Gene expression profiling in acute myeloid leukaemia

H.J.M. de Jonge1,2*, G. Huls3, E.S.J.M. de Bont4

1Division of Paediatric Oncology/Haematology, Department of Paediatrics, Beatrix Children’s Hospital, Groningen, the Netherlands, *corresponding authors: tel.: +31 (0)53 61 41 46, fax: +31 (0)53 61 42 35, e-mail: e.s.j.m.de.bont@bkk.umcg.nl or h.j.m.de.jonge@int.umcg.nl

ABSTRACT

Acute myeloid leukaemia (AML) is a heterogeneous disease characterised by clonal malignant haematopoiesis with a differentiation arrest and excessive proliferation of leukaemic blasts. Over the past decades, the heterogeneity of AML has been illustrated by evolving classifications based on morphology (French-American-British classification (FAB classification), cytogenetic abnormalities (e.g. t(8;21), monosomies etc.), phenotype and/or molecular abnormalities (e.g. Fms-like tyrosine kinase 3 gene internal tandem duplication (FLT3-ITD), mutations in nucleophosmin 1 (NPM1) and the transcription factor CCAAT/enhancer binding protein α (CEBPA), etc.). The current World Health Organisation (WHO) 2008 classification has integrated these classification modalities. Clinically, dissection of AML into various subtypes allows better survival prediction, but has still limited impact on treatment strategies, with the exception of all-trans retinoic acid treatment for AML-M3 and no allogeneic haematopoietic cell transplantation in complete remission (CR1) for patients with normal karyotype bearing an NPM1 mutation without FLT3-ITD. However, enhanced understanding of the molecular biology of AML will likely result in more ‘tailor-made’ therapies, for example by adding specific tyrosine kinase inhibitors to standard chemotherapy.

In this review, we summarise the variables currently used to classify AML. Specifically, the contribution of microarrays in classification, prognosis and understanding of pathobiology of AML is discussed.

KEYWORDS

Acute myeloid leukaemia, gene expression profiling, microarray, prognostic factors

INTRODUCTION

ACUTE MYELOID LEUKAEMIA

Acute myeloid leukaemia (AML) is defined as a clonal disorder caused by malignant transformation of a bone marrow-derived, self-renewing stem or progenitor cell, which demonstrates an enhanced proliferation as well as aberrant differentiation resulting in haematopoietic insufficiency (i.e. granulocytopenia, thrombocytopenia or anaemia). The clinical signs and symptoms of AML are diverse and nonspecific, but they are usually directly caused by the leukaemic infiltration of the bone marrow, with resultant cytopenia. AML is considered to be a heterogeneous group of disorders with variable underlying abnormalities and clinical behaviour, including responses to treatment. Therefore, classification of the disease is important and several classification systems exist to subdivide AML.

FAB classification

Historically, AMLs were divided into subtypes based on the type of cell from which the leukaemia developed and the level of maturation (i.e. French-American-British (FAB) classification). In addition, cytogenetic analysis of leukaemic blasts has resulted in the identification of non-random clonal chromosomal aberrations, of which some have been correlated to specific FAB subtypes (e.g. t(15;17) with AML-M3).

WHO classification

Nowadays, the World Health Organization (WHO) provides a classification system in which morphology, cytogenetics, molecular genetics, and immunological markers are incorporated and interrelated. Recently, for the first time, specific gene mutations (i.e. mutations in CEBPA and NPM1) have been included as ‘provisional entities’ in
the revised WHO 2008 classification for AML. There is growing evidence that these two gene mutations represent primary genetic lesions (so-called class II mutations) that impair haematopoietic differentiation. Mutations in the fms-related tyrosine kinase 3 (FLT3) gene (e.g. FLT3-ITD or FLT3 kinase domain mutations) are considered class I mutations conferring a proliferation and/or survival advantage. AML with FLT3 mutations is not considered a distinct entity, although determining the presence of such mutations is recommended because they have prognostic significance.

Prognostic factors
A number of clinical and biological features that reflect the heterogeneity of AML are used to predict the likelihood that a patient will have a response to treatment or relapse. Adverse prognostic factors in AML include increasing age, a poor performance before treatment, unfavourable cytogenetic abnormalities and a high white blood cell count. Furthermore, therapy-related AML or AML arising after a myelodysplastic or myeloproliferative syndrome is usually more resistant to standard treatment than de novo AML.

Cytogenetics
Important predictors of disease outcome are the pre-treatment cytogenetic and molecular findings in AML blasts. To date, in AML approximately 200 different structural and numerical aberrations have been described. Cytogenetic findings permit patient risk to be categorised as favourable, intermediate or unfavourable, with very different cure rates. Although there may be (subtle) differences in the criteria used to define these risk groups among different study groups, the presence of for instance t(8;21)(q22;q22), t(15;17)(q22;q21) and inv(16)(p13q22)/t(16;16)(p13;q22) is generally classified as favourable-risk AML (with leucocytes <20 x 10⁹). On the other end of the spectrum is the unfavourable-risk group, which includes blasts showing e.g. monosomies of chromosome 5 or 7, deletion of the long arm of chromosomes 3, 5 and 7 and complex karyotypes. Of note, the monosomal karyotype, defined as non-core-binding factor (CBF) leukaemias with a karyotype with at least two autosomal monosomies or one single autosomal monosomy in the presence of one or more structural cytogenetic abnormalities, is considered to be a better predictor of (very) poor outcome than the traditionally defined complex karyotype. The intermediate-risk group includes AMLs with a normal karyotype and AMLs which are not classified in the other two risk groups.

Molecular genetics
In recent years, the discovery of mutations in e.g. genes encoding FLT3, NPM1 and CEBPA has shown to be of major importance (table 1). Nowadays, it is increasingly possible to distinguish subsets of patients with differing outcomes from the large cohort with a normal karyotype AML or miscellaneous cytogenetic abnormalities considered as intermediate-risk cytogenetics. The majority of FLT3 receptor tyrosine kinase gene mutations are internal tandem duplications (ITD); less frequent are mutations involving the tyrosine kinase domain (TKD). Several groups have consistently reported that FLT3-ITD is a major independent adverse risk factor in AML. The prognostic relevance of FLT3-TKD mutations, however, remains controversial. FLT3-ITD has a prevalence of 20 to 25% in young adults and nearly 35% in the older adult population. The ratio of the FLT3-ITD and the wild-type FLT3 (measured by polymerase chain reaction, PCR) varies from patient to patient, and this difference may have clinical implications. Thiede et al. found that patients with an allelic ratio (AR) above the median (0.78) had significantly shorter overall and disease-free survival, whereas survival in patients with ratios below 0.78 did not differ from those without FLT3 aberrations. CEBPA, a transcription factor involved in normal myelopoiesis, is mutated in ~10% of AML cases and predicts a relatively favourable outcome in paediatric patients. The revised WHO 2008 classification for AML: There is growing evidence that these two gene mutations represent primary genetic lesions (so-called class II mutations) that impair haematopoietic differentiation. Mutations in the fms-related tyrosine kinase 3 (FLT3) gene (e.g. FLT3-ITD or FLT3 kinase domain mutations) are considered class I mutations conferring a proliferation and/or survival advantage. AML with FLT3 mutations is not considered a distinct entity, although determining the presence of such mutations is recommended because they have prognostic significance.

Prognostic factors
A number of clinical and biological features that reflect the heterogeneity of AML are used to predict the likelihood that a patient will have a response to treatment or relapse. Adverse prognostic factors in AML include increasing age, a poor performance before treatment, unfavourable cytogenetic abnormalities and a high white blood cell count. Furthermore, therapy-related AML or AML arising after a myelodysplastic or myeloproliferative syndrome is usually more resistant to standard treatment than de novo AML.

Cytogenetics
Important predictors of disease outcome are the pre-treatment cytogenetic and molecular findings in AML blasts. To date, in AML approximately 200 different structural and numerical aberrations have been described. Cytogenetic findings permit patient risk to be categorised as favourable, intermediate or unfavourable, with very different cure rates. Although there may be (subtle) differences in the criteria used to define these risk groups among different study groups, the presence of for instance t(8;21)(q22;q22), t(15;17)(q22;q21) and inv(16)(p13q22)/t(16;16)(p13;q22) is generally classified as favourable-risk AML (with leucocytes <20 x 10⁹). On the other end of the spectrum is the unfavourable-risk group, which includes blasts showing e.g. monosomies of chromosome 5 or 7, deletion of the long arm of chromosomes 3, 5 and 7 and complex karyotypes. Of note, the monosomal karyotype, defined as non-core-binding factor (CBF) leukaemias with a karyotype with at least two autosomal monosomies or one single autosomal monosomy in the presence of one or more structural cytogenetic abnormalities, is considered to be a better predictor of (very) poor outcome than the traditionally defined complex karyotype. The intermediate-risk group includes AMLs with a normal karyotype and AMLs which are not classified in the other two risk groups.

Molecular genetics
In recent years, the discovery of mutations in e.g. genes encoding FLT3, NPM1 and CEBPA has shown to be of major importance (table 1). Nowadays, it is increasingly possible to distinguish subsets of patients with differing outcomes from the large cohort with a normal karyotype AML or miscellaneous cytogenetic abnormalities considered as intermediate-risk cytogenetics. The majority of FLT3 receptor tyrosine kinase gene mutations are internal tandem duplications (ITD); less frequent are mutations involving the tyrosine kinase domain (TKD). Several groups have consistently reported that FLT3-ITD is a major independent adverse risk factor in AML. The prognostic relevance of FLT3-TKD mutations, however, remains controversial. FLT3-ITD has a prevalence of 20 to 25% in young adults and nearly 35% in the older adult population. The ratio of the FLT3-ITD and the wild-type FLT3 (measured by polymerase chain reaction, PCR) varies from patient to patient, and this difference may have clinical implications. Thiede et al. found that patients with an allelic ratio (AR) above the median (0.78) had significantly shorter overall and disease-free survival, whereas survival in patients with ratios below 0.78 did not differ from those without FLT3 aberrations. CEBPA, a transcription factor involved in normal myelopoiesis, is mutated in ~10% of AML cases and predicts a relatively favourable outcome in paediatric patients.

Table 1. Recurrent molecular abnormalities in adult AML

<table>
<thead>
<tr>
<th>Gene mutation</th>
<th>Percentage of cases</th>
<th>Prognostic significance</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fms-related tyrosine kinase 3 (FLT3), internal tandem duplication (ITD)</td>
<td>20-35</td>
<td>Unfavourable</td>
<td>27-31</td>
</tr>
<tr>
<td>CCAAT/enhancer binding protein alpha (CEBPA)</td>
<td>5-10</td>
<td>Favourable, when mutated on both alleles</td>
<td>32-37</td>
</tr>
<tr>
<td>Nucleophosmin (NPM1)</td>
<td>25-35</td>
<td>Favourable in absence of FLT3-ITD</td>
<td>34,35,38,39</td>
</tr>
<tr>
<td>Wilms tumour 1 (WT1)</td>
<td>10-15</td>
<td>Unfavourable?</td>
<td>40-42</td>
</tr>
<tr>
<td>RAS</td>
<td>10-15</td>
<td>-</td>
<td>34</td>
</tr>
<tr>
<td>Cytosolic isocitrate dehydrogenase 1/2 (IDH1, IDH2)</td>
<td>10-25</td>
<td>In subsets unfavourable?</td>
<td>47-50</td>
</tr>
<tr>
<td>Tet oncogene family member 2 (TET2)</td>
<td>12-20</td>
<td>Unfavourable?</td>
<td>51-53</td>
</tr>
<tr>
<td>KIT</td>
<td>2-8</td>
<td>Unfavourable?</td>
<td>54-58</td>
</tr>
<tr>
<td>DNA (cytosine-5)-methyltransferase 3 alpha (DNMT3A)</td>
<td>22</td>
<td>Unfavourable?</td>
<td>59</td>
</tr>
<tr>
<td>Protein tyrosine phosphatase, non-receptor type 11 (PTPN11)</td>
<td>&lt;5</td>
<td>-</td>
<td>60</td>
</tr>
<tr>
<td>Runt related transcription factor 1 (RUNX1)</td>
<td>&lt;5</td>
<td>-</td>
<td>60</td>
</tr>
</tbody>
</table>

*Due to space limitations, only a selected number are given for each abnormality.
and adult AML, however, only when CEBPA is mutated on both alleles.\textsuperscript{54-57} Approximately 50\% of adult normal karyotype AMLs harbour an NPM1 mutation, which leads to delocalisation of the NPM1 protein to the cytoplasm.\textsuperscript{58} NPM1 and FLT3-ITD commonly co-exist in normal karyotype AML suggesting that they may cooperate in generating the leukemic phenotype. The presence of an NPM1 mutation (in the absence of an FLT3-ITD mutation) is associated with better outcome in terms of higher complete response rates and increased long-term survival compared with patients lacking the mutation.\textsuperscript{54,55,59} Consequently, it has been suggested that cytogenetically normal AML involving the genotype of mutant NPM1 without FLT3-ITD should no longer be classified as intermediate-risk leukaemia but rather should be classified as favourable-risk leukaemia.\textsuperscript{39} Furthermore, patients with mutant NPM1 without FLT3-ITD may not benefit from related-donor transplantation as first-line treatment.\textsuperscript{19} Mutations in the Wilms' tumour gene (WT1), present in ~10\% of patients with normal karyotype AML, have been found to be associated with poor outcome, especially in combination with an FLT3-ITD.\textsuperscript{40-42} RAS mutations, occurring in ~15\% of cases, are suggested to be prognostically neutral.\textsuperscript{14} Recently, mutations in genes involved in metabolism have been discovered.\textsuperscript{44,45} In AML, but also in low-grade gliomas and secondary glioblastoma multiforme (GBM), mutations in cytosolic isocitrate dehydrogenase 1 (IDH1) and its mitochondrial homolog IDH2 have been identified. Both IDH1 and IDH2 are important enzymes in the citrate cycle (Krebs cycle). Two distinct alterations are caused by the tumour-derived mutations in IDH1 or IDH2: loss of its normal catalytic activity in the production of \(\alpha\)-ketoglutarate (\(\alpha\)-KG) and gain of the catalytic activity to produce 2-hydroxyglutarate (2-HG). Consequently, less \(\alpha\)-ketoglutarate is available for biological processes in which it functions as a co-factor. Remarkably, IDH1/2 mutations, occurring in ~10 to 25\% of AML cases,\textsuperscript{67-69} were mutually exclusive with mutations in gene encoding the \(\alpha\)-ketoglutarate-dependent enzyme tet oncogene family member 2 (TET2) (occurring in 12 to 20\% of AML cases).\textsuperscript{15} Loss-of-function mutations in TET2 were associated with similar epigenetic defects as IDH1/2 mutants. Interestingly, a shared proleukaemogenic effect between TET2 mutations and mutations in IDH1 and IDH2 was suggested since \(\alpha\)-ketoglutarate is a co-factor for TET2 in the hydroxylation of 5-methylcytosine and thus effects the methylation process.\textsuperscript{46}

In cytogenetically favourable core binding factor (CBF) AML (i.e. AML with t(8;21) or inv(16)(t(16;16)), the presence of a mutation in the KIT receptor tyrosine kinase has been shown to have an unfavourable influence on outcome in retrospective studies.\textsuperscript{54-58} Recently, highly recurrent mutations in the DNA methyltransferase gene DNMT3A have been discovered and were found to be independently associated with poor outcome in AML.\textsuperscript{39} Other mutations as those involving protein tyrosine phosphatase, non-receptor type 11 (PTPN11) and runt-related transcription factor 1 (RUNX1) are relatively rare (i.e. <5\% of cases), making their relevance to risk-stratified treatment approaches uncertain at the present time.\textsuperscript{50}

### Effect of over-expressed genes on outcome

Quantitative expression levels of several genes (e.g. Brain And Acute Leukaemia Cytoplasmic gene BAALC),\textsuperscript{61-63} Ets-related gene (ERG),\textsuperscript{64,65} Meningioma-1 gene (MNI),\textsuperscript{66,67} and Ecotropic Viral Integration-1 gene (EVI1)\textsuperscript{68-70} have been shown to carry prognostic information in patients with (normal karyotype) AML (table 2). Except for EVI1, the molecular basis of up-regulation of these genes remains, however, poorly understood. Recently, it was shown that expression levels of ERG, BAALC and MNI are strongly correlated, which suggests that their prognostic significance may be overlapping.\textsuperscript{64} Several studies have evaluated the prognostic significance of expression of multidrug resistance (MDR) genes with varying conclusions.\textsuperscript{71-74} Expression of factors that may relate to interaction of leukemic cells with bone marrow microenvironment (e.g. vascular endothelial growth factor A (VEGF), and chemokine (C-X-C motif) receptor 4 (CXCR4)) as well as VEGFC have also been found to impact on outcome.\textsuperscript{75-77} Finally, high expression of p16\(^{INK4A}\) was found as a prognostic parameter for overall survival in older patients with AML.\textsuperscript{80}

| Table 2. Effect of quantitative expression levels of genes on outcome |
|------------------------|----------------|-----------------|----------------|
| Gene overexpression    | Percentage of cases\(\times\) | Prognostic significance | Reference\(\times\) |
| Brain and acute leukaemia cytoplasmic gene (BAALC) | -30 | Unfavourable | 61-63 |
| Ets-related gene (ERG) | -25 | Unfavourable | 64,65 |
| Meningioma-1 gene (MNI) | -25-50 | Unfavourable | 66,67 |
| Ecotropic viral integration-1 gene (EVI1) | 6-11 | Unfavourable | 68-70 |
| Chemokine (C-X-C motif) receptor 4 (CXCR4) | -33 | Unfavourable | 7778 |
| Vascular endothelial growth factor C (VEGFC) | -50 | Unfavourable | 79 |
| Cyclin-dependent kinase inhibitor 2A (CDKN2A, p16\(^{INK4A}\)) | -75 | Unfavourable | 80 |

Due to space limitations, only a selected number are given for each abnormality. * in case of overexpression, the percentage is based on the cut-off used in the referenced papers. This may involve simple dichotomisation (e.g. BAALC), resulting in 50\% of the cases by definition exhibiting overexpression. Of note, also continuous expression levels of VEGFC correlated with poor outcome.
GENE EXPRESSION PROFILING

Although an increasing number of prognostically relevant (cyto) genetic variables have been identified in AML, not all cases are currently classified adequately. To date, tremendous evidence exists that DNA microarray-based gene expression profiling adds an important new facet to the study of AML, e.g. in relation to classification opportunities. In the past decade, microarrays, together with the availability of the complete nucleotide sequence of the human genome, have made it possible to measure expression levels of thousands of different mRNA transcripts simultaneously.\textsuperscript{5,54} There are several (potential) applications for gene expression profiling (GEP) studies. GEP studies are well suited to reveal characteristic patterns (signatures) of activation or silencing or both of multiple genes that may reflect underlying biology of disease subtypes. Subsequently, this may provide diagnostic/prognostic information, and potentially reveal novel molecular targets for therapeutic intervention.

**Prediction of known classes: ‘class prediction’**

In an early landmark study in 1999, researchers demonstrated for the first time the power of GEP in leukaemias.\textsuperscript{56} In that particular study, GEP profiles were used to distinguish AML samples from those with acute lymphoblastic leukaemia in an unsupervised approach. Of note, the grouping of cases according to similar gene expression profiles is known as clustering.\textsuperscript{56,57} Clustering in an unsupervised approach is done in an unbiased way, i.e. without the use of external information such as patient baseline characteristics, mutations or cytogenetics. Class prediction refers to the possibility to predict leukaemia subtypes, as defined by their phenotypes and genotypes, with the use of GEP signatures. For instance, it was demonstrated that the prognostically favourable AML subtypes (i.e. t(8;21), t(15;17) and inv(16)) have distinctive GEP profiles which have consistently been found to be predictable with almost 100\% accuracy using GEP.\textsuperscript{58,59} Interestingly, paediatric AML GEP profiles could also be used to predict adult AML samples with identical cytogenetic abnormalities.\textsuperscript{60}

In addition, GEP profiles have a high accuracy to predict subgroups with rare translocations, as shown for the t(8;16) (p11;p13) with CBP and MOZ (monocytic leukemia zinc finger protein) re-arrangements.\textsuperscript{60,62} Moreover, unsupervised clustering revealed that mutations in CEBPA and also NPM1 correlated with gene expression signatures.\textsuperscript{62,63} However, the accuracy of prediction for other cytogenetic AML subtypes, such as those with abnormalities involving band 11q23, abnormalities involving 3q, -5/5q-, -7/7q- or t(9;22) was lower.\textsuperscript{58,59,60}

Similarly, the prediction accuracy for specific molecular subsets of patients such as those harbouring FLT3-ITD, FLT3-TKD and mutations in KRAS and NRAS genes was lower.\textsuperscript{59,60}

**Prediction of new AML subgroups: ‘class discovery’**

GEP studies also have the potential to uncover new subgroups in AML.\textsuperscript{58,59,60} This procedure is representative of class discovery. For example, Valk and colleagues identified 16 subgroups in 285 AMLs, several of which lacked previously known denominators.\textsuperscript{59} In addition, at least five other GEP studies revealed previously unrecognised heterogeneity within established paediatric as well as adult AML subtypes.\textsuperscript{58,59,60,62,63} Recently, it was demonstrated that a subset of AML patients who did not harbour CEBPA mutations could be characterised by a GEP signature resembling that of AML patients with CEBPA mutations.\textsuperscript{64} Interestingly, further experiments revealed that in these cases, CEBPA was epigenetically silenced, which indicates that the detection of a distinct gene expression subtype had indeed led to the discovery of a biologically meaningful subgroup.

From a clinical point of view, one of the most important challenges in AML is to enlarge insight into the pathobiology of AML in the elderly. In recent decades, survival of paediatric and adult AML patients has improved significantly, while survival of older AML patients (≥60 years) has remained virtually unchanged over the past decades resulting from the combination of poor chemotherapeutic tolerance and inherent chemotherapy resistance compared with younger AML patients.\textsuperscript{1,4,15} Moreover, AML in older patients shows a lower frequency of favourable core-binding chromosomal abnormalities and a higher incidence of complex aberrant karyotypes.

Recently, two studies showed that older patients with AML show distinct GEP signatures compared with younger patients with AML.\textsuperscript{60,63} The latter study described that, unlike healthy cells, AML-derived blasts show a down-regulation of p16\textsuperscript{INK4A} mRNA with increasing age. Based on this observation it was hypothesised that suppression of defence mechanisms which protect older cells against cellular and DNA damage might facilitate oncogenesis in older individuals.\textsuperscript{58,60}

So, GEP could help researchers to discover hidden heterogeneity within AML subtypes.

**GEP and predicting outcome in AML**

GEP has also been applied to derive prognostic signatures for AML that would identify subsets of patients with differing outcomes. In these studies treatment outcome or resistance were used to define a prognostic predictor.\textsuperscript{107,108} Hierarchical clustering analysis in 93 patients with core-binding factor AML revealed the stratification of two clusters with significantly different survival.\textsuperscript{109} In cytogenetically normal AML, Bullinger et al. were able to divide cytogenetically normal samples into two diverse prognostically relevant clusters using GEP.\textsuperscript{10} Importantly, the prognostic impact of this signature was independently validated in another cohort of AML samples using a
different platform and a longer follow-up.\textsuperscript{99} Of note, the prognostic effect of the signature was in part related to the occurrence of FLT3-ITD mutations, only 84 of 133 probes could be validated due to differences in platforms and the prediction accuracy of the classifier was overall modest, with approximately 60% of the patients having their outcome predicted correctly.\textsuperscript{99,100} Recently, another study in cytogenetically normal karyotype AMLs revealed a gene signature of 86-probe sets correlating significantly with overall survival.\textsuperscript{101} The prognostic effect of this classifier was independent of age, FLT3-ITD and NPM1 mutation status. In paediatric AML, a GEP study in 54 AML patients revealed 36 probe sets to be associated with prognosis.\textsuperscript{112} However, in an independent paediatric AML GEP study this prognostic signature could not be confirmed.\textsuperscript{98}

**Remarks and limitations**

Gene expression analysis can be performed on microarray platforms with varying kinds of probes (cDNA, short-oligonucleotide, long-oligonucleotide, etc.), production and labelling method (microbeads, spotting, in situ polymerisation, etc.). Specificity is highest for DNA-oligonucleotide microarrays of 40-60-mer probe length as they have a lower risk of cross-hybridisation.\textsuperscript{103} The widely-used Affymetrix microarrays rely on 25-mer in situ synthesised probes.\textsuperscript{104} The interpretation of the fluorescence intensity signals requires sophisticated computational methods for data normalisation and classification.\textsuperscript{105} because each study generates large datasets. GEP is a multistep procedure that can only be briefly outlined here. Initially, data pre-processing and quality control steps are performed for detection of array artefacts and the evaluation of the homogeneity of experimental groups. Furthermore, it is important to be aware of interstudy variations with regard to data normalisation, gene filtering and clustering procedures, which could influence the outcome of the analysis.\textsuperscript{106,107} Notably, significant efforts have led to the establishment of proposed guidelines to describe the minimum information about a microarray experiment (MIAME) that is needed to enable the interpretation of the results of the experiment unambiguously and potentially to reproduce the experiment. This is particularly important information if microarray data are deposited in a public database, such as the Gene Expression Omnibus.\textsuperscript{107,108}

GEP holds promise for developing molecular portraits of cancer subtypes with different clinical outcomes that could not be sub-classified or identified upon (initial) clinical presentation. One of the possible challenges in GEP studies is the (low) number of samples as compared with the number of genes tested, the so-called ‘curse of dimensionality’ (i.e., overfitting).\textsuperscript{109} In addition, there may be small numbers of genes whose expression discriminate cancer subtypes but they may not be driving causes of cancer initiation/progression and therefore provide little survival information. Another not surprising issue is that independent studies can identify different panels of genes with similar discriminatory specificity and power. Furthermore, the number of genes expected to be differentially expressed between two (or more) classes of interest within a single cancer subtype is probably small, and the differences in expression may not be large (enough) in relation to experimental noise.\textsuperscript{110} We have introduced the concept of TSR profiling that might improve the performance of predictive profiles.\textsuperscript{111} These transcriptional system regulators (TSRs) allowed one to characterise the expression profile of an individual microarray with just 50 TSR scores instead of using ten thousands of individual genes: i.e., a >500-fold reduction of complexity, thus avoiding the problem of overfitting. There is a second advantage of TSR profiling: i.e., when signals of multiple genes are added to calculate TSR scores the signal-to-noise ratio improves because noise cancels out. Further studies are needed to investigate whether TSR scores may be more reproducible input variables for prediction models than expression signals of selected individual genes.

**Biological versus statistics**

A pending question in GEP studies is whether large-fold changes in individual genes have more biological relevance than smaller but coordinated fold-changes in a set of genes (particularly along a single biological pathway). The assumption that (only) changes of more than twofold are significant is still surprisingly widespread.\textsuperscript{112} This threshold is based on initial publications by the Stanford group who found, from concordance analyses, that a more than twofold variation was significant for a particular set of experiments.\textsuperscript{113} This factor of two was subsequently referred to by others as a universal significance threshold, without realising its development. Moreover, in principle, the particular changes in gene expression between classes of samples may be less informative than the pathways they impact. Finally, it is important to realise that relative levels of mRNA expression do not necessarily reflect biological activity, as the latter may be highly dependent on other factors, such as posttranslational modifications.

**Clinical application**

Following the introduction of GEP in leukaemia research a decade ago by Golub and colleagues, various study groups worldwide have consistently shown that GEP can be used to predict molecularly defined subtypes of AML.\textsuperscript{114-118} However, from a clinical point of view, several questions surround GEP in AML: e.g., can GEP improve current diagnostics and risk classification schemes in AML, or the ability to predict outcome in AML patients beyond that currently provided by well-established
prognostic variables such as age, presenting white blood cell count and the presence of cytogenetic or molecular (e.g. mutations) abnormalities? To be able to answer such questions properly at least two important prerequisites should be met. Firstly, appropriate validation of GEP results in independent (prospective) study cohorts is needed. Secondly, for successful subgroup discovery it is crucial to have access to sufficiently large series of cases representing the various subtypes of AML. It may be unlikely that gene expression arrays will be used to diagnose cytogenetic and molecular abnormalities in the clinical setting when direct diagnostic assays are available and are more cost-effective. However, it is important to realise that the particular value of GEP-based classification lies in its comprehensiveness (i.e. the ability to measure tens of thousands of transcripts at one time) and its possibility to uncover (hidden) heterogeneity (e.g. related to differing outcome) within established cytogenetic and/or molecular subtypes of AML. However, the latter is highly dependent on the availability of high-quality samples and robustly annotated clinical data, which often have to be collected over many years. Ultimately, once intensively (prospectively) validated and standardised, measuring a panel of selected genes in combination with clinical (e.g. age, WBC count) and established variables (e.g. cytogenetics, and mutations) might be of importance in guiding doctors (therapeutic) decisions. Finally, from a cell biological point of view, particular efforts should be directed towards proper understanding of the biological mechanism and regulation of ‘genes with prognostic significance’. This aspect will clearly need to be further studied, also in terms of targeted therapy development and testing.

Which cells to profile?
There is not only heterogeneity among AML patients, heterogeneity is also evident within the AML cells of one patient. AML is thought to be initiated and maintained by a few leukaemia-initiating cells (LICs) that have an enhanced self-renewal capacity, can engraft in nonobese diabetic/severe combined immunodeficient mice and are, nowadays, believed to be restricted to the CD14⁺/CD38⁺ or CD34⁺/CD38⁺ fraction. However, there is evidence from mouse studies that mixed lineage leukaemia-associated human leukaemias can also arise from more progenitor cells. Furthermore, a recent study suggested that for some NPM1 mutated AMLs the LICs are also present in the CD14⁺ fraction. Most AML GEP studies, however, have been performed with the total AML mononuclear cell (MNC) fraction. Because cell lineage and differentiation stages might (theoretically) affect gene-expression based clustering, the differential expression of genes associated with the differentiation stage might obscure more basic gene expression information related to tumour initiation and maintenance. Consequently, profiling of more purified cell populations, instead of total MNC fractions, might enhance the possibilities of GEP in identifying novel prognostic markers or subgroup discovery. However, this approach directly depends on the accepted definition of immunophenotypic markers of leukaemia-initiating cells. Finally, there is compelling emerging evidence that cell nonautonomous contributions to leukaemia play a pivotal role in disease maintenance and propagation (i.e. the microenvironment, the niche).

Conclusions and future perspectives
Gene expression profiling using microarrays is currently the standard for analysing the transcriptome. However, profiling of e.g. microRNA (miRNA) levels, chromosomal copy number changes and epigenetic modifications have also played a pivotal role in enhanced molecular understanding of the (patho)biology of cancer, including AML. For example, similarly to mRNA profiling, miRNA profiling has revealed that specific subgroups of AML share distinctive miRNA signatures with prognostic significance. Furthermore, methylation profiling of a large series of AML patients identified several clusters, of which some could not be explained by the enrichment of any currently known recurrent cytogenetic, molecular, or clinical features. In recent times, next-generation sequencing (NGS) technologies have become available that enable gene expression analysis by direct shotgun sequencing of complementary DNA synthesised from RNA samples. NGS technologies have an impressive range of applications, and are increasingly being developed. In contrast to microarrays, sequencing technologies do not depend on predefined sequences, thus allowing for detection of, for example, new splicing variants or single-nucleotide polymorphisms. Furthermore, it allows genome-wide profiling of epigenetic marks. It is hypothesised that in the near future, NGS technologies could be used to obtain high-quality sequence data from a genome isolated from a single cell, which would be a substantial breakthrough, particularly for cancer genomics. Once we know the genomic landscape of cancer more adequately, what should follow? While genome-wide characterisation of cancer subtypes will likely reveal significant clues about genes that play a role in cancer progression, it is important to follow-up on these clues by carrying out functional screens of altered genes. Functional screening would aim to identify those (somatic) alterations that are imperative in tumour initiation and progression. Furthermore, functionally relevant mutations must be distinguished from passenger

Mutations (i.e. unimportant genetic changes caused by genomic instability of cancer cells). Finally, functional screening may establish candidate genes and their protein products for targeted therapy development or testing, as well as for diagnostic/prognostic assay development.

Acknowledgement

This work was partially supported by the Dutch Cancer Society (grant 5601; ES)/MB, grant 4566; GH) and a Netherlands Organisation for Scientific Research (NWO)–VENI grant (GH). None of the authors have a conflict of interest to disclose.

References


