-risk group: 17 patients with >1 leukemic cell in 1,000 versus 12 with ≤1 in 1,000.](image-url)
Fig. 3. Genes differentially expressed between very early and late relapse of ALL and GO terms overrepresented within this gene list. A, expression levels of 87 probe sets (rows) identified by comparison of very early and late relapse of ALL in individual patients (columns). Expression levels are expressed as fold changes with respect to the mean overall patients (n = 52). Top, red/green bar, samples representing very early or late relapse. B, parts of the GO biological process organizing principle that contain terms (nodes) significantly overrepresented within the 87 probe sets shown above. GO nodes are colored according to their overrepresentation P values; red highlight, nodes overrepresented among the probe sets up-regulated in very early relapse; green highlight, those overrepresented among the down-regulated genes.
Discussion

We did a genome-wide transcription profiling study of 60 children, with first relapse of ALL enrolled in the relapse trial ALL-REZ BFM 2002 of the BFM study group. Genetic and immunologic subclasses described by gene expression profiling studies of initial ALL (10, 12) were correctly predicted from microarray data in relapse patients, thus proving consistency of microarray-based leukemia subtype classification across different stages of disease. To identify molecular determinants of the major prognostic factors at ALL relapse, we compared prognostic groups of B-cell precursor ALL relapse classified according to each of these factors. No significant gene expression patterns were found to correlate with the prognostic factors site of relapse and MRD. About the most evident prognostic factor time of relapse, we identified significant differences in gene expression between patients with very early and late relapse. We obtained a list of 83 differentially expressed genes...
genes mostly up-regulated in very early relapse, and in this list, genes coding for proteins with functions in cell cycle and cell division and in particular in mitosis were significantly overrepresented.

The significance of the changes in gene expression observed was determined by a confidence interval–based analysis that enabled us to delineate relevant changes between two groups from random noise. According to these stringent criteria, not more than random differences were detected between prognostic groups according to site of relapse and MRD. In case of site of relapse, this might be ascribed to the small number of patients with combined bone marrow relapse (n = 7) compared with 45 patients with isolated bone marrow relapse. On the other hand, the site of ALL relapse and MRD might depend on mechanisms that are not well reflected by an informative gene expression pattern. In initial childhood ALL, an expression signature has been identified recently that distinguishes between patients with high and low MRD (26). However, the prognostic value of MRD depends on criteria that strongly differ between frontline and relapse treatment (7, 27). Furthermore, at first presentation, MRD is used for general-risk group stratification, whereas, at relapse, MRD is solely determined in the intermediate-risk group to indicate stem cell transplantation. Thus, the groups of patients compared strongly differ between the study of Cario et al. (26) and the present analysis. Still, the lack of a MRD signature at relapse suggests distinct molecular mechanisms determining MRD in initial and relapsed ALL.

In samples of patients with very early relapse, we identified a gene expression pattern that shows a significant overrepresentation of genes relevant in the late cell cycle. Consistently, we observed a significant increase of S and G2-M phase cells in very early relapse samples using DNA flow cytometry. A predictive regression model based on the expression of late cell cycle genes accurately predicted the fraction of S and G2-M phase cells in a given sample. This supports the view that the S and G2-M phase fraction of blasts contributes significantly to the enhanced expression of these genes, suggesting that the proliferative capacity of ALL blasts is increased in very early relapse patients. Although we analysed only a few T-lineage samples, the evidence obtained supports that T-cell ALLs also conform to this model. Thus, an increased proliferative capacity may constitute an overall biological feature of very early relapse of ALL that is typically associated with a particularly low event-free survival rate (4).

Interestingly, at first presentation of childhood ALL, the prognostic value of the fraction of proliferating cells has remained controversial and there is no clear evidence for an association with outcome (see ref. 28 and references therein). However, in in vitro experiments, an increased proliferation rate was associated with good response to treatment (28). In initial ALL, a higher fraction of proliferating cells seems to result in a higher efficacy of chemotherapeutic drugs targeting the cell division cycle. Consistently, gene expression profiling studies of initial ALL identified cell cycle genes to be up-regulated in the good responsive group (26, 29). At ALL relapse, however, we consider leukemic blasts that have escaped frontline treatment and progressed in malignancy reflected by higher rates of poor or nonresponse to treatment in particular in very early relapsed patients (4). Accordingly, higher proportions of cycling cells did not correlate with in vitro chemosensitivity at ALL relapse (28). At relapse, leukemic blasts may have gained signaling pathways that enable a more aggressive growth of cancer cells. This hypothesis needs to be validated by comparing gene expression profiles of paired initial ALL and relapse samples obtained from the same patient in a future study.

In breast cancer, an increased proliferation rate is strongly associated with an increased risk of metastasis and death (30). Remarkably, gene expression profiling has also identified recently an up-regulation of cell cycle and cell proliferation relevant genes in a distinct poor prognosis subgroup of patients in this tumor entity (31). In particular, there is a considerable overlap in genes or gene families (CCNB2, CDCA5L, MAD2L1, KIF23, PTTG2, ORC6L, and KNSL6) identified between the current and the breast cancer study (31). Moreover, similar results have been obtained recently for aggressive neuroblastoma (32) and poor prognosis lung squamous cell carcinoma (33). Therefore, a cell cycle or cell proliferation signature might be a common theme of aggressive or in particular of advanced stages of several malignancies associated with poor therapy outcome.

Interestingly, it has been noted earlier that the proportion of cycling cells is increased in relapsed versus initial ALL (28). Our data now reveal that this mainly applies to very early relapse, whereas late relapse displays an even lower fraction of cycling cells in the bone marrow than samples at first presentation (median, 3.9% S phase cells at late relapse versus 6.9% at first presentation in the study of Kaaik et al.; ref. 28). In contrast to very early relapse, we found that only a very small number of genes is commonly deregulated in late relapse. This suggests that very early ALL relapse represents a mechanistically more homogenous subgroup compared with late relapse and is consistent with the view of more diverse mechanisms resulting in recurrent disease when following a longer event-free period.

A large fraction of cell cycle genes identified in our study code for proteins involved in mitosis. These are molecules of the chromosomal passenger complex (CDC28 and Aurora kinase B; ref. 34) of the spindle checkpoint (MAD2L1 and PTTG1; ref. 35), spindle components (Spc25; ref. 36), spindle assembly and organization proteins (TPX2 and NUSAP1; refs. 37, 38), and several members of the kinesin superfamily of mitotic motor proteins involved in spindle function and cytokinesis (KIFC1, KIF23, KIF11, and KNSL7; refs. 39–43). Up-regulation of these genes opens up the possibility of additional therapeutic strategies considering novel antimitotic drugs in the treatment of patients with very early relapsed ALL. For example, a novel inhibitor of Aurora kinases, such as Aurora kinase B as identified in this screen, has been shown recently to be effective on leukemia cell lines and on primary cells of patients with acute myelogenous leukemia refractory to standard therapy (46). Kinesin motor protein inhibitors designed to circumvent resistance mechanisms and cytotoxicity of conventional antimitotic drugs (47) may represent an alternative treatment option. Whether patients would benefit from novel antimitotic drugs needs to be carefully elucidated, but our findings indicate that it might be worthwhile to consider these therapeutic strategies to improve the response and outcome for patients with very early relapse of ALL that face a particularly dismal prognosis.

Acknowledgments

We thank the study participants for enabling this research.
Molecular Characterization of Relapsed Childhood ALL

References


Expression of Late Cell Cycle Genes and an Increased Proliferative Capacity Characterize Very Early Relapse of Childhood Acute Lymphoblastic Leukemia

Renate Kirschner-Schwabe, Claudio Lottaz, Jörn Tödling, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/12/15/4553

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2006/08/08/12.15.4553.DC1

Cited articles
This article cites 45 articles, 19 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/12/15/4553.full.html#ref-list-1

Citing articles
This article has been cited by 9 HighWire-hosted articles. Access the articles at:
/content/12/15/4553.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.