Gene expression changes associated with progression and response in chronic myeloid leukemia

Jerald P. Radich†‡, Hongyue Dai§, Mao Mao§, Vivian Oehler*, Jan Schelter§, Brian Druker†¶, Charles Sawyers‡, Neil Shah†, Wendy Stock**, Cheryl L. Willman†, Stephen Friend§, and Peter S. Linsley§

†Divisions of Clinical Research, Fred Hutchinson Cancer Research Center, Seattle, WA 98109; §Rosetta Inpharmatics, Seattle, WA 98109; Oregon Health & Science University, University Center, Portland, OR 97239; University of California, Los Angeles, CA 90095; **University of Chicago School of Medicine, Chicago, IL 60637; †University of New Mexico Cancer Research and Treatment Center, Albuquerque, NM 87131; and §Southwest Oncology Group, Ann Arbor, MI 48109

Communicated by E. Donnall Thomas, Fred Hutchinson Cancer Research Center, Seattle, WA, December 13, 2005 (received for review June 24, 2005)

Chronic myeloid leukemia (CML) is a hematopoietic stem cell disease with distinct biological and clinical features. The biologic basis of the stereotypical progression from chronic phase through accelerated phase to blast crisis is poorly understood. We used DNA microarrays to compare gene expression in 91 cases of CML in chronic (42 cases), accelerated (17 cases), and blast phases (32 cases). Three thousand genes were found to be significantly (P < 10−10) associated with phase of disease. A comparison of the gene signatures of chronic, accelerated, and blast phases suggest that the progression of chronic phase CML to advanced phase (accelerated and blast crisis) CML is a two-step rather than a three-step process, with new gene expression changes occurring early in accelerated phase before the accumulation of increased numbers of leukemia blast cells. Especially noteworthy and potentially significant in the progression program were the deregulation of the WNT/β-catenin pathway, the decreased expression of Jun B and Fos, alternative kinase deregulation, such as Arg (Abi2), and an increased expression of PRAME. Studies of CML patients who relapsed after initially successful treatment with imatinib demonstrated a gene expression pattern closely related to advanced phase disease. These studies point to specific gene pathways that might be exploited for both prognostic indicators as well as new targets for therapy.

Conflict of interest statement: No conflicts declared.

Freely available online through the PNAS open access option.

Abbreviations: CML, chronic myeloid leukemia; CCR, complete cytogenetic remission.

Data deposition: The sequence reported in this paper has been deposited in the Gene Ontology (GO) database (accession no. GSE4170).

†To whom correspondence should be addressed. E-mail: jradich@fhcrc.org.

© 2006 by The National Academy of Sciences of the USA

PNAS | February 21, 2006 | vol. 103 | no. 8 | www.pnas.org/cgi/doi/10.1073/pnas.0510423103
statistical significance cutoff of $P < 10^{-11}$ (Fig. 1 and Table 4, which is published as supporting information on the PNAS web site). This set of genes identified in the progression from chronic to blast phase is referred to here as the “phase reporter” gene set.

We examined the proposed biologic function of the genes associated with CML phases by applying a biological annotation program based on the gene ontology (GO) and KEGG annotations. The major functional groups of genes in the phase reporter gene set are shown in Table 5, which is published as supporting information on the PNAS web site. The functional groups most highly correlated with disease phase (accelerated/blast phase relative to chronic phase) included increased expression of nuclear genes, mitochondrial genes, RNA-binding genes, and protein biosynthesis genes, reflecting the increased proliferation and metabolism of progressive disease. Advanced-phase CML, compared to chronic phase, exhibited a decreased expression of genes involved in structural integrity and adhesion, as well as decreases in expression of genes involved in inflammatory and immune response. In addition, several protooncogenes and tumor suppressor genes are differentially expressed in advanced phase CML, including N- and H-ras, FLT3, yes, AF1q, CBFB, WT1, ORALOV1, Bcl-2, and PTPN11.

CML Appears to Be a “Two-Step” Disease. We examined whether the progression of CML best fit a three-step model as often described.

Fig. 1. Genes associated with CML progression. Samples from patients with of CML cases in chronic phase, accelerated by cytogenetic criteria only, accelerated phase, blast crisis, and blast crisis “in remission” were compared to a pool of chronic phase RNA (See Materials and Methods for details). Approximately 3,500 genes were significantly associated with progressive disease at a significance level of $P < 10^{-11}$. Each row represents one sample, and each column represents one gene. Red color indicates overexpression relative to the control pool, and green color indicates low expression.

Fig. 2. Comparison of CML blasts to normal CD34$^+$ cells and correction for the pattern of gene expression. (A) Phase genes were corrected for normal CD34$^+$ gene expression (ANOVA, $P < 1 \times 10^{-9}$). The gene expression of normal CD34$^+$ cells was subtracted from each disease sample. The resulting pattern reflects genes associated with progression independent of normal blast biology. (B) Relative gene expression of the “top ten” genes found to be up-regulated in advanced phase disease compared to chronic phase.
Identification of “Progression-Specific” Gene Expression. We next investigated how the phase reporter gene set was influenced by the gene expression signature of leukemia blasts and how these CML blasts compared to normal immature CD34+ cells. First, we made a direct comparison of the gene expression signature of six samples of normal CD34+ cells with CML blast samples containing >70% blasts (Fig. 5, which is published as supporting information on the PNAS web site). We found that the gene expression pattern of CML blast cells was very similar to normal CD34+ blasts (Fig. 5, which is published as supporting information on the PNAS web site). The correlation of expression level between accelerated and blast phase gene expression was strong (r = 0.81). In addition, the correlation of gene expression between accelerated phase and cases of accelerated phase defined by new cytogenetic abnormalities alone (i.e., chronic phase morphology but additional chromosomal changes besides the Ph) was also high (r = 0.61). These observations suggest that the difference of gene expression between accelerated and blast phase is a quantitative change in expression rather a qualitative change.

**Table 1. Functional annotation of “progression” genes**

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Number (%)*</th>
<th>Examples</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribosome</td>
<td>18 (21)</td>
<td>ROK13A</td>
<td>9 × 10^{-11}</td>
</tr>
<tr>
<td>Wnt signaling</td>
<td>16 (11)</td>
<td>Cadherin, MDI1, Prickle 1, Fzd2</td>
<td>2 × 10^{-5}</td>
</tr>
<tr>
<td>Nucleosome</td>
<td>22 (22)</td>
<td>BZ1A, HIST1H2AE</td>
<td>3 × 10^{-11}</td>
</tr>
<tr>
<td>Sugar</td>
<td>45 (26)</td>
<td>RPIA, ALD0C</td>
<td>4 × 10^{-9}</td>
</tr>
<tr>
<td>Metabolism</td>
<td>27 (14)</td>
<td>G6PD, CEBPA, CEBPE, FOXO3A</td>
<td>3 × 10^{-4}</td>
</tr>
<tr>
<td>Myeloid differentiation</td>
<td>42 (10)</td>
<td>GADOD45G, BCL2, FOXO3A, MCL1</td>
<td>2 × 10^{-7}</td>
</tr>
<tr>
<td>Apoptosis</td>
<td>36 (10)</td>
<td>GASOD45G, FANC, XRN2</td>
<td>2 × 10^{-7}</td>
</tr>
</tbody>
</table>

Bold functions are up-regulated in progression; italic functions are down-regulated.

*The number of genes present on the array with the specific keyword; in parentheses are the percentage of significant genes found differentially expressed among the keyword class.

MZF- and EF1α-_controlled Genes Are Aberrantly Expressed in Progression. Gene expression networks defined by specific promoter regulation were analyzed in the progression and phase reporter gene sets (Table 2). The most significantly network of genes showing aberrant control in progression with those that contained a MZF promoter or a EF1α promoter sequence (P < 10^{-11} and <10^{-11}; respectively, Figs. 8 and 9, which are published as supporting information on the PNAS web site). Also significantly associated with progression were genes bearing SPI-B, Yin Yang, and AHR-ARNT promoter sequences (all with P values <10^{-9}).

Gene Expression Profiles of Cases Treated with Imatinib. We studied the gene expression of CML patients for whom imatinib therapy was ineffective. We examined 15 cases of CML treated with imatinib, including nine patients who initially achieved a CCR on imatinib therapy, but then relapsed back into an apparent chronic phase by morphologic examination; three cases who had achieved a complete hematologic but no cytogenetic response; and three late chronic phase patients. All but one of these chronic phase patients who relapsed after a CCR had a point mutation in abl1, presumably abrogating imatinib activity (Table 3). In addition, only one of the relapsing cases had additional cytogenetic lesions.

**Table 2. Promoter sequences differentially represented in progression**

<table>
<thead>
<tr>
<th>Promoters</th>
<th>P value in progression genes</th>
<th>P value in phase genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>MZF-1</td>
<td>&lt;10^{-15}</td>
<td>&lt;10^{-15}</td>
</tr>
<tr>
<td>EF1α</td>
<td>9 × 10^{-11}</td>
<td>1 × 10^{-9}</td>
</tr>
<tr>
<td>SPI-B</td>
<td>1 × 10^{-10}</td>
<td>2 × 10^{-11}</td>
</tr>
<tr>
<td>Yin Yang</td>
<td>1 × 10^{-9}</td>
<td>7 × 10^{-11}</td>
</tr>
<tr>
<td>AHR-ARNT</td>
<td>2 × 10^{-9}</td>
<td>5 × 10^{-10}</td>
</tr>
</tbody>
</table>
However, there were gene expression features that were unique to the imatinib relapers. Examples of genes unique to imatinib resistance failure are those of serine threonine kinases (CTRL, MAP21K14, CLK3), mitogen-activated protein kinase (MKNK2), the tyrosine kinase oncogene FYN, and the drug exporter ABCC3. Fig. 10, which is published as supporting information on the PNAS web site, compares relapsed imatinib cases against blast crisis cases, and several areas (boxed) indicate genes that are expressed in reverse direction in these two states (Tables 12 and 13, which are published as supporting information on the PNAS web site, lists genes in the first box. See Table 12 for the rest of the imatinib resistance genes).

Discussion
The biology and treatment of CML is dictated by the phase of disease, as the efficacy of all therapies (transplantation, IFN, or imatinib) are most successful when used during the chronic phase, in comparison to accelerated phase and blast crisis.

Understanding the biology of progression may provide clinical diagnostic markers of progression and offer insights into new strategies for treatment. The data presented here suggest that the progression of chronic phase CML to advanced phase CML is a two-step process, with progression associated a block of differentiation and apoptosis, a shift toward turning on expression of genes involved in the nucleosome, with alterations in cell adhesion, and activation of alternative signaling pathways. In addition, it appears that relapse after initial successful treatment with imatinib may be associated with gene expression patterns similar to advanced phase CML, suggesting that the process of progression persists in a subpopulation of CML cells during therapy.

In our analysis, we first used unsorted CML samples to develop gene signatures associated with the phase of disease. This data set gives a more global picture of the transformation of disease, and is important because any diagnostic assay for progression would optimally be performed on whole blood or marrow, rather than selected cell subpopulations. To control for any gene signatures that might be common to an immature blast phenotype, we subtracted a normal CD34+ signature from the CML samples. This left us with a progression gene set that should be highly enriched in genes associated with CML progression, but not reflecting normal CD34+ cell gene expression. This gene set may be more indicative of the biology of CML progression, and give more insight to pathways that might be exploited by new therapeutic agents. Moreover, the progression gene set may contain genes able to distinguish early evidence of progression in immunophenotypically similar immature cells.

The demonstration that the gene expression pattern between accelerated and blast phases are very similar suggests that the crucial steps in progression are at the transition of chronic to accelerated phase, before obvious morphologic, cytogenetic, or clinical evidence of progression. This has clinical implications, because these patients might benefit from more aggressive therapy. In addition, the observation that gene signatures of blast crisis can be seen in accelerated phase patients by cytogenetic criteria only and blast crisis cases in remission (both of which have similarly low blast counts as patients with chronic phase disease), demonstrates two important points. First, it points out the difficulty of correlating...
morphology with the biology of the disease. Secondly, the penetration of a progression gene expression signature into a “chronic phase” appearing bone marrow suggests that progression is not merely an absolute block of differentiation, but that abnormal gene expression signals persist in normal appearing differentiated cells. This is critically important because it provides the basis for testing for progression genes in unsorted bone marrow samples from “chronic phase” patients.

Bcr-Abl has a wholesale range of biological activities. Critical in the transformation process is the activation of the Ras/mitogen-activated protein kinase (MAPK) pathways, which has broad effects on changes in cell adhesion (through Rho), proliferation (MAPK pathway), and apoptosis (through Akt) (16, 17). Imatinib’s efficacy results from a blockade of these effects. Imatinib is less successful for advanced-phase disease; this may be because these leukemia have become less dependent on the pathways that imatinib blocks. Thus, we found that the MAPK pathway was relatively underexpressed in advanced disease compared to chronic phase, but other signaling pathways, including those involving cytokines (IL3RA, SOCS2), alternative ras pathways (Rras2), and those involved in cell adhesion (WNT/β-catenin) are activated. The activation of these pathways may allow progression even in the face of therapeutic blockade of Bcr-Abl activated pathways. In addition, abl2 (Abl related gene, or ARG) is also up-regulated in progression. As opposed to abl1, ARG is a cytoplasmic protein (18). The signaling targets of ARG are unknown, and although broadly expressed in tissue, its only known functional role apparent from knockout mouse models appears to be in the nervous system (19, 20). ARG has been associated with myeloid leukemia in the context of TEL/ARG translocations (21, 22). ARG shares >90% homology of its tyrosine kinase domain with abl1, and ARG tyrosine kinase activity is inhibited by imatinib at similar drug concentrations (18, 23). However, given that Bcr-Abl amplification is considered to play a role in imatinib resistance (24, 25), ARG overexpression in blast crisis could theoretically contribute to the relative resistance to imatinib found with progressive disease, either as independently acting on aberrant signaling, or acting as an imatinib “sink.”

Two recent observations on the molecular biology of progression in CML are relevant to this study. First, activation of the WNT/β-catenin pathway was observed in primary cell samples from patients with CML (26). Secondly, it was recently observed that mice deficient in Jun B develop a disease much like CML (27). Our data complement these findings, as we found broad dysregulation of WNT/β-catenin pathway as well as decreased Jun B expression. A link between these pathways may be the gene MDFI (I-mfa), an inhibitor of myogenic basic helix–loop–helix transcription factors (28). MDFI interacts with axin, which is involved in binding and modulating free β-catenin, and thus changes in MDFI may influence β-catenin mediated gene activation (29). In addition, MDFI and axin interaction also influences WNT and Jun signaling (30).

Our analysis suggests that genes controlled by MZF1 and EF1α may be particularly important in progression. MZF1 is a member of the Kruppel family of zinc finger proteins, and was originally cloned from a cDNA library from a blast crisis CML patient (31). MZF1 appears to play a critical role in hematopoietic stem cell differentiation, including modulation of CD34 and c-myc expression (32, 33). MZF1 in knockout mice display an increase in hematopoietic progenitor proliferation, which continues in long-term culture conditions (32). These data suggest MZF1 deregulation may disrupt normal differentiation, promoting the progression to advanced disease. EF1α is related to the Smad zinc finger proteins that play an important role in TGF-β gene regulation. EF1α has been shown to compete with basic helix–loop–helix activators, and is implicated in modification of MyoD regulated pathways (34, 35). It is not known whether EF1α directly influences MDFI expression. Of note is that both MZF1 and EF1α have been shown to influence cadherin expression (36–38). Thus, the further study of the control of MZF1 and EF1α may be particularly fruitful in understanding the molecular mechanisms of CML progression.

Given that efficacy of treatment in CML (be it with IFN, imatinib, or transplantation) is so intimately associated with phase of disease, one wonders whether those chronic phase patients in whom imatinib therapy is ineffective have genetic features of advanced phase invisible to routine pathological and cytogenetic examination. Imatinib failures are a reasonable setting to explore this possibility. Although imatinib can cause cytogenetic remissions in the majority of chronic phase cases, treatment failure, especially secondary to point mutations, is an increasingly important problem. It has previously been demonstrated that the probability of developing a point mutation depends largely on the time from diagnosis to initiation of therapy (8). This finding implies that the genetic mechanisms that lead to point mutations are relentless, and therefore the treatment of “late” chronic phase patients (i.e., > 1 year from diagnosis) may be undermined by genetic changes that have already occurred. Branford et al. (8) demonstrated that patients who developed P-loop point mutations had a very poor outcome, with approximately half dying within a year of relapse. These observations are in keeping with our demonstration that many imatinib failures have gene expression changes similar to advanced disease, despite their benign pathological appearance. The predictive power of genetic markers of response and progression will be important to test prospective in future clinical studies of imatinib and similar compounds.

Several of the genes found in the progression set might serve as early markers of progression in diagnostic assays, and may serve as therapeutic targets, as well. For example, PRAME was originally identified as a tumor antigen recognized by cytotoxic T cells against a melanoma surface antigen (13, 39). Like similar antigens MAGE, BAGE, or GAGE, PRAME is expressed in some solid tumors; unlike these other antigens, however, it has been found to be overexpressed in >25% of leukemia, and has been found to be induced by Bcr-Abl in CML cell lines (14, 40). Indeed, PRAME overexpression has been described as one of the few features that characterize the transient myeloproliferative syndrome of Down’s syndrome from the progressive acute megakaryoblastic leukemia found in that disorder (41). Very recently, insights into the function of PRAME make it an attractive target for potential therapy. PRAME is a nuclear protein, which seems to act as a dominant repressor of retinoic acid receptor (RAR) signaling (42). In cell line models, the forced overexpression of PRAME blocked RAR-mediated differentiation, growth arrest, and apoptosis. Thus, therapeutic maneuvers to block the PRAME/RAR effect may allow for differentiation and cell death, and be particularly effective in attacking advanced phase CML.

In comparison to other types of leukemia, there have been few papers exploring the use of microarrays on the biology of CML. It is difficult to make a direct comparison of these studies and ours, given the different types of samples obtained (some unsorted, some AC133+ or CD34+ selected), the difference in the array platforms, etc. Nonetheless, in general, the functional changes of progression (changes in differentiation, apoptosis, cell adhesion, and inflammatory response) remained as common themes across these studies. Moreover, in keeping with a recent publication that found that the high expression of elastase (ELA2) was associated with a long period of indolent disease (45), in our data, ELA2 decreased expression is associated with progression.
Functional Annotation of Gene Lists. Genes represented on the microarray were annotated by assignment to GO Biological Process or Molecular Function categories (www.ebi.ac.uk/GOA), or to KEGG pathways (www.genome.jp/kegg/pathway.html). Gene lists (input sets) were queried to enrichment for members of specific functional classes or pathways relative to the background frequency. The significance ($P$) value of enrichment was computed by using the hypergeometric probability distribution. Reported in each case are the number of genes in the input set (inlying gene count), number of genes in a particular category or pathway in the input set (overlap gene count), and number of genes in a particular category among all genes present on the array (set gene count). The total number of unique genes on the array is 24,132.

Methods to Common Promote Site Analysis. The common promoter sites were based on the predictions derived from the database (oPSSUM) by Wasserman et al. (www.cisreg.ca). The hypergeometric $P$ value for enrichment of a particular binding site was computed by comparing the number of genes with the binding site from a signature gene set to that from a background set (i.e., all genes represented on the microarray).

Controls. We compared gene expression signatures from the sites contributing samples to confirm that there were no site-specific confounding the analysis.

Because some samples of blast crisis came from peripheral blood rather than bone marrow, we compared three samples in which simultaneous samples were available from bone marrow and peripheral blood. Gene expression was extremely well correlated ($r = 0.97–0.99$; Fig. 11, which is published as supporting information on the PNAS web site).

This work was supported in part by National Institutes of Health Grants CA-18029 and CA-85053 (to J.P.R.) and the Howard Hughes Medical Institute (B.D. and C.S.).
Supplemental Figure 1. Comparison of gene expression of accelerated and blast phase CML
Supplemental Figure 2. Comparison of blast crisis and normal CD34+ gene expression
Supplemental Figure 3. Expression differences between CML blasts and normal CD34+ cells
Supplemental Figure 4. WNT/B-catenin pathway involvement in CML
Supplemental Figure 5. Gene expression of genes with MZF promoter sequences
Supplemental Figure 6. Gene expression of genes with delta EF1 promoter sequences
Supplemental Figure 7. Gene expression differences between imatinib failures and CML blast crisis

IM vs. BC
ANOVA P < 1e-5
Supplemental Figure 8. Correlation of bone marrow and peripheral blood expression

r = 0.96

Bone Marrow

Peripheral Blood

r = 0.97

r = 0.99