Clinical and Biological Relevance of Gene Expression Profiling in Acute Myeloid Leukemia

Alicja M. Gruszka¹ and Myriam Alcalay¹,²

¹Istituto Europeo di Oncologia; ²Università degli Studi di Milano
Italy

1. Introduction

Over the last decade, considerable effort has gone into defining global gene expression profiles (GEP) in many different types of malignancies. There is a dual aim behind these studies: on the one hand, to identify molecular signatures that correlate with clinically useful parameters and, on the other hand, to increase knowledge concerning the biology of the respective diseases. Some of these studies yielded molecular classifications of specific cancer types that better correlate with disease progression and/or response to therapy, whereas others revealed yet unknown biological properties of cancer cells that may represent the starting point for novel therapeutic approaches.

Acute myeloid leukemias (AML) represent a highly heterogeneous set of malignancies whose pathogenesis is linked to specific genetic abnormalities, including chromosome translocations and point mutations that involve genes encoding for key regulators of hematopoiesis (Marcucci, Haferlach, & Dohner 2011). Genetic information is the most relevant parameter for the correct classification of AML patients at diagnosis into three prognostic risk groups (favorable, intermediate and adverse) and, consequently, for directing therapeutic choices (Lo-Coco et al. 2008; Dohner et al. 2010). In fact, current diagnostic approaches include cytogenetic and molecular analyses for correct stratification of AML patients according to the World Health Organization (WHO) recommendations (Vardiman et al. 2009). However, these approaches are not fully satisfactory, particularly within the significant group of cytogenetically normal AML (CN-AML), where known prognostic markers are lacking. This likely reflects the genetic heterogeneity of the CN-AML group, the existence of yet unidentified genetic lesions and the co-existence of different genetic mutations in a significant number of cases.

AML was the first type of malignancy to be studied with a GEP approach, and hundreds of reports addressing specific clinical issues (classification, prognosis, response to therapy) have been published. An equally significant number of studies have addressed the molecular pathogenesis of AML by analyzing GEP in functionally characterized AML model systems, including transgenic mice, primary or established cell lines expressing leukemogenic oncogenes. The analysis of their specific target genes has been exploited to
unravel the functional consequences of AML-associated oncogene expression, including the arrest of myeloid differentiation and enhanced cell survival. In recent years, the analysis of AML has been extended to other genomic approaches, including microRNA profiling and epigenetic studies. The rapid technological advancement, and in particular the advent of next-generation sequencing has imposed a dramatic change of outlook to the molecular basis of cancer, including AML, and it is likely that GEP approaches so far used will become obsolete, favoring more focused, clinically relevant expression studies.

What have GEP studies taught us about AML? We here propose an overview of the progress that has been made through GEP both in terms of clinical utility and of insight to the biology of the disease, and discuss future perspectives (Figure 1).

![Gene expression profile in AML (microarray technology)](image)

**Fig. 1.** Main points covered in this review.

### 2. Clinical utility of GEP analysis in AML

The first evidence that GEP could be employed as a tool for the correct classification of cancer was reported by Golub and collaborators in 1999, using acute leukemias as a test case (Golub et al. 1999). The authors were able to discriminate AML samples from acute lymphoblastic leukemia (ALL) samples without prior knowledge concerning the respective diagnosis, and suggested two important applications for GEP: “class discovery”, which refers to the identification of new prognostically relevant tumor subtypes, and “class prediction”, which assigns tumor samples to already known subtypes on the basis of their specific gene expression signature. Successive studies introduced the possibility to exploit GEP for predicting response to therapy (“outcome prediction”) (Theilgaard-Monch et al. 2011). Technically, class discovery implies the search for significant similarities and differences in a cohort of samples, assuming that similar gene expression signatures will correspond to the same disease subtype, and relies on an unsupervised approach (i.e. no
prior knowledge of patient characteristics, such as age, cytogenetics, molecular abnormalities, etc.). Class prediction, instead, takes into account patient information to derive gene expression signatures that are specific for given parameters, which can then be used for predicting disease subtypes in samples of unknown status (i.e. supervised approach).

In de novo AML, chromosomal abnormalities can be detected at diagnosis in approximately 55% of cases by cytogenetics analysis, and specific genetic mutations can be identified in 85% of the remaining CN-AML. The most frequent chromosomal rearrangements include reciprocal translocations and inversions, such as t(8;21), which fuses the AML1 and ETO genes, inv(16), which results in the CBFβ/MYH11 chimeric gene, t(15;17), which generates the PML/RARα fusion specific of acute promyelocytic leukemia (APL), and a variety of translocations involving the MLL gene on chromosome 11q23, the most frequent being t(9;11) (Look 1997). The resulting fusion proteins possess oncogenic properties and frequently function as transcriptional regulators (Alcalay et al. 2001). It is therefore perhaps not surprising that AML blasts bearing such rearrangements display specific gene expression signatures, as demonstrated by several studies (Bullinger et al. 2004; Debernardi et al. 2003; Schoch et al. 2002; Valk et al. 2004). In fact, GEP can actually predict favorable cytogenetic AML subtypes, i.e. t(8;21), t(15;17) and inv(16), with 100% accuracy, and with >90% accuracy for AML with MLL rearrangements (Haferlach et al. 2005b; Ross et al. 2004), whereas the correlation is less stringent for other molecular subtypes (Verhaak et al. 2009).

The function of group-specific genes often reflects characteristics of the corresponding disease: for example, the t(15;17) signature of APL, which is clinically characterized by a hemorrhagic diathesis and a response to treatment with Retinoic Acid (RA), includes genes involved in hemostasis and suggests an impairment in the response to RA-induced differentiation (Bullinger et al. 2004). Interestingly, GEP can segregate APL cases (M3 subtype) from variant cases (M3v), which are characterized by a specific morphology of blasts and by a more severe prognosis (Haferlach et al. 2005a). The genes that are differentially expressed between APL-M3 and APL-M3v encode for functions such as granulation and maturation of blood cells, which are coherent with the morphological and clinical observations.

Specific somatic mutations are frequent events in AML, particularly in CN-AML, which are classified in the group with intermediate risk even though there are important differences. Currently, only mutations of the NPM1, CEBPA and FLT3 genes have an impact on the clinics because of their correlation with prognosis. Other recurrent mutations in AML include N-RAS, KIT, IDH1, IDH2, WT1, RUNX1 and MLL, but their clinical significance is either controversial or unknown, and their identification is at the moment not used for guiding therapeutic choices. Although these mutations are prevalent in CN-AML, they can also be found in association to other cytogenetic abnormalities (Marcucci, Haferlach, & Dohner 2011).

Specific GEP signatures have been described for AML that carry mutations, but their predictive accuracy appears to be lower than the one described for the recurrent cytogenetic rearrangements described above (Verhaak et al. 2009), likely due to the frequent co-occurrence of different genetic abnormalities or to the presence of other yet unknown mutations. Mutations in the NPM1 gene represent the most frequent genetic abnormality in CN-AML (Falini et al. 2005), and are associated to mutations of the FLT3 gene in a significant proportion of cases. AML with mutated NPM1 without concurrent FLT3
mutations are characterized by a better response to induction therapy and a favorable prognosis (Falini et al. 2007). These cases present a gene expression signature characterized by over-expression of HOX and TALE genes (Alcalay et al. 2005; Verhaak et al. 2005). Other genes involved in maintenance of hematopoietic stem cells, such as the NOTCH ligand JAG1, are also over-expressed, suggesting that the cell of origin of AML with mutated NPM1 may be an early hematopoietic progenitor, as further indicated by frequent multilineage involvement (Alcalay et al. 2005; Pasqualucci et al. 2006).

AML with CEBPA mutations are also associated with a favorable outcome. Generally, AML in this group carry mutations in both CEBPA alleles, whereas heterozygous mutations are less frequent. A specific GEP signature has been described for AML with biallelic CEBPA mutations, while no discriminating gene expression pattern was detectable in AML carrying single mutations (Wouters et al. 2009). Of note, in this study, only AML with biallelic CEBPA mutations correlated with a favorable outcome, whereas single mutations did not, suggesting a prognostic value for the specific gene expression signature. Interestingly, in a group of CN-AML patients displaying a GEP signature resembling that of AML with mutant CEBPA, but lacking such mutations, the CEBPA gene was found to be silenced as a consequence of promoter hypermethylation (Wouters et al. 2007). This result suggests that GEP may actually be more efficient than mutational analysis in identifying functional pathways that are perturbed in specific AML cases, and may be a useful tool for correct molecular classification of AML.

Two types of mutations involving FLT3 can be found in AML: internal tandem duplications (FLT3-ITD), which are present in 20% of AML, and point mutations within the tyrosine kinase domain (FLT3-TKD) that can be detected in an additional 5-10% of cases. FLT3-ITD are associated with a poor prognosis in CN-AML, whereas the prognostic relevance of FLT3-TKD is not clear (Marcucci, Haferlach, & Dohner 2011; Mrozek et al. 2007). GEP analyses in AML with mutated FLT3 have yielded controversial results. One study described an accurate separation of samples with FLT3-ITD from those with FLT3-TKD (Neben et al. 2005), while another study reported a specific gene expression signature that discriminates between FLT3-TKD and FLT3-wild type CN-AML (Whitman et al. 2008). On the other hand, other studies reported difficulties in predicting FLT3 mutations from GEP results (Valk et al. 2004; Verhaak et al. 2009). Such conflicting results may reflect the frequent coincidence of FLT3 mutations with other mutations in CN-AML (Dohner et al. 2010), where the co-occurrence of several genetic abnormalities is likely to impact on the phenotype and the specific GEP. Interestingly, however, a specific gene expression signature derived from FLT3-ITD CN-AML, although not highly accurate in predicting the FLT3-ITD genotype, proved to be extremely accurate in predicting clinical outcome (Bullinger et al. 2008), suggesting that activation of the FLT3 pathway may be mediated by other yet unidentified genetic alterations.

In summary, GEP has proven to be extremely reliable in identifying AML cases with recurrent chromosomal abnormalities, but less predictive in identifying specific gene mutations in CN-AML (Verhaak et al. 2009), and this raises doubts as to its applicability in a clinical setting. However, a recent multicenter study involving 11 laboratories that use different microarray platforms (MILE – Miocarray Innovations in Leukemia) demonstrated that GEP is a robust technology for the diagnosis of hematologic malignancies with high accuracy (Haferlach et al. 2010), and that in some cases GEP outperformed routine diagnostic procedures. The analysis of larger cohorts of AML cases, in particular of the less
frequent molecular subtypes, will likely be necessary for the identification of reliable gene expression signatures of clinical utility. Whether GEP-derived predictors can be of use for prognosis, and in particular whether GEP presents a concrete advantage over standard cytogenetic or molecular markers in terms of prognostic value, remains an open question. Different studies have identified specific gene expression signatures that correlate with clinical outcome in AML (Bullinger et al. 2004; Metzeler et al. 2008; Radmacher et al. 2006), suggesting that GEP may, in fact, yield diagnostic and prognostic information simultaneously.

3. GEP and the biology of AML

As discussed above, GEP results derived from AML patients have been exploited to derive information concerning the biology underlying the disease. For the purpose of identifying specific functional pathways relevant to leukemogenesis, this approach is, however, partially hampered by properties that are intrinsic to patient-derived material, including individual genetic variability and the presence of heterogeneous cellular populations in each sample. It is, therefore, possible to exploit different experimental model systems, including purified cellular subpopulations, cell lines or animal models, with the aim of unraveling the molecular mechanisms underlying leukemic transformation. Such approaches have been widely proved to be reliable by extensive validation through a variety of independent methods. A significant number of reports have identified transcriptional targets deregulated in specific types of AML, and it is not possible to discuss all the data and their implications in due detail. We will instead briefly review specific aspects that emerged from these studies, focusing on their possible relevance to AML pathogenesis and management.

3.1 Common functions: Myeloid differentiation and stem cell maintenance

A general feature of AML-associated oncogenes is the capacity to block the process of myeloid differentiation and to promote self-renewal of hematopoietic precursor/stem cells. A variety of model systems have been employed for GEP analyses with the aim of identifying specific genes and pathways underlying these properties. Some of these studies have highlighted the existence of target genes that are common to several AML-associated oncogenes, suggesting that diverse genetic mutations can lead to the deregulation of overlapping downstream functional pathways. This observation is of potential clinical importance, since it suggests the existence of common therapeutic targets, regardless of the specific initiating oncogenic lesion. For example, expression of AML fusion proteins such as AML1/ETO, PML/RAR, and PLZF/RAR in the U937 hematopoietic cell line resulted in deregulation of a large set of common targets (Alcalay et al. 2003). These included a decreased expression of genes involved in myeloid differentiation, such as GFI1, CSF3R, STAT5A and others, and the activation of pathways leading to increased stem cell renewal (in particular, the Jagged1/Notch pathway). A similar approach led to the identification of Wnt signaling activation by diverse AML oncogenes, through upregulation of plakoglobin expression (Muller-Tidow et al. 2004). The existence of common leukemogenic functions is further suggested by the observation that a specific cellular subpopulation derived from CEBPA-deficient leukemia, which is capable of transferring AML to recipient mice, revealed a GEP signature shared with MLL-AF9-transformed AML (Kirstetter et al. 2008).
The identification of overlapping functions deregulated by AML oncogenes is perhaps to be expected. In fact, the genetic lesions underlying AML pathogenesis mostly involve transcriptional regulators that function during myeloid differentiation in an orchestrated manner, and are physiologically cross-regulating each other’s expression. For example, CEBPA is a crucial factor in myeloid differentiation, and its expression is often attenuated or repressed by oncogenic transcription factors such as AML1/ETO (Pabst et al. 2001). Therefore, part of the transcriptional response elicited by AML1/ETO may be due to a decrease in CEBPA activity. Overexpression of CEBPA in human CD34+ hematopoietic precursors induces the expression of genes involved in myeloid differentiation (Cammenga et al. 2003), which are presumably targets of deregulation not only in AML with mutated CEBPA, but also in all AML that present decreased levels of CEBPA expression.

Interestingly, mutations involving genes that do not directly regulate gene expression, such as FLT3 or NPM1, are associated to alterations in global gene expression that partially overlap with those described for oncogenic transcription factors. Expression of FLT3 mutants in murine 32Dcl3 cells resulted in repression of genes involved in myeloid differentiation, including CEBPA, and in the activation of a transcriptional program that partially overlaps with that induced by IL-3, a potent hematopoietic cytokine (Mizuki et al. 2003). Recently, the expression of an NPM1 mutant allele in mouse HSC was shown to result in HOX gene overexpression, reproducing the situation of primary AML with mutated NPM1 (Vassiliou et al. 2011). The molecular mechanisms through which these mutants elicit a transcriptional response remain to be elucidated.

3.2 Characterization of Leukemic Stem Cells

AML derives from the transformation of a single hematopoietic progenitor/stem cell, known as leukemic stem cell (LSC), which shares important properties with normal hematopoietic stem cells (HSC), including unlimited self-renewal and the capacity to give origin to a hierarchy of hematopoietic cells. The molecular characterization of LSC and the identification of functions that are specific of LSC with respect to normal HSC are clearly instrumental for designing novel therapeutic strategies aimed at eradicating AML. The transforming genetic event does not necessarily occur in HSC, but may take place in more differentiated progenitors that reacquire stem cell characteristics (Passegue et al. 2003). In favor of the latter possibility, Krivtsov et al. isolated leukemic stem cells (LSC) from a mouse model of AML generated by the MLL-AF9 fusion protein, which revealed a GEP that was reminiscent of normal granulocyte/macrophage progenitors (Krivtsov et al. 2006). However, a subset of genes that is highly expressed in normal HSC appeared to be re-activated in LSC, including several HOXA genes, STAT1 and CD44. Another study conducted on CD34+ AML revealed two distinct subpopulations of LSC, the more mature resembling normal granulocyte-macrophage progenitors in terms of GEP, and the immature LSC population reminiscent of lymphoid-primed multipotent progenitors (Goardon et al.). Taken together, these studies and the ones discussed in the previous section suggest that AML initiates in progenitor cells that re-acquire specific stem cell characteristics, such as activation of Notch and/or Wnt signaling and/or over-expression of HOX genes, which are tightly linked to the acquisition of an unlimited self-renewal capacity and cause an arrest in the differentiation program.

However, the identification of functions that are specific of LSC with respect to normal HSC is clearly of importance for the identification of novel therapeutic strategies aimed at
eradicating AML. In terms of global gene expression, LSCs are not simply characterized by the re-activation of stem cell pathways and maintenance of self-renewal. A direct comparison of the GEP of highly enriched normal human HSC and LSC from AML of diverse subtypes revealed differences in relevant functional pathways, including Wnt signaling, MAP Kinase signaling, and Adherens Junction (Majeti et al. 2009). The latter is particularly intriguing, since it suggests specific abnormalities in the relationship between LSCs and the microenvironment (“niche”).

3.3 Response to therapy
AML are characterized by a heterogeneous response to therapy, and although there has been notable progress in the past decades, most patients still succumb to the disease. The search for new therapeutic strategies is therefore of paramount importance, and GEP studies have also been exploited to dissect the molecular basis underlying the response to AML therapy.

One way to identify genes/pathways that may determine response to therapy is to compare GEP of treated versus untreated AML cells. Among AML, APL represents an exception in that its exquisite sensitivity to RA and arsenic trioxide treatment has dramatically changed its prognosis to a 5-year survival rate of 90%. Several studies have described specific transcriptional programs that are modulated by RA in APL cells, with the aim of identifying targets that may be of wider use in AML treatment. GEP analysis of APL blasts and PML/RAR-expressing U937 cells treated with RA in vitro revealed that the transcriptional response to RA is characterized by regulation of genes involved in the control of differentiation, stem cell self-renewal and chromatin remodeling, suggesting that specific structural changes in local chromatin domains may be required to promote RA-mediated differentiation (Meani et al. 2005).

Another possible approach is to compare the GEP of sensitive versus resistant AML cells after treatment with drugs. Tagliafico et al. derived a molecular signature that predicts the resistance or sensitivity to differentiation induced by RA or vitamin D in six myeloid cell lines, and proved its validity in a set of primary AML blasts using an in vitro differentiation assay (Tagliafico et al. 2006). Similarly, Zuber at al., described the differences between a chemosensitive and a chemoresistant AML model: murine AML expressing the AML1/ETO fusion protein, which show a dramatic response to chemotheraphy, displayed activation of the p53 tumor suppressor function. Murine AML expressing MLL fusion proteins are instead drug-resistant and present an attenuated p53 response. It appears, therefore, that the p53 network has a central role in the response to chemotherapy in AML (Zuber et al. 2009).

Importantly, GEP information can also be exploited for the identification of new therapeutic options. Corsello et al. defined an AML1/ETO GEP signature by comparing a t(8;21) bearing cell line before and after siRNA-mediated inhibition of the fusion protein, and used the resulting signature to screen a set of drug-induced expression profiles (Corsello et al. 2009). In a recent study, publicly available GEP data sets derived from APL patients were exploited for the identification of an “APL signature”, which was then compared to a collection of expression profiles for more than 1300 bioactive compounds for the discovery of relevant drug candidates (Marstrand et al. 2010). Although these studies have not yet been transferred to the clinics, they open a concrete possibility for exploiting GEP data alongside chemical genomics approaches for the in silico identification of molecularly targeted drugs.
4. Conclusion and perspectives

Have GEP studies truly had an impact on the management of AML? Currently, no GEP-based diagnostic/prognostic tests are available for AML in clinical practice. However, tests that reliably predict the outcome for cancer patients based on the expression pattern of a selected subset of genes identified through GEP are available for other malignancies, such as breast cancer (van’t Veer&Bernards 2008), and available evidence suggests that a reliable set of prognostic predictors could be established for AML as well. Furthermore, GEP can accurately sub-classify most AML according to the underlying genetic abnormality even when histopathological data are ambiguous, and outperforms routine diagnostic tests in certain cases, raising the possibility to introduce GEP-derived approaches in a diagnostic setting. A particularly exciting perspective is to exploit GEP data in combination with chemical genomics for the design of novel therapeutic strategies aimed at molecular targets. Many factors concur in determining the AML phenotype, and in recent years there has been a growing interest in high-throughput approaches other than GEP to analyze microRNA (miRNA) expression, epigenetic modifications and whole-genome DNA sequencing in AML. MiRNAs are small non coding RNAs that can regulate the expression levels of numerous target mRNAs both at the transcriptional and post-transcriptional level. Similarly to what has been observed for mRNAs, specific miRNA signatures have been associated with AML subtypes (Garzon et al. 2008; Li et al. 2008; Marcucci et al. 2008), and may therefore represent an additional option for the development of clinically useful tools.

Epigenetic modifications including DNA methylation and covalent histone modifications, such as acetylation, methylation and ubiquitination, are known to play a crucial role in the regulation of gene expression, and different epigenetic alterations have been described in leukemias (Plass et al. 2008). Recently, specific DNA methylation profiles have been described for distinct cytogenetic and molecular AML subtypes, and a 15-gene methylation classifier was found to be predictive of overall patient survival (Figueroa et al. 2010). Interestingly, the integration of DNA methylation data with GEP was shown to further improve prognostication in AML, suggesting that integration of genomic approaches may prove of clinical importance (Bullinger et al. 2010).

The advent of next-generation sequencing technology has opened the possibility to investigate the complexity of cancer genomes, and the first complete sequence of a human malignancy reported was that of an AML genome (Ley et al. 2008). The authors identified ten somatic mutations, two of which had already been described (NPM1 and FLT3), while the other eight were novel. None of the latter were, however, detected in a cohort of 187 AML cases, casting serious doubts as to their relevance in determining the leukemic phenotype. Successive studies using the same approach identified novel mutations in the IDH1 and DNMT3 genes that are instead recurrent in AML, underlining the power of this approach for the discovery of genetic alterations in cancer (Ley et al. 2010; Mardis et al. 2009). However, these studies also highlighted the relevant genetic heterogeneity among AML patients within the same subtype, since most of the mutations described were actually specific to the single patient under analysis, and their contribution to AML progression remains to be defined. One distinct possibility is that they represent “passenger” mutations that arise as a consequence of cancer-associated genomic instability, and bear no functional relevance to the malignant phenotype. On the other hand, such mutations may instead involve different players within complex functions (for example, regulators of proliferation or differentiation), and although the mutated genes are different, the functional

www.intechopen.com
consequence(s) may be the same. In any case, the co-existence of several genetic alterations within the same cell is bound to have an impact on gene expression, and it will therefore be necessary to integrate GEP data with the corresponding results from mutational analyses. Finally, microarray technology has inherent limitations, and with the advancement of current sequencing approaches may rapidly become obsolete. The possibility to sequence entire transcriptomes (RNA-seq) has several advantages over microarray-based GEP, including transcription start site mapping, gene fusion detection, small RNA identification and detection of alternative splicing events (Ozsolak & Milos 2011). The first RNA-seq analysis of an AML model reported an unexpected level of transcriptome variation between phenotypically similar LSC, including a large number of structural differences such as alternative splicing and promoter usage (Wilhelm et al. 2011). These results suggest a broad transcriptional heterogeneity in AML that is not limited to differences in mRNA levels. In the next few years there will inevitably be an explosion of genomic data in AML, describing yet unknown molecular mechanisms underlying the disease. The large amount of already available GEP data will have to be integrated with the new findings to increase its value in generating knowledge that can ultimately be translated into clinically useful tools.

5. References


Haferlach, T., A. Kohlmann, S. Schnittger, M. Dugas, W. Hiddemann, W. Kern, & C. Schoch. (2005a). "AML M3 and AML M3 variant each have a distinct gene expression signature but also share patterns different from other genetically defined AML subtypes." Genes Chromosomes Cancer no. 43 (2):113-27, ISSN: 1045-2257


Clinical and Biological Relevance of Gene Expression Profiling in Acute Myeloid Leukemia


This book comprises a series of chapters from experts in the field of diagnosis and treatment of myeloid leukemias from all over the world, including America, Europe, Africa and Asia. It contains both reviews on clinical aspects of acute (AML) and chronic myeloid leukemias (CML) and original publications covering specific clinical aspects of these important diseases. Covering the specifics of myeloid leukemia epidemiology, diagnosis, risk stratification and management by authors from different parts of the world, this book will be of interest to experienced hematologists as well as physicians in training and students from all around the globe.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
