Routine use of microarray-based gene expression profiling to identify patients with low cytogenetic risk acute myeloid leukemia: accurate results can be obtained even with suboptimal samples.

Diane Raingeard de la Blétière, Odile Blanchet, Pascale Cornillet-Lefèbvre, Anne Coutolleau, Laurence Baranger, Franck Geneviève, Isabelle Luquet, Mathilde Hunault-Berger, Annaelle Beucher, Aline Schmidt-Tanguy, et al.

To cite this version:
Diane Raingeard de la Blétière, Odile Blanchet, Pascale Cornillet-Lefèbvre, Anne Coutolleau, Laurence Baranger, et al.. Routine use of microarray-based gene expression profiling to identify patients with low cytogenetic risk acute myeloid leukemia: accurate results can be obtained even with suboptimal samples.. BMC Medical Genomics, BioMed Central, 2012, 5 (1), pp.6. 10.1186/1755-8794-5-6 . inserm-00673125

HAL Id: inserm-00673125
https://www.hal.inserm.fr/inserm-00673125
Submitted on 22 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Routine use of microarray-based gene expression profiling to identify patients with low cytogenetic risk acute myeloid leukemia: accurate results can be obtained even with suboptimal samples

Diane Raingeard de la Blétière1,2†, Odile Blanchet2,3†, Pascale Cornillet-Lefèbvre4, Anne Coutolleau1, Laurence Baranger5, Franck Geneviève3, Isabelle Luquet4, Mathilde Hunault-Berger5, Annaelle Beucher3, Aline Schmidt-Tanguy2,6, Marc Zandecki3, Yves Delneste7, Norbert Ifrah2,6 and Philippe Guardiola1,2,6*

Abstract

**Background:** Gene expression profiling has shown its ability to identify with high accuracy low cytogenetic risk acute myeloid leukemia such as acute promyelocytic leukemia and leukemias with t(8;21) or inv(16). The aim of this gene expression profiling study was to evaluate to what extent suboptimal samples with low leukemic blast load (range, 2-59%) and/or poor quality control criteria could also be correctly identified.

**Methods:** Specific signatures were first defined so that all 71 acute promyelocytic leukemia, leukemia with t(8;21) or inv(16)-AML as well as cytogenetically normal acute myeloid leukemia samples with at least 60% blasts and good quality control criteria were correctly classified (training set). The classifiers were then evaluated for their ability to assign to the expected class 111 samples considered as suboptimal because of a low leukemic blast load (n = 101) and/or poor quality control criteria (n = 10) (test set).

**Results:** With 10-marker classifiers, all training set samples as well as 97 of the 101 test samples with a low blast load, and all 10 samples with poor quality control criteria were correctly classified. Regarding test set samples, the overall error rate of the class prediction was below 4 percent, even though the leukemic blast load was as low as 2%. Sensitivity, specificity, negative and positive predictive values of the class assignments ranged from 91% to 100%. Of note, for acute promyelocytic leukemia and leukemias with t(8;21) or inv(16), the confidence level of the class assignment was influenced by the leukemic blast load.

**Conclusion:** Gene expression profiling and a supervised method requiring 10-marker classifiers enable the identification of favorable cytogenetic risk acute myeloid leukemia even when samples contain low leukemic blast loads or display poor quality control criterion.

Background

Prognostic evaluation is a critical step in newly diagnosed patients with acute myeloid leukemia (AML) in order to identify those at high risk of relapse. For AML patients, cytogenetic abnormalities as well as gene mutations and/or hyper-expressions detected at diagnosis are the main prognostic factors guiding the initial treatment strategy in a risk-oriented manner [1-4]. Hypergranular acute promyelocytic leukemia (APL), as well as AMLs with either translocation t(8;21)(q22;q22) [t(8;21)-AMLS] or inversion inv(16)(p13q22)/t(16;16)(p13;q22) [inv(16)-AMLS], are well-defined entities associated with a favorable outcome [1,3]. They can be distinguished from all other AML subtypes based on specific chromosomal alterations and fusion genes: PML/RAR-alpha fusion gene with reciprocal translocation t(15;17)(q24;q11) for APLs, AML1/ETO (also called RUNX1/RUNX1T1)...
fusion gene for t(8;21)-AMLs, and CBFB/MYH11 fusion gene for AMLs with either inv(16)(q21;q22) or balanced reciprocal translocation t(16;16)(q21;q22). Of note, in up to 15 percent of APLs, no translocation t(15;17)(q24; q21) is detected by conventional cytogenetics at diagnosis, despite PML/RARA fusion gene is detected using molecular assays [5,6]. Similarly, cryptic t(8;21)(q22;q22) and inv(16)(q21;q22), undetected by conventional cytogenetics, have also been reported [7-9].

With samples containing a high leukemic blast load, microarray-based gene expression profiling (GEP) and class prediction analyses have demonstrated their ability to assign AML samples to one of these three well characterized favorable cytogenetic risk AML subtypes, with high accuracy and low error rates [10-18]. One of the largest class prediction analyses in AML achieved 100 percent classification accuracy with respect to APL, t(8;21)-AML, and inv(16)-AML subtypes, indeed [10]. However, in the majority of those studies, the minimum percentage of leukemic cells within each sample is most often above 60 percent [10-18], an arbitrary threshold that is significantly different from the one used by cytologists for the diagnosis of AML, i.e., excess of blasts greater or equal to 20 percent.

To our knowledge, the impact of the leukemic blast load and/or of the sample quality on microarray-derived class prediction results has not been specifically studied in AMLs. For centers wishing to integrate microarray-based GEP in a routine prognostic workflow for newly diagnosed AML patients, one critical issue is the ability to perform accurate class prediction analysis with suboptimal samples, i.e., containing as low as 20 percent to 15 percent of APLs, no translocation t(15;17), t(8;21)-AML or inv(16)-AML, respectively detected. Analysis of NPM1, FLT3 and CEBPA for mutations was also performed for all patients [20-23]. All participants gave their written informed consent, and the study was approved by the Ethical Committee of Angers University Hospital.

The gene expression dataset, generated from 206 samples (182 AML samples, duplicated samples of NB4, Kasumi-1, and ME-1 cell lines, 18 samples of unique or combined normal bone marrows), was divided into a Training Set (n = 89 samples) and a Test Set (n = 117 samples). The main characteristics of AML patients and samples are summarized in Tables 1 and 2. Conventional cytogenetic banding, fluorescence in situ hybridization, and RT-PCR analysis of fusion gene transcripts were used to identify patients with APL, t(8;21)-AML or inv(16)-AML, as previously reported [19]. Analysis of NPM1, FLT3 and CEBPA for mutations was also performed for all patients [20-23]. All participants gave their written informed consent, and the study was approved by the Ethical Committee of Angers University Hospital.

### Methods

#### Characteristics of the patients and samples

A total of 182 bone marrow or peripheral blood samples from 97 AML patients, diagnosed with either APL (n = 18), t(8;21)-AML (n = 19), inv(16)-AML (n = 29), or NK-AML (n = 31), followed at Angers University Hospital (n = 72) or Reims University Hospital (n = 25), were analyzed, as well as duplicated samples of NB4, Kasumi-1, and ME-1 cell lines (n = 6 samples), which are derived from patients with APL, t(8;21)-AML and inv(16)-AML, respectively. In addition, 18 samples of unique (n = 9) or pooled (n = 9) normal bone marrows obtained from 12 healthy volunteers were included in the study. The main characteristics of AML patients and samples are summarized in Tables 1 and 2. Conventional cytogenetic banding, fluorescence in situ hybridization, and RT-PCR analysis of fusion gene transcripts were used to identify patients with APL, t(8;21)-AML or inv(16)-AML, as previously reported [19]. Analysis of NPM1, FLT3 and CEBPA for mutations was also performed for all patients [20-23]. All participants gave their written informed consent, and the study was approved by the Ethical Committee of Angers University Hospital.

### Table 1 Characteristics of the patients.

| Covariates          | Overall | APLs | t(8;21)-AMLs | inv(16)-AMLs | NK-AMLs |
|---------------------|---------|------|--------------|--------------|---------|
| Patients            | N = 97  | N = 18 | N = 19       | N = 29       | N = 31  |
| Gender              |         |       |              |              |         |
| Males               | 44      | 7     | 13           | 15           | 9       |
| Females             | 53      | 11    | 6            | 14           | 22      |
| Age at diagnosis    |         |       |              |              |         |
| Median (years)      | 54      | 56    | 53           | 38           | 60      |
| Range               | 18-87   | 19-87 | 18-84        | 18-70        | 25-78   |
| Leukocytosis at diagnosis |     |       |              |              |         |
| WBC ≥ 30 G/L        | 39      | 6     | 12           | 19           |         |
| FAB classification  |         |       |              |              |         |
| M1 or M2            | 45      | 0     | 18           | 6            | 21      |
| M3                  | 18      | 18    | 0            | 0            | 0       |
| M4 or M5            | 34      | 0     | 0            | 24           | 10      |
| Cyrognetics         |         |       |              |              |         |
| Expected anomaly*   | 62      | 16    | 19           | 27           | 0       |
| Normal karyotype    | 34      | 2     | 0            | 1            | 31      |
| Karyotype failure   | 1       | 0     | 0            | 1            | 0       |
| Gene mutations      |         |       |              |              |         |
| FLT3-ITD            | 19      | 2     | 1            | 2            | 14      |
| FLT3-D835<sup>a</sup> | 7      | 2     | 2            | 2            | 1       |
| NPM1                | 18      | 0     | 0            | 1            | 17      |
| CEBPA (monon-allelic) | 4     | 0     | 2            | 1            | 1       |

WBC, white blood cells. * t(15;17), t(8;21), inv(16) or t(16;16) for APL, t(8;21)-AML and inv(16)-AML, respectively detected. <sup>a</sup> Internal tandem duplication.
Table 2 Characteristics of the samples.

| Covariates | Overall | APLs | t(8;21)-AMLs | inv(16)-AMLs | NK-AMLs |
|------------|---------|------|-------------|-------------|---------|
| Groups     | N = 206 | N = 40 | N = 39      | N = 52      | N = 57  |
| Training Set AML samples | 71 | 14 | 14 | 15 | 28 |
| Training Set NBM<sup>†</sup> samples | 18 | – | – | – | – |
| Test Set AML samples* | 111 | 24 | 23 | 35 | 29 |
| Test Set AML cell lines | 6 | 2 | 2 | 2 | – |

- **Samples**<sup>*</sup> with optimal QCC<sup>⊥</sup>
  - **Blast % < 60% (undiluted)**
    - Overall, blasts < 60%
      - Overall, blasts ≥ 40% < 60%
        - Overall, blasts ≥ 20% < 40%
          - Overall, blasts ≥ 10% < 20%
            - Overall, blasts ≥ 5% < 10%
              - Overall, blasts < 5%
    - Diluted at 50%
    - Diluted at 75%
    - Diluted at 90%
    - Diluted at 95%
    - Overall, blasts < 60%

- **Samples**<sup>†</sup> with poor QCC<sup>⊥</sup>
  - **cRNA < 750 ng**
  - **Low RIN**
  - **Low RIN + low cRNA**

<sup>*</sup> including the three AML cell lines. <sup>†</sup> quality control criteria. <sup>⊥</sup> labeled cRNA to be hybridized on Illumina Beadchips.<br><br>samples). The Training Set, used to identify specific class markers (classifiers), included 18 samples of unique or pooled normal bone marrows (see details about the pools below) as well as 71 AML samples obtained from patients (bone marrow samples, n = 48; peripheral blood samples, n = 23) newly diagnosed with APL (n = 14 patients), t(8;21)-AML (n = 14 patients), inv(16)-AML (n = 15 patients) or NK-AML (n = 28 patients), which contained at least 60 percent blasts. The class prediction analysis was performed on a Test Set including 111 suboptimal AML samples (bone marrow samples, n = 78; peripheral blood samples, n = 33 obtained from 53 patients), i.e., with a low leukemic blast load (n = 101 samples containing from 2 to 56 percent blasts) and/or poor quality control criteria (n = 10 samples) as well as duplicated samples of NB4, Kasumi-1, and ME-1 cell lines (n = 6 samples). Among test samples with optimal quality control criteria, 22 originally contained less than 60% blasts (range, 5-56 percent blasts), whereas 79 were high blast load ones artificially diluted within one of the five pools of normal bone marrows (see details below and Tables 2 and 3). The 10 samples with suboptimal quality control criteria were characterized by either a low RNA integrity number suggestive of RNA degradation (n = 2), a low cRNA amount obtained after total RNA amplification and labeling (cRNA hybridized on BeadChips < 750 micrograms) (n = 7) or both (n = 1) (Tables 2 and 3).

**Sample preparation for gene expression profiling**

Blasts and mononuclear cells were purified by Ficoll-Hypaque density gradient centrifugation from bone marrow or peripheral blood samples (Nygaard, Oslo, Norway). Isolated cell samples were then immediately cryopreserved. Total RNAs were extracted from 10<sup>7</sup> thawed cells using RNeasy<sup>®</sup> Mini Kits (Qiagen Incorporation, Valencia, USA). Total RNA quantification was performed using the Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific Incorporation, Waltham, USA) according to manufacturer recommendations. Integrity of the extracted RNAs was assessed with the Bioanalyzer 2100 and the RNA6000 Nano kit (Agilent Technologies Incorporation, Santa Clara, USA). A RNA integrity number (RIN) greater or equal to 7.00 was achieved for 203 samples (RNA degradation observed for 3 AML samples). No sign of DNA contamination was detected in any of the 206 samples analyzed. The starting amount of total RNA used for the reactions was 200 nanograms per sample, for all samples. The Illumina Total Prep RNA Amplification Kit (Applied Biosystems/Ambion, Austin, USA) was used to generate biotinylated, amplified cRNA according to the manufacturer recommendations (see additional file 1 for details). To study the leukemic blast load effect on the class prediction accuracy, besides including 22 samples with good quality control criteria that originally contained less than 60 percent of leukemic blasts, 79 artificially generated low leukemic blast load samples, via a “dilution/mixture” approach, were also assessed (see additional file 1 for details for the generation of low leukemic blast load AML samples by a “dilution/mixture” approach). Hybridization on Illumina HumanHT-12 v3 Expression BeadChips, staining and detection of cRNAs on microarrays using an I-Scan system were performed according to Illumina’s protocol (see additional file 1 for details).

**Data analyses**

GenomeStudio 2010.3 software (Illumina Inc., San Diego, USA) and its Gene Expression Analysis Module
Table 3 Characteristics and class assignment of the 117 Test Set samples.

| UPN | Cell Source | Time | Blasts % | Dilution | QCC  | Real Class          | Assigned Class | Confidence |
|-----|-------------|------|----------|----------|------|---------------------|----------------|------------|
| NB4 | BM Diagnosis| 100  | No Dilution | OK       | APLs Test Set | APLs            | 10.802       |
| NB4 | BM Diagnosis| 100  | No Dilution | OK       | APLs Test Set | APLs            | 17.507       |
| UPN53| BM Diagnosis| 43   | Dilution 50% | OK       | APLs Test Set | APLs            | 26.028       |
| UPN22| BM Diagnosis| 40   | Dilution 50% | OK       | APLs Test Set | APLs            | 29.91        |
| UPN65| BM Diagnosis| 39   | Dilution 50% | OK       | APLs Test Set | APLs            | 26.403       |
| UPN40| BM Diagnosis| 33   | Dilution 50% | OK       | APLs Test Set | APLs            | 29.91        |
| UPN50| PB Diagnosis| 27   | Dilution 50% | OK       | APLs Test Set | APLs            | 19.2         |
| UPN53| BM Diagnosis| 22   | Dilution 75% | OK       | APLs Test Set | APLs            | 20.904       |
| UPN22| BM Diagnosis| 20   | Dilution 75% | OK       | APLs Test Set | APLs            | 29.91        |
| UPN65| BM Diagnosis| 20   | Dilution 75% | OK       | APLs Test Set | APLs            | 15.763       |
| UPN40| BM Diagnosis| 16   | Dilution 75% | OK       | APLs Test Set | APLs            | 26.425       |
| UPN5  | PB Diagnosis| 15   | No Dilution | OK       | APLs Test Set | APLs            | 17.193       |
| UPN50| PB Diagnosis| 14   | Dilution 75% | OK       | APLs Test Set | APLs            | 10.46        |
| UPN48| PB Diagnosis| 11   | No Dilution | OK       | APLs Test Set | APLs            | 23.005       |
| UPN53| BM Diagnosis| 9    | Dilution 90% | OK       | APLs Test Set | APLs            | 16.749       |
| UPN22| BM Diagnosis| 8    | Dilution 90% | OK       | APLs Test Set | APLs            | 15.781       |
| UPN40| BM Diagnosis| 7    | Dilution 90% | OK       | APLs Test Set | APLs            | 10.143       |
| UPN86| PB Diagnosis| 7    | No Dilution | OK       | APLs Test Set | APLs            | 11.641       |
| UPN50| PB Diagnosis| 6    | Dilution 90% | OK       | APLs Test Set | APLs            | 2.661        |
| UPN22| BM Diagnosis| 4    | Dilution 95% | OK       | APLs Test Set | APLs            | 7.665        |
| UPN53| BM Diagnosis| 4    | Dilution 95% | OK       | APLs Test Set | APLs            | 1.201        |
| UPN40| BM Diagnosis| 3    | Dilution 95% | OK       | APLs Test Set | APLs            | 10.667       |
| UPN50| PB Diagnosis| 3    | Dilution 95% | OK       | APLs Test Set | APLs            | 4.337        |
| UPN59| PB Diagnosis| 70   | No Dilution | Low RIN | APLs Test Set | APLs            | 26.497       |
| UPN82| BM Diagnosis| 64   | No Dilution | Low RIN | APLs Test Set | APLs            | 22.822       |
| UPN13| BM Diagnosis| 30   | No Dilution | Low cRNA | APLs Test Set | APLs            | 18.651       |
| Kasumi-1| BM Diagnosis| 100  | No Dilution | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 5.184        |
| Kasumi-1| BM Diagnosis| 100  | No Dilution | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 1.525        |
| UPN30| PB Diagnosis| 49   | Dilution 50% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 19.199       |
| UPN69| PB Diagnosis| 41   | Dilution 50% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 30.674       |
| UPN51| BM Relapse | 38   | Dilution 50% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 20.609       |
| UPN76| PB Diagnosis| 31   | No Dilution | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 32.067       |
| UPN69| BM Diagnosis| 30   | Dilution 50% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 25.183       |
| UPN26| BM Diagnosis| 27   | No Dilution | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 14.89        |
| UPN30| PB Diagnosis| 25   | Dilution 75% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 17.719       |
| UPN61| BM Diagnosis| 24   | Dilution 50% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 30.674       |
| UPN69| PB Diagnosis| 21   | Dilution 75% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 27.738       |
| UPN21| BM Diagnosis| 21   | No Dilution | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 24.116       |
| UPN95| BM Diagnosis| 20   | No Dilution | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 30.674       |
| UPN51| BM Relapse | 19   | Dilution 75% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 10.519       |
| UPN94| PB Diagnosis| 16   | No Dilution | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 28.631       |
| UPN69| BM Diagnosis| 15   | Dilution 75% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 22.744       |
| UPN30| PB Diagnosis| 10   | Dilution 90% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 16.958       |
| UPN69| PB Diagnosis| 6    | Dilution 90% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 13.468       |
| UPN30| PB Diagnosis| 5    | Dilution 90% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 13.581       |
| UPN61| BM Diagnosis| 5    | Dilution 90% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 28.443       |
| UPN51| BM Diagnosis| 4    | Dilution 90% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 3.096        |
| UPN69| PB Diagnosis| 3    | Dilution 95% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 12.875       |
| UPN51| BM Diagnosis| 2    | Dilution 95% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 4.6          |
| UPN61| BM Diagnosis| 2    | Dilution 95% | OK       | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 19.237       |
| UPN30| PB Diagnosis| 25   | No Dilution | Low cRNA | t(8;21)-AMLs Test Set | t(8;21)-AMLs | 15.256       |
Table 3 Characteristics and class assignment of the 117 Test Set samples. (Continued)

| UPN | BM/Diagnosis | 100 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
|-----|---------------|-----|-------------|----|-----------------------|--------------|
| ME-1 BM Diagnosis | 100 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN52 BM Diagnosis | 56 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN45 BM Diagnosis | 50 | Dilution 50% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN31 BM Diagnosis | 50 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN88 BM Diagnosis | 46 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN47 BM Diagnosis | 41 | Dilution 50% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN37 BM Diagnosis | 40 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN9 BM Diagnosis | 35 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN87 PB Diagnosis | 35 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN37 BM Diagnosis | 28 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN87 PB Diagnosis | 25 | Dilution 75% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN67 PB Diagnosis | 24 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN36 PB Diagnosis | 23 | No Dilution | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN47 BM Diagnosis | 20 | Dilution 75% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN37 BM Diagnosis | 20 | Dilution 50% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN10 BM Diagnosis | 16 | Dilution 75% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN52 BM Diagnosis | 14 | Dilution 75% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN37 PB Relapse | 14 | No Dilution | OK | inv(16)-AMLs Test set | t(8;21)-AMLs |
| UPN45 BM Diagnosis | 10 | Dilution 90% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN47 BM Diagnosis | 10 | Dilution 90% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN37 BM Diagnosis | 10 | Dilution 75% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN10 BM Diagnosis | 7 | Dilution 90% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN52 BM Diagnosis | 6 | Dilution 90% | OK | inv(16)-AMLs Test set | APLs |
| UPN45 BM Diagnosis | 5 | Dilution 95% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN47 BM Diagnosis | 5 | Dilution 95% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN7 BM Relapse | 5 | No Dilution | OK | inv(16)-AMLs Test set | t(8;21)-AMLs |
| UPN10 BM Diagnosis | 4 | Dilution 95% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN52 BM Diagnosis | 3 | Dilution 95% | OK | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN6 BM Diagnosis | 74 | No Dilution | Low cRNA | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN45 BM Diagnosis | 25 | No Dilution | Low cRNA | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN47 BM Diagnosis | 20 | No Dilution | Low cRNA | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN79 BM Diagnosis | 47 | No Dilution | Low cRNA | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN90 PB Diagnosis | 9 | No Dilution | Low cRNA | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN23 BM Diagnosis | 27 | No Dilution | Low RIN + Low cRNA | inv(16)-AMLs Test set | inv(16)-AMLs |
| UPN60 PB Diagnosis | 50 | No Dilution | OK | NK-AMLs Test Set | inv(16)-AMLs |
| UPN72 BM Diagnosis | 49 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN8 BM Diagnosis | 48 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN44 PB Diagnosis | 47 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN73 PB Diagnosis | 47 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN97 BM Diagnosis | 47 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN1 PB Diagnosis | 46 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN66 BM Diagnosis | 46 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN78 PB Diagnosis | 46 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN25 BM Diagnosis | 43 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN38 BM Diagnosis | 40 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN74 BM Diagnosis | 40 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN89 BM Diagnosis | 40 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
| UPN19 BM Diagnosis | 39 | No Dilution | OK | NK-AMLs Test Set | NK-AMLs |
| UPN11 BM Diagnosis | 37 | Dilution 50% | OK | NK-AMLs Test Set | NK-AMLs |
More precisely, for each sample, the following probe/marker landed each sample into its predicted class. Briefly, the confidence level of the class prediction analysis reflected the strength with which the overall model. This estimator enabled optimizing the number of probes/markers to be used per class. Briefly, the confidence level of the class prediction analysis reflected the strength with which the probe/marker landed each sample into its predicted class. More precisely, for each sample, the following happened: (1) each probe/marker casted its vote for the sample being in each of the train classes, (2) the votes were summed up over all probes/markers, yielding a vote for each of the classes, (3) the score of the winner class (call it S1) was compared to the second best (call it S2), and the confidence level of the classification of the sample was computed as (S1-S2)/(S1+S2). In the Cross-validation analysis, the same was performed, except that the probe/marker set was recomputed each time, leaving out the sample being tested in order to yield an honest estimate of the worth of the method. The fitness of a class prediction model was computed on the test samples only, in order to estimate how well the class prediction (of the “unclassified” samples) was, and reflected the success with which the model was able to correctly (re)classify the already “classified” samples. The fitness of the model was in fact the result of a cross-validation (as described above) on the Training samples, with the integral part being equal to the number of correct (re)classifications, and the fractional part being computed as follows: (1) if all samples were correctly classified, the average confidence level (as defined above) of all samples, (2) if some samples were misclassified, the average confidence level of the correctly classified samples minus the average confidence level of the misclassified ones. If all samples were misclassified, the fitness of the model was zero. Omics Explorer 2.2 software (Qlucore, Lund, Sweden - http://www.qlucore.com) was used for principal-component analyses. S-Plus® 8.0 Enterprise Developer software (Insightful Corporation, Seattle, USA) was used for all other statistical analyses. The two datasets (Training Set and Test Set) discussed in this publication have been deposited in NCBI’s Gene Expression Omnibus and are accessible.

### Table 3 Characteristics and class assignment of the 117 Test Set samples. (Continued)

| UPN | PB | Diagnosis | Dilution | OK | NK-AMLs Test Set | NK-AMLs |
|-----|----|-----------|----------|----|-----------------|---------|
| UPN96 | PB | Diagnosis | 35 | No Dilution | OK | NK-AMLs Test Set | NK-AMLs 27.303 |
| UPN72 | BM | Diagnosis | 25 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 26.707 |
| UPN8 | BM | Diagnosis | 24 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 27.357 |
| UPN44 | PB | Diagnosis | 24 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 29.235 |
| UPN97 | BM | Diagnosis | 24 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 24.035 |
| UPN1 | PB | Diagnosis | 23 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 26.707 |
| UPN66 | BM | Diagnosis | 23 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 19 |
| UPN73 | PB | Diagnosis | 23 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 24.104 |
| UPN78 | PB | Diagnosis | 23 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 28.541 |
| UPN25 | BM | Diagnosis | 21 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 18.99 |
| UPN38 | BM | Diagnosis | 20 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 25.107 |
| UPN74 | BM | Diagnosis | 20 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 19.705 |
| UPN89 | BM | Diagnosis | 20 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 26.555 |
| UPN11 | BM | Diagnosis | 19 | Dilution 75% | OK | NK-AMLs Test Set | NK-AMLs 24.587 |

UPN, unique patient number. BM, bone marrow. PB, peripheral blood. QCC, quality control criteria.

Low cRNA, amount of labeled cRNA hybridized on the BeadChip below 750 ng. Low RIN, RNA integrity number below 7.00.
Results
Identification of classifiers from optimal APL, t(8;21)-AML, inv(16)-AML and NK-AML samples
First, using the Class Prediction Module of ArrayMiner 5.3.3 software, classifiers associated with the APL, t(8;21)-AML, inv(16)-AML, NK-AML classes were defined based on 71 samples containing at least 60% of leukemic blasts and characterized by good quality control criteria (Optimal samples - Training Set). At this stage, unique and pooled normal bone marrow samples were also included in the model as the objective was to enhance the predictive capacity of the classifiers, especially for Test Set samples containing a majority of residual normal cells and a low leukemic blast load. Evaluating classifiers which included from 1 to 100 markers per class, all Training Set samples were assigned to the correct class when selecting from 3 to 46 markers per class (class prediction accuracy, sensitivity, specificity, negative and positive predictive values, 100% for each class - error rate, 0%), whether the source of leukemic cells was bone marrow or peripheral blood. The best model fitness was obtained with 8, 9 or 10 markers per classifier. To select the optimal model among these three, a principal-component analysis was performed. The highest cumulative percent of variance accounted for by the first three components was obtained with the 10-marker classifiers (cumulative percent of variance with classifiers including 10 markers, 79 percent - additional file 1, Figure S1), which led to select these for subsequent analyses (Figure 1). The best median confidence levels for the correct assignment of the Training Set samples were achieved with 10, 2, 1, and 9 markers for APLs, t(8;21)-AMLs, inv(16)-AMLs, and NK-AMLs, respectively (additional file 1, Figure S2). Regarding NK-AMLs and in agreement with previous reports, the highest median confidence levels were observed for NK-AMLs with mutated NPM1 and the lowest ones for NK-AMLs with neither NPM1 nor FLT3 mutations (additional file 1, Figure S3) [22-26].

Class prediction analysis for suboptimal AML samples - Impact of the leukemic blast load
Using the 10-marker classifiers, the GGA-based supervised method was subsequently applied to a series of 101 Test Set samples, with optimal quality control criteria, for which the leukemic blast load was originally below 60 percent or had been artificially lowered to less than 60 percent blasts by dilution series. Only the classifiers associated with one of the four AML classes were
All APL and t(8;21)-AML Test Set samples were correctly classified, even when containing as low as 2 percent blasts (Table 3 - Figure 2). Twenty six of the 29 inv(16)-AML samples were assigned to the expected class, including the 95 percent-diluted ones which contained as low as 3 percent blasts (error rate, 10 percent). In all three inv(16)-AML Test Set samples that were incorrectly classified, CBFB-MYH11 fusion gene was detected by FISH assay. For APL, t(8;21)-AML and inv(16)-AML classes, the blast load significantly influenced the confidence level of the class assignment (Figures 3 - additional file 1, Figure S4). For NK-AMLs, all but 1 of the 29 Test Set samples were correctly classified (error rate, 3 percent) (Table 3). This misclassified sample, obtained at diagnosis from UPN60, had no NPM1, FLT3 or CEBPA mutations. Its class assignment, characterized by a low confidence level (given its blast load - Table 1, additional file 1, Figure S4), was not confirmed as no PML/RARA fusion gene was detected by FISH analysis and RT-PCR assay.

Across the range of leukemic blast loads studied in the Test Set, the confidence level associated with the class assignment of diluted AML samples was not different from the one achieved for undiluted low leukemic blast load samples (Table 3 - additional file 1, Figure S5). For those 101 AML Test set samples with low leukemic blast load and optimal quality control criteria, the overall error rate was 3.9 percent.

Class prediction analysis for suboptimal AML samples - Impact of quality control criteria

When the performance of the 10-marker classifiers was assessed on 10 samples that did not fulfill all quality control criteria, all suboptimal samples were assigned to the expected class, even though their leukemic blast content was as low as 9 percent (median, 55 percent; range 9 to 100 percent) (Table 3 - additional file 1, Figure S6). The confidence level for the class assignment of these samples was not different from the one achieved for the other Test Set samples (Data not shown).

Overall, the error rate was 3.6 percent for the entire Test Set (excluding AML cell lines that were considered as controls), and for each class, sensitivity and specificity, negative and positive predictive values of the class assignments ranged from 91 to 100 percent (additional file 1, Figure S7).

Discussion

The integration of microarray-derived data to the current workflow dealing with the prognostic evaluation of AML patients requires the technology to deliver informative data for the majority of samples, including those

![Figure 2 Results of the class assignment for the 107 AML Test Set samples fulfilling all quality control criteria based on the 10-marker classifiers characterizing the APL, t(8;21)-AML, inv(16)-AML and NK-AML classes (all three AML cell lines run in duplicates included). Each column represents a sample; each row represents a marker (gene transcript). The log2 relative gene expression scale is depicted on the bottom left.](image)
with suboptimal characteristics. The current study focused on this kind of samples - either with a low leukemic blast load, poor quality control criteria or both - and the ability to identify favorable-risk karyotype AMLs in such situations, using microarray-based GEP and a class prediction method based on GGA. Apart from the study of de Ridder et al, which addressed the impact of random, fixed and group-specific impurities on the results of differential gene expression analyses, using a limited set of samples and computer simulations, the influence of the leukemic blast load on class prediction accuracy based on GEP data has never been specifically studied so far [25]. To evaluate the capacity of microarray-derived predictors to correctly assigned samples with a low leukemic blast load, besides using “real” AML samples with low blast percentage, i.e., less than 60 percent, a “dilution/mixture” approach was considered [26,27]. This strategy has already been successfully applied to assess the efficiency of various normalization methods applied to microarray datasets. It consisted in diluting labeled cRNAs of originally high blast load AML samples within pools of normal bone marrow labeled cRNAs. This approach was also considered as in some AML subtypes, such as t(8;21)-AMLs, more differentiated cells (not counted as blasts) are frequently of leukemic origin, leading to under-estimate the leukemic load of some “low blast content” samples. Consequently, it became possible to accurately control the percentage of leukemic blasts within the Test Set sample down to 2 percent.

Overall, considering the 111 Test Set samples with low leukemic blast load and/or suboptimal quality control criteria, whether being peripheral blood or bone marrow samples, the predictive capacity of the GGA-derived model with 10-marker classifiers was encouraging, with specificities ranging from 0.98 to 1.00, and sensitivities ranging from 0.91 to 1.00. Moreover, considering AML samples containing 20 to 56 percent blasts, the overall error rate was lower than 1.4 percent (1 misclassified NK-AML sample out of 72). Finally, APL, t(8;21)-AML, and inv(16)-AML samples containing as low as 2 percent blasts could be correctly classified, whether the classifiers mainly relied on a single marker, such as for inv(16)-AMLs and t(8;21)-AMLs, or on the overall set of 10 markers defining the classifier, as for APL samples. However, regarding APL, t(8;21)-AML and inv(16)-AML samples, the confidence level of the class assignment was correlated to the percentage of leukemic blasts within the studied samples. For NK-AML samples, such a finding was not observed, probably because NK-AMLs represent a more heterogeneous group of leukemias as compared to t(8;21)-AMLs, inv(16)-AMLs or APLs in term of oncogenic processes. This heterogeneity likely led to a higher variance of gene expression profiles within the NK-AML class, to the detriment of the interclass variance, hence influencing the class prediction confidence level more than the leukemic blast load. This hypothesis was supported by the fact that, within the NK-AML class, the best predictive capacity of the GGA was achieved for inv(16)-mutated samples, in accordance with previous studies, and the worst one for samples with no NPM1 mutations [28-31].

Unexpectedly, the worst results of the class prediction analysis were achieved for inv(16)-AML Test set samples (Sensitivity, 0.91; negative predictive value, 0.96). Among the three misclassified inv(16)-AML samples and for which the presence of CBFB/MYH11 fusion gene was confirmed by FISH analysis, all contained less than 20 percent blasts and two were obtained at the time of relapse - from patients UPN7 and UPN37 -. Regarding these two samples, which respectively contained 5 and 14 percent blasts, they were characterized by a low expression level of most markers defining the inv(16)-AML specific signature; especially MYH11, which expression level was within the estimated background signal range. It is noteworthy that for one of those two patients, UPN37, the bone marrow sample obtained at diagnosis was correctly classified, even when diluted at 50 and 75 percent (blast content, 20 and 10 percent,
For this patient, a clonal evolution of the disease was suspected as CD2, CD13, CD34 and CD117 expression values, assessed by flow cytometry on gated leukemic cells, and MYH11 expression level were significantly lower in the relapse sample as compared to the ones observed in the diagnostic sample. Verhaak et al in their study using Affymetrix GeneChips reported that MYH11 expression level was sufficient to identify inv(16)-AMLs [10]. The present results suggest that the use of MYH11 as the sole marker for the inv(16)-AML class could lead to unstable results in a class prediction model, especially in low leukemic blast load samples or in case of a clonal evolution at relapse. Of note, the inv(16)-AML signature developed on the AMLProfiler™ kit http://www.skylinediagnostics.com, which uses the Affymetrix™ platform, also includes a set of 19 markers [32].

As reported by Kohlmann et al., sub-optimal AML samples with poor quality control criteria, mainly because of low amounts of hybridized cRNAs, were all correctly classified, even though half of these contained less than 60 percent blasts (down to 9 percent for UPN90) [33]. Studying the impact of RNA degradation on GEP analyses, similar results have been recently reported by Opitz et al on a different microarray platform [34]. As in the present study, they showed that useful information could still be obtained from thermally degraded RNA; at least to a certain extent, and depending on the length of the mRNA molecules. Interestingly, even AML samples for which a third of the required cRNA amount was hybridized on BeadChips could be assigned to the correct class in the present study, suggesting that the biological variance between the AML classes was higher than the one related to technical variability or artifacts. This result could also be related to the fact that, during the sample processing, the TotalPrep RNA Amplification kit that was used, did not require any fragmentation step, which could further alter poor quality RNA or lower the already reduced cRNA amounts. These encouraging results, on a limited set of poor quality samples, could also be related to the classification method used, as the GGA-based class prediction strategy aimed at finding the most influential markers for each AML class (top markers with the highest inter-class variance and the lowest intra-class variance). This GGA-based class prediction method was indeed associated with the lowest error rate and the highest predictive accuracy when compared to other supervised methods, such as support vector machines, random forests, artificial neural networks, k-Nearest Neighbor or nearest shrunken centroids (also known as PAM) (data not shown).

Although, in most cases, these favorable prognostic chromosomal abnormalities and/or their related gene products can be detected by karyotype, FISH or PCR assay, they represent main targets for a microarray-based class prediction analysis to be identified before considering GEP as a useful tool in a routine workflow for prognostic assessment of AML patients. The fact that the AMLProfiler, a microarray-derived kit based on the Affymetrix™ platform, has been recently commercialized with the aim of achieving a molecular diagnosis for APLs, (t(8;21))-AMLs and inv(16)-AMLs, respectively using 27, 31 and 19 markers, emphasizes this assumption [32].

Finally, regarding the classifiers associated with APLs, (t(8;21))-AMLs or inv(16)-AMLs, 18 of the 28 markers identified in this study (64 percent) had already been reported in previous studies using various microarray technologies (additional file 1, Table S1). These findings (1) confirm that the Illumina bead-based technology is as reliable, robust and sensitive as other microarray technologies developed by commercial manufacturers or academic facilities, and (2) suggest that the GGA-derived class prediction approach is a highly efficient one as it required a limited set of ten markers per class to achieve accurate class assignments. Furthermore, this study has identified new AML markers that will need further studies to delineate their role in the leukemogenic events involved in APLs (CERCAM, COL23A1, LOC643201, LOXL4, SLC39A11), (t(8;21))-AMLs (LOC440030, TNFRSF21), and inv(16)-AMLs (EFHC2, GPR12, MEGF10).

Conclusions

In more than 96 percent of the suboptimal cases, using a microarray-derived GEP approach and a GGA-based class prediction method, favorable cytogenetic risk AML samples with low leukemic blast load and/or poor quality control criteria could be correctly assigned to the appropriate class with a limited set of markers, allowing to consider GEP as a useful tool in a routine workflow for prognostic assessment of AML patients.

Additional material

Additional file 1: Sample preparation for gene expression profiling

- Generation of biotinylated, amplified cRNA using The Illumina Total Prep RNA Amplification Kit (Applied Biosystems/Ambion, Austin, USA).
- Generation of low leukemic blast load AML samples by a "dilution/mixture" approach.
- Hybridization on Illumina HumanHT-12 v3 Expression BeadChips, staining and detection of cRNAs on microarrays using an I-Scan system.

Additional table S1 - Citations in previous studies of the identified markers. Additional figures S1 to S7. Figure S1. Three-dimensional projection of the 3 principal components in a principal-components analysis of APL, (t(8;21))-AML, inv(16)-AML, NK-AML, and Normal Bone Marrow samples belonging to the Training Set, with the use of the 10-marker classifiers. Figure S2. Per class median confidence level of the assignment for APL, (t(8;21))-AML, inv(16)-AML, NK-AML and normal bone marrow samples belonging to the Training Set according to the number of markers per class (from 1 to 100 markers per class).
1. Döhner H, Estey EH, Amadori S, Appelbaum FR, Büchner T, Burnett AK, de la Blétière.

Authors

3. Grimwade D, Walker H, Oliver F, Wellems T, Harrison C, Harrison G, Rees J, Hahn I, Stevens R, Burnett A, Goldstone A: The importance of diagnostic cytogenetics on outcome in AML: analysis of 1612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. Blood. 1998; 92:2322-33.

4. Breems DA, Van Putten WL, De Groot GE, Van Zelder-Rhoëla SL, Gerrissen-School KB, Mellink CH, Nieuwint A, Jotterand M, Hagemeijer A, Beverloo HB, Löwenberg B: Monosomy karyotype in acute myeloid leukemia: a better indicator of poor prognosis than a complex karyotype. J Clin Oncol 2008, 26:4791-9.

5. Isajil S, Grimwade D, Chase A, Goldstone A, Burnett A, Goldman JM, Swerdlow S: Identification of PML/RARA rearrangements in suspected acute promyelocytic leukemia using fluorescence in situ hybridization of bone marrow smears: a comparison with cytogenetics and RT-PCR in MRC ATRA trial patients. MRC Adult Leukaemia Working Party. Leukemia 2000, 14:950-3.

6. Grimwade D, Biondi A, Mazzoncacci MJ, Hagemeijer A, Berger R, Neat M, Howe K, Dastugue N, Jansen J, Radford-Weiss I, Lo Coco F, Lessard M, Hernandez JM, Delabesse E, Head D, Liso V, Sainty D, Flandrin G, Solomon E, Birg F, Lafage-Pochitaloff M: Characterization of acute promyelocytic leukemia cases lacking the classic t(15;17): results of the European Working Party. Groupe Francais de Cytogenetique Hématologique, Groupe de Francais d'Hematologie Cellulaire, UK Cancer Cytogenetic Group and BIOMED 1 European Community-Concerted Action "Molecular Cytogenetic Diagnosis in Haematological Malignancies". Blood. 2000, 96:1297-308.

7. Mitterbauer M, Kusec R, Schwarzer S, Haas OA, Lechner K, Jaeger U: Comparison of karyotype analysis and RT-PCR for AML/ETO in 204 unselected patients with AML. Ann Hematol 1998, 76:139-43.

8. Rowie D, Courtell SJ, Ross PM, Bunyan DJ, Vickers SJ, Bryon J, McMullan DJ, Griffiths MJ, Reilly JT, Vandenberghe EA, Wilson G, Watmore AE, Bowen NP: Cytogenetically cryptic AML-1-ETO and CBF beta-MYH11 gene rearrangements: incidence in 412 cases of acute myeloid leukaemia. Br J Haematol 2000, 111:103-1.

9. Mrózek K, Prior TW, Edwards C, Marucci G, Carroll AJ, Snyder PJ, Koduru PR, Thel KS, Pettenati MJ, Archer KJ, Caligiuri MA, Vardiman JW, Kolitz JE, Larson RA, Bloomfield CD: Comparison of cytogenetic and molecular genetic detection of t(8;21) and inv(16) in a prospective series of adults with de novo acute myeloid leukaemia: a Cancer Leukemia Group B Study. J Clin Oncol 2001, 19:3482-92.

10. Verhaak RG, Wenters BJ, Erpelincok CA, Abbas S, Beverloo HB, Lughart S, Löwenberg B, Delwel R, Valk PJ: Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. Haematologica 2009, 94:131-4.

11. Valk PJ, Verhaak RG, Beijen MA, Erpelincok CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, Beverloo HB, Moonhouse MJ, van der Spek PJ, Nieuwint A, Jotterand M, Hagemeijer A, Beverloo HB, Delwel R, Valk PJ: Prognostically useful gene expression profiles in acute myeloid leukemia. N Engl J Med 2004, 350:1617-28.

12. Bullinger L, Döhner K, Bier E, Föhlisch S, Schlenk RF, Tibshirani R, Döhner H, Pollack JR: Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med 2004, 350:1605-16.

13. Haffelbach T, Kholmann A, Wieczorek L, Basso G, Kronie GT, Béné MC, De Vos J, Hernández JM, Hofmann W, Mills K, Gikes A, Chiaretti S, Shurtleff SA, Kipps TJ, Rassenti LZ, Yeoh AE, Papenhausen PR, Liu WM, Williams PM, Foà R: Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: report from the International Microarray Innovations in Leukemia Study Group. J Clin Oncol 2010, 28:2529-37.

14. Haffelbach T, Kholmann A, Schnitter S, Dugas M, Hiddemann W, Kern W, Hiddemann W, Kern W, Schoch C: Global approach to the diagnosis of leukemia using gene expression profiling. Blood 2005, 106:1189-98.

15. Schoch C, Kholmann A, Schnitter S, Brods D, Dugas M, Mengenthaler K, Kern W, Hiddemann W, Eils R, Haffelbach T: Acute myeloid leukemias with from an international expert panel, on behalf of the European LeukemiaNet. Blood 2010, 115:453-74.

2. Baldus CD, Mrózek K, Marcucci G, Bloomfield CD: Clinical outcome of de novo acute myeloid leukaemia patients with normal cytogenetics is affected by molecular genetic alterations: a concise review. Br J Haematol 2007, 137:387-400.

References

1. Döhner H, Estey EH, Amadori S, Appelbaum FR, Büchner T, Burnett AK, Dombret H, Fenaux P, Grimwade D, Larson RA, Lo-Coco F, Rabbit T, Niedierwieser D, Oserenkoppel G, Saroz MA, Sierra J, Tallman MS, Löwenberg B, Bloomfield CD, European LeukemiaNet: Diagnosis and management of acute myeloid leukemia in adults: recommendations..
reciprocal rearrangements can be distinguished by specific gene expression profiles. Proc Natl Acad Sci USA 2002, 99:10008-13.

16. Kohlmann A, Schoch C, Schnittger S, Dugas M, Hiddemann W, Kern W, Haferlach T. Molecular characterization of acute leukemias by use of microarray technology. Genes Chromosomes Cancer 2003, 37:396-405.

17. Gutiérrez NC, López-Pérez R, Hernández JM, Isidro J, González B, Delgado M, Fernández E, García J, Vázquez L, González M, San Miguel JF. Gene expression profile reveals deregulation of genes with relevant functions in the different subclasses of acute myeloid leukemia. Leukemia 2005, 19:402-9.

18. Ross ME, Mahfouz R, Onciu M, Liu HC, Zhou X, Song G, Shurtleff SA, Pounds S, Ching C, Ma J, Ribeiro RC, Rubnitz JE, Girtman K, Williams WK, Ramondi SC, Liang DC, Shih LY, Pui CH, Downing JR. Gene expression profiling of pediatric acute myelogenous leukemia. Blood 2004, 104:3679-87.

19. Gabert J, Beillard E, van der Velden VH, Bi W, Grimwade D, Pallisgaard N, Barbany G, Cazzaniga G, Cayaula JM, Cavel H, Pane F, Aerts JL, De Micheli D, Thirion X, Pradel V, González M, Velthuys S, Malec M, Saglio G, van Dongen JJ. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion genes transcripts for residual disease detection in leukemia - a Europe Against Cancer program. Leukemia 2003, 17:2518-57.

20. Boisnel N, Rennieville A, Biggo V, Philippe N, Thomas X, Cayaule JM, Cavel H, Pane F, Aerts JL, De Micheli D, Thirion X, Pradel V, González M, Velthuys S, Malec M, Saglio G, Shurtleff SA, Pounds S, Ching C, Ma J, Ribeiro RC, Rubnitz JE, Pui CH, Downing JR. Gene expression profiling of pediatric acute myelogenous leukemia. Blood 2004, 104:3679-87.

21. Nakao M, Yokota S, Iwai T, Kaneko H, Horiike S, Kashima K, Sonoda Y, Fujimoto M, Misawa S. Internal tandem duplication of the FLT3 gene found in acute myeloid leukemia. Leukemia 1996, 10:1911-8.

22. Abu-Duhier FM, Goodeve AC, Wilson GA, Care RS, Peake IR, Reilly JT. Identification of novel FLT3 Aspa835 mutations in adult acute myeloid leukaemia. Br J Haematol 2001, 113:983-8.

23. Pabst T, Mueller BU, Zhang P, Radomiska S, Schnittger S, Behre G, Hiddemann W, Tenen DG. Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in adult myeloid leukemia. Nat Genet 2001, 27:263-70.

24. Falkenauer E. The grouping genetic algorithm. In Genetic Algorithms and Grouping Problems. First edition. Edited by: Falkenauer E. Chichester: Wiley J 1998:81-110.

25. de Riddere D, van der Linden CE, Schonewille T, Dik WA, Reinders MJ, van Dongen JJ, Staal FJ. Purity for clarity: the need for purification of tumor cells in DNA microarray studies. Leukemia 2005, 19:618-27.

26. Lemon WJ, Palatini JJ, Krahe R, Wright FA. Theoretical and experimental comparisons of gene expression indexes for oligonucleotide arrays. Bioinformatics 2002, 18:1470-6.

27. Ittazar RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003, 31:e15.

28. Mullighan CG, Kennedy A, Zhou X, Radtke I, Phillips LA, Shurtleff SA, Preudhomme C, Falini B, Pelicci PG. Impact of RNA degradation on gene expression profiling. Proc Natl Acad Sci USA 2005, 102:9987-92.

29. Opitz L, Salinas-Riester G, Grade M, Jung K, Jo P, Emons G, Ghadimi BM, Beissbarth T, Gaedcke J. Pattern robustness of diagnostic gene expression signatures in leukemia. Genes Chromosomes Cancer 2005, 42:299-307.

30. Veltor H, Döhner H, Lowenberg B. A validated diagnostic microarray for newly diagnosed adult AML patients, Oral Communication at European Hematology Association annual meeting London 2011. Abstract C469.

Pre-publication history

The pre-publication history for this paper can be accessed here: http://www.biomedcentral.com/10.1186/1755-8794-5-6/prepub

Cite this article as: de la Blétière et al.: Routine use of microarray-based gene expression profiling to identify patients with low cytogenetic risk acute myeloid leukemia: accurate results can be obtained even with suboptimal samples. BMC Medical Genomics 2012 5:6.

Submit your next manuscript to BioMed Central and take full advantage of:

• Convenient online submission
• Thorough peer review
• No space constraints or color figure charges
• Immediate publication on acceptance
• Inclusion in PubMed, CAS, Scopus and Google Scholar
• Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit