npg

SHORT COMMUNICATION

Oncogenic fusion proteins expressed in immature hematopoietic cells fail to recapitulate the transcriptional changes observed in human AML

N Rapin[1,2,3,4] and BT Porse[1,2,4]

Reciprocal chromosomal translocations are observed in one-third of acute myeloid leukemia (AML) cases. Targeting and understanding the effects of the resulting aberrant oncogenic fusion proteins may help developing drugs against specific leukemic subtypes, as demonstrated earlier by the use of ATRA in acute promyelocytic leukemia. Hematopoietic stem/progenitor (HSPCs) cells transduced with oncogenic fusion genes are regarded as promising *in vitro* models of their corresponding AML subtypes. Here, we critically assessed the potential of such *in vitro* models using an integrative bioinformatics approach. Surprisingly, we found that the gene-expression profiles of CD34 + human HSPCs transformed with the potent oncogenic fusion proteins AML-ETO or MLL-AF9, only weakly resembled those derived from primary AML samples. Hence, our work raises concerns as to the relevance of the use of *in vitro* transduced cells to study the impact of transcriptional deregulation in human AML.

INTRODUCTION

Acute myeloid leukemia (AML) arises from the transformation of normal hematopoietic stem/progenitor cells (HSPCs) mediated by mutations or chromosomal aberrations that are essential drivers of the leukemic process[1–3] In particular, chromosomal rearrangements leading to the expression of oncogenic fusion proteins are frequent in AML and have been widely studied.[3,4] Genes targeted by chromosomal rearrangements often encode transcriptional or epigenetic regulators and, consequently, expression of the resulting fusion proteins is associated with transcriptional deregulation of gene expression programs in the targeted cells. Indeed, gene expression profiling studies have shown that distinct oncogenic fusion proteins are associated with distinct gene-expression signatures thus, providing a molecular explanation for their differential impact on prognosis and survival.[5,6]

The importance of oncogenic fusion proteins in leukemogenesis is further supported by the ability of several of these to transform HSPCs *in vitro*.[7–13] Examples are fusion proteins such as MLL-ENL and MLL-AF9 (representative of translocations involving the *MLL1* locus at 11q23) as well as AML-ETO (t(8;21)), which can transform HSPCs from both mouse and human.[4,14,15] This have led to widespread use of fusion protein-driven tissue culture models of AML; however, the extent to which these faithfully mirror the transcriptional changes in primary human AMLs have not been rigorously tested.

We have recently developed a bioinformatics pipeline that allow us to identify transcriptional differences between cancer and its nearest normal counterpart.[16] In the present study, we used this approach to ask to what extent the gene expression changes conferred by the expression of either MLL-AF9 or AML-ETO in CD34 + HSPCs mirrored those observed in primary AML patient blasts.

RESULTS AND DISCUSSION

In order to determine the extent to which fusion protein-expressing HSPCs cultured *in vitro* mirrored the transcriptional changes observed in primary leukemic blasts, we collected microarray-based gene expression data from several sources (Table 1). These include normal HSPCs, empty vector-, MLL-AF9- and AML-ETO-transduced CD34 + cells cultured *in vitro* (6 h, 3 d or 8 d after transduction) as well as primary leukemic blast from patients with corresponding karyotypic lesions. Using our recent cancer versus normal (CvN) approach based on principal component analysis (PCA), we mapped gene expression profiles from *in vitro* cultured cells and patient samples onto the gene expression landscape of normal hematopoiesis (Figure 1a).[16] Strikingly, we find that the transduced cells cluster tightly as a function of time but independent of the expression of the transforming oncogene. Moreover, the oncogene-transduced CD34 + cells map nowhere near their respective patient counterparts. Therefore, these findings suggest that the main driver of the transcriptional changes of transduced cells is related to the culturing process and not to the expression of the oncogenic fusion protein. Indeed, when we quantify the extent of differentiation using a newly defined 'stemness' score, we find that the transduced cells progressively lose stemness (that is, they differentiate) over time in a manner independent of the expression of the fusion oncogene (Figure 1b).

We next used hierarchical clustering of the gene expression data sets of normal HSPCs, transduced cells and primary AML blasts to further assess their relationship (Figure 1c). In line with the PCA, we find that transduced CD34 + cells form a distinct cluster embedded in the normal HSPCs and that their behavior was independent of the expression of oncogenic fusion protein. Importantly, the patient samples harboring either MLL

[1]The Finsen Laboratory, Rigshospitalet, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark; [2]Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Copenhagen, Denmark; [3]Department of Biology, Faculty of Natural Sciences, The Bioinformatics Centre, University of Copenhagen, Copenhagen, Denmark and [4]Danish Stem Cell Centre (DanStem) Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark. Correspondence: Professor BT Porse, The Finsen Laboratory, Rigshospitalet, Faculty of Health Sciences, University of Copenhagen, Ole Maaløesvej 5, Copenhagen 2200, Denmark.
E-mail: bo.porse@finsenlab.dk

rearrangements or the AML-ETO translocation formed distinct clusters clearly separated from normal HSPCs and transduced cells (Figure 1c).

We next focused on the top 1% genes that were selectively upregulated in AML blasts versus their nearest counterpart using the CvN approach, as well as on the top 1% genes upregulated in cultured cells following fusion gene expression (Figure 1d). Specifically, we used a hypergeometric test to score the enrichment of genes belonging to previously reported gene-expression signatures associated with MLL-rearranged or t(8; 21)-driven AML. Strikingly, and in contrast to the primary AML samples, the expression of these signatures was only marginally enriched in fusion protein-expressing cultured cells, clearly demonstrating that they fail to induced the transcriptional program associated with the presence of these lesions in human AML (Figure 1d).

Next, we took a more gene-centric approach and identified deregulated genes ($|\log2 FC| > 1$, $P < 5e - 3$, moderated t-test) in either fusion protein-transduced cells or in AML blasts. This analysis revealed a very limited overlap between genes exhibiting deregulated expression in both transduced CD34 + cells and in AML blasts harboring corresponding karyotypic lesions (Figures 1e and f). Furthermore, when we attempted to correlate the extent of deregulation of aberrantly transcribed genes in transduced CD34 + cells with that observed in leukemic blasts, the correlation was very poor (Figures 1g and h). However, we do note that a subset of well-known MLL-fusion target genes, including HOXA9 and MEIS1,[17,18] exhibit very similar patterns of deregulation in MLL-rearranged AML and MLL-AF9-transduced CD34 + cells. This suggests that enforced expression of MLL-AF9 in CD34 + is able to recapitulate some, albeit a minor fraction, of the transcriptional changes associated with MLL-rearranged AML.

Finally, we took a pathway-centric approach to compare the transcriptional changes in AML blasts with those in transduced CD34 + cells. Specifically, we first identified gene signatures found to be significantly ($P < 1e - 5$, hypergeometric test) deregulated in AML blasts versus normal cells.[16] We next report their median fold changes, relative to control and nearest normal counterparts, in transduced CD34 + cells and in a subset of AML blast samples, respectively (Figure 2). Strikingly, while gene signatures associated with cell cycle processes are generally downregulated in AML blasts, we note that the transduced CD34 + cells only exhibit a transient downregulation of these pathways. Such behavior could potentially be associated with an adaptation to the culture conditions. Similarly, pathways found to be frequently upregulated in cancer patients,[19–21] such as immune response and various signaling pathways, are either unaffected or show opposite trends in transduced CD34 + cells compared with AML blasts. Hence in conclusion, the pathway-centric analyses clearly demonstrate that AML blasts and oncogene-transduced CD34 + cells express distinct transcriptional programs.

Previously, Wei et al.[15] reported good correlations between the transcriptional profiles derived from long-term cultures of transformed human cells and leukemic cells from patients with either MLL-rearranged or core-binding factor AML. However, only a small selection of genes found to be differentially expressed between the two AML subtypes were used to classify the in vitro-transformed cells, meaning that only a limited selection of the entire transcriptome was probed in that study. In contrast, here we have used an unbiased bioinformatics approach to analyze gene-expression profiles of AML blasts, normal HSPCs as well as transduced CD34 + to assess the extent to which the latter are likely to constitute good experimental surrogates of human AML.

Table 1. Source of the data

| Type	Cell types included	Reference (GEO accession)
Normal hematopoiesis	Hematopoietic stem cells; multi-potent progenitors; common myeloid progenitors; granulocyte–monocyte progenitors; megakaryocyte–erythrocyte progenitors; early promyelocytes; late promyelocytes; myelocytes; metamyelocytes; band cells; polymorphonuclear neutrophilic granulocytes; monocytes	Rapin et al.[16] (GSE42519)
Transformed CD34 + hematopoietic cells	Empty vector control CD34 + cells at 6 h, 3 days and 8 days; MLL-AF9-transformed CD34 + cells at 6 h, 3 days and 8 days; AML1-ETO-transformed CD34 + cells at 6 h, 3 days and 8 days; AML1-ETO, MLL-AF9	Abdul-Nabi et al.[22] (GSE57194)
Acute myeloid leukemia	Whole-bone marrow extracts from patients with AML with t(11q23)/MLL and AML with t(8;21)	Haferlach et al.,[11] Verhaak et al.[12,13] (GSE13159, GSE6891)

Abbreviation: AML, acute myeloid leukemia.

Figure 1. Side-by-side comparison of gene expression profiles derived from AML blasts and fusion gene-transduced CD34 + cells cultured in vitro. (**a**) Mapping of relevant samples into the PCA space of the hierarchy of normal hematopoiesis. The replicates of the different populations have been averaged into one data point for readability. Hematopoietic stem cells (HSCs); multi-potent progenitors (MPPs); common myeloid progenitors (CMPs); granulocyte–monocyte progenitors (GMPs); megakaryocyte–erythrocyte progenitors (MEPs); early promyelocytes (early PMs); late promyelocytes (late PMs); myelocytes (MYs); metamyelocytes (MMs); band cells (BCs); polymorphonuclear neutrophilic granulocytes (PMN_BM); monocytes (Mono); empty vector control CD34 + cells at 6 h (c_6 h), 3 days (c_3 d) and 8 days (c_8 d); MLL-AF9-expressing CD34 + cells at 6 h (mll_6 h), 3 days (mll_3 d) and 8 days (mll_8 d); AML-1ETO-expressing CD34 + cells at 6 h (eto_6 h), 3 days (eto_3 d) and 8 days (eto_8 d); leukemic blasts from patients with t(8;21) AML (AML with t(8;21)); leukemic blasts from patients with MLL-rearranged AML (AML with t(11q23)/MLL). The PCA was performed on 2119 probe sets selected by variance filtering.[16] (**b**) Stem cell score of gene expression profiles of transformed cells, AML blasts and normal HSPCs. (**c**) Hierarchical clustering of samples in (**a**) using genes from the gene signatures RAPIN_CVN_t(8;21)_up/_dn and RAPIN_CVN_t(11q23)_MLL_up/_dn.[16] (**d**) AML1-ETO- and MLL-published gene signatures enrichment represented as –log(P-value) for transformed cells and AML blasts (MLL signatures *$P < 0.05$, **$P < 0.001$, ***$P < 1e5$; AML1-ETO signatures °$P < 0.05$, °°$P < 0.001$, °°°$P < 1e5$). (**e**) Overlap between genes deregulated ($|\log2$-fold change$| > 1$, $P < 5e - 3$, moderated t-test) in AML with t(8;21) versus normal cells and AML1-ETO-transduced CD34 + cells versus control after 8 days of culture. (**f**) Overlap between genes deregulated ($|\log2$-fold change$| > 1$, $P < 5e - 3$, moderated t-test) in AML with t(11q23)/MLL versus normal cells and MLL-AF9-transduced CD34 + cells versus control after 8 days of culture. (**g**) Correlation between the extent of deregulation in AML with t(8;21) and transduced CD34 + cells of the genes selected in **e**. Genes displaying good correlation (AML blast fold change = transduced CD34 + cells fold change ± 0.25) are depicted. (**h**) Same as **g** for MLL-rearranged AML using genes selected in **f**.

Strikingly, we only observed a limited overlap between the fusion protein-driven gene expression changes in culture and those observed in AML blasts. We can only speculate as to the underlying reasons, but likely candidates could be the lack of additional mutations in the transduced CD34 + cells or the failure of the culture conditions to recapitulate the 'leukemic' niche in which AML blasts are normally residing. AML is not a monogenic disease and different driver mutations may deregulate distinct transcriptional programs or even collaborate to deregulate others. Similarly, human leukemic blasts are inherently difficult to grow in culture, suggesting that they receive distinct signals from the 'leukemic' niche that current culture systems fail to recapitulate. In conclusion, whereas fusion protein-expressing *in vitro* models certainly mimic some features of their corresponding AML, our work clearly demonstrate that they only recapitulate parts of the transcriptional deregulation observed in primary patient-derived leukemic material. Hence, our findings raise concerns as to the widespread use of fusion protein-expressing cultured HSPCs as a tool to understand the biology of human AML.

MATERIALS AND METHODS

Microarray-based gene expression data from *in vitro* transformed HSPCs (CD34 + controls and MLL-AF9/AML-ETO-transduced CD34 + cells at three distinct time points following transduction)[22] and primary AML blasts with either MLL rearrangements or t(8;21) translocations[11–13] were normalized alongside with distinct subsets of normal human HSPCs, as described in Rapin *et al.*[16]
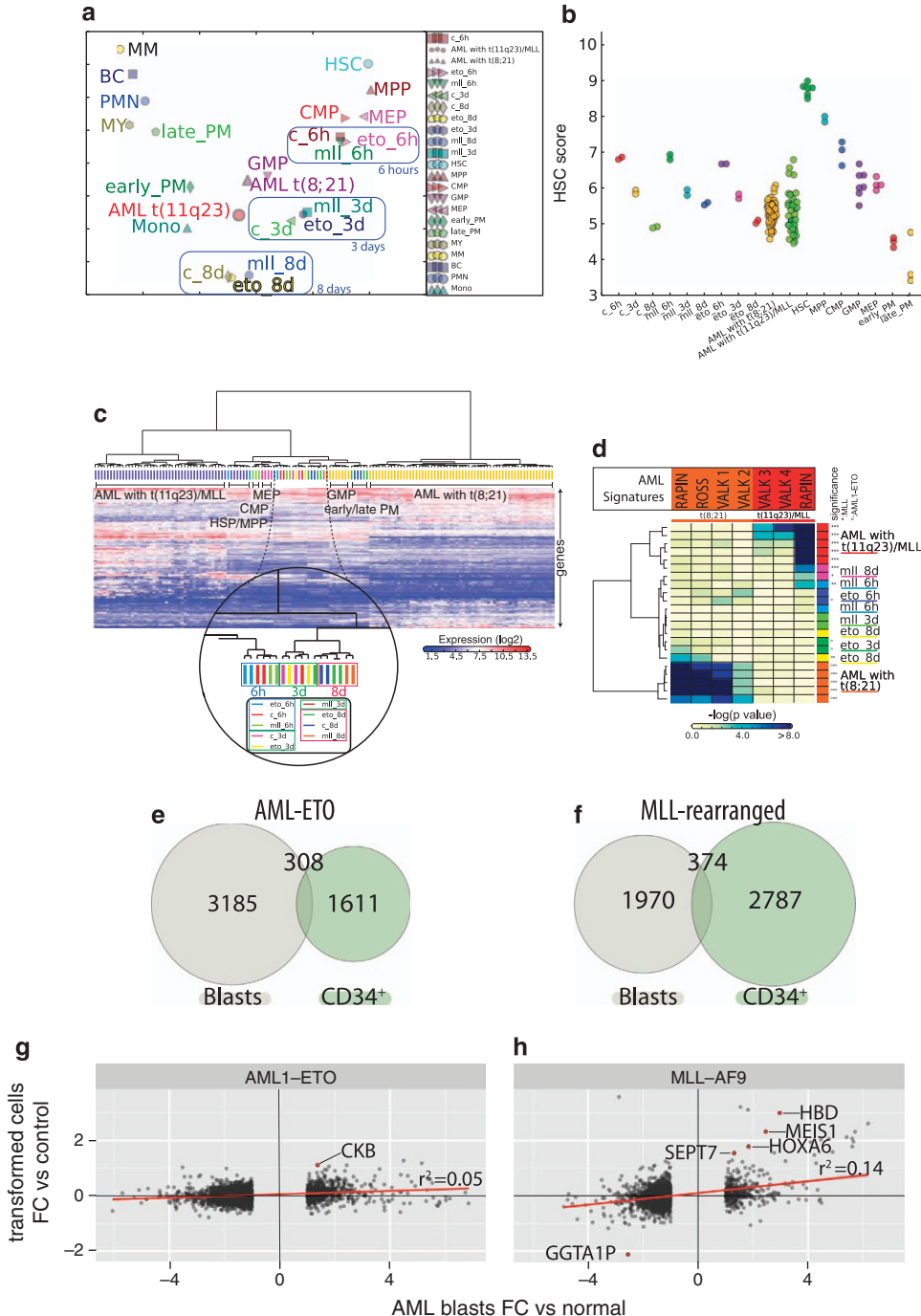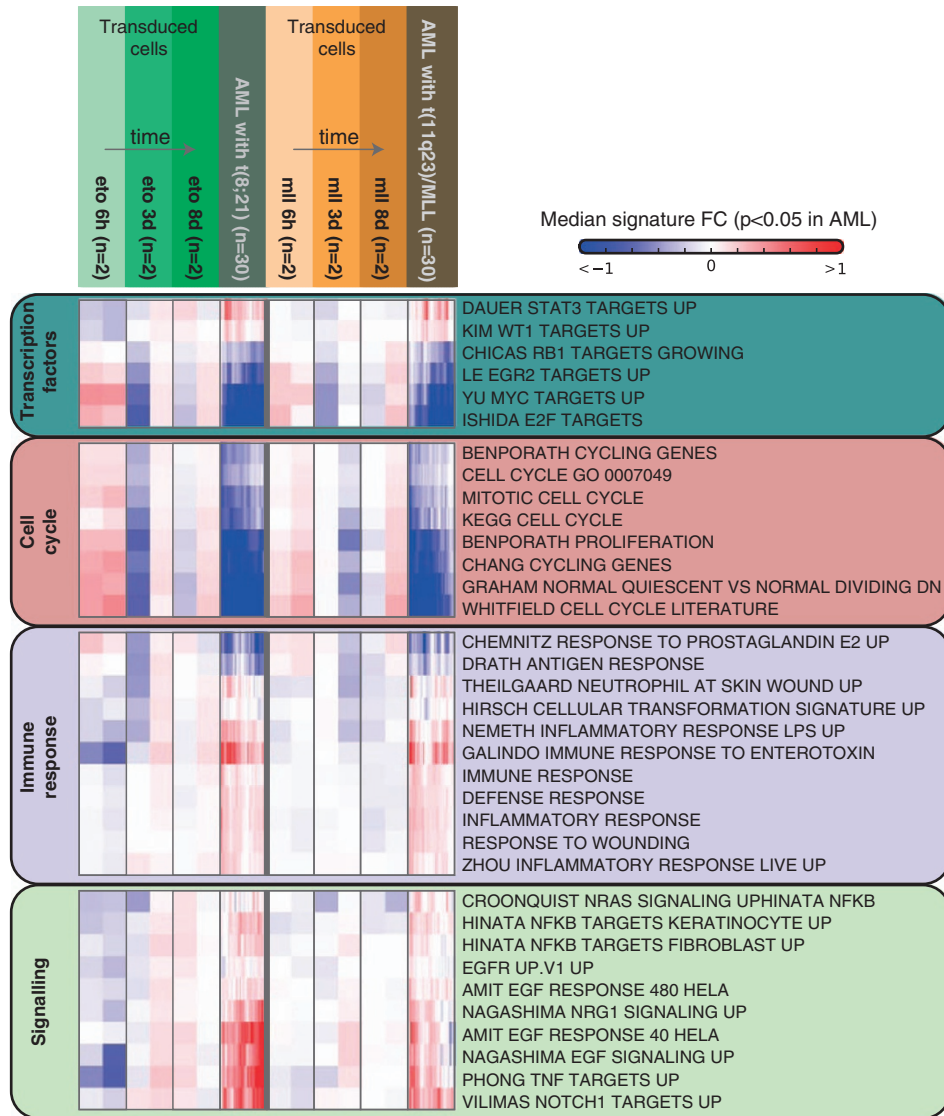
Figure 2. Pathway-centric comparison of transcriptional changes in AML blasts and CD34+ cells transduced with the same oncogenes. Median gene-expression fold change of selected MsigDB gene signatures that are significantly ($P < 1e - 5$) enriched among genes found to be deregulated in AML blasts samples. Fold change is computed relative to controls for transduced cells, and relative to the closest normal counterpart for AMLs blasts.

Following batch correction, we used PCA to construct a gene expression landscape of normal hematopoiesis and subsequently mapped individual CD34+-derived and AML samples to this space, thereby allowing us to identify the nearest normal counterpart for each sample. Fusion protein-mediated gene expression differences within the *in vitro* data set were calculated by comparing the fusion protein samples with vector controls sampled at identical time points. Gene expression differences between leukemic samples and corresponding normal were determined as described previously.[16,23] Genes were defined as deregulated by the following criteria: |log2-fold change > 1|, $P < 5e - 3$, Smyth's moderated t-test.[24]

To approximate the extent of cellular differentiation, we generated a stem cell signature defined by genes that are downregulated continuously at all stages of normal hematopoiesis, from stem cells to early promyelocytes ($P < 0.05$, t-test, see Rapin *et al*.[16] for details and gene list). We report the mean expression of the signature genes to quantify the degree of maturation.

Using a hypergeometric test with the top 1% upregulated genes in fusion protein-expressing CD34+ cells versus empty vector controls at each time point, or AML blasts versus normal counterpart, we quantify the enrichment of published AML t(8;21), MLL-fusion gene signatures[9,16,25,26] as well as MSigdb[26] signatures as described previously.[16] For the gene signatures, we also report the median fold change of all the genes in the signature for each individual sample.[16]

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1 Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult *de novo* acute myeloid leukemia. *N Engl J Med* 2013; **368**: 2059–2074.

2 Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC *et al.* The origin and evolution of mutations in acute myeloid leukemia. *Cell* 2012; **150**: 264–278.

3 Döhner K, Dohner H. Molecular characterization of acute myeloid leukemia. *Haematologica*. Ferrata Storti Foundation 2008; **93**: 976–982.

4 Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer* 2007; **7**: 823–833.

5 Grimwade D, Hills RK. Independent prognostic factors for AML outcome. hematology/the education program of the American Society of Hematology. American Society of Hematology. *Education Program*. American Society of Hematology 2009; **2009**: 385–395.

6 Metzeler KH, Hummel M, Bloomfield CD, Spiekermann K, Braess J, Sauerland M-C *et al.* An 86-probe-set gene-expression signature predicts survival in cytogenetically normal acute myeloid leukemia. *Blood* 2008; **112**: 4193–4201.

7 Smith ML, Hills RK, Grimwade D. Independent prognostic variables in acute myeloid leukaemia. *Blood Rev* 2011; **25**: 39–51.

8 Grimwade D, Hills RK, Moorman AV, Walker H, Chatters S, Goldstone AH *et al.* Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood* 2010; **116**: 354–365.

9 Valk PJM, Verhaak RGW, Beijen MA, Erpelinck CAJ, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM *et al.* Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med* 2004; **350**: 1617–1628.

10 Bullinger L, Döhner K, Bair E, Fröhling S, Schlenk RF, Tibshirani R *et al.* Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 2004; **350**: 1605–1616.

11 Haferlach T, Kohlmann A, Wieczorek L, Basso G, Kronnie GT, Béné M-C *et al.* 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**: 2529–2537.

12 de Jonge HJM, Valk PJM, Veeger NJGM, Elst ter A, Boer den ML, Cloos J *et al.* High VEGFC expression is associated with unique gene expression profiles and predicts adverse prognosis in pediatric and adult acute myeloid leukemia. *Blood* 2010; **116**: 1747–1754.

13 Verhaak RGW, Wouters BJ, Erpelinck CAJ, Abbas S, Beverloo HB, Lugthart S *et al.* Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. *Haematologica*. Ferrata Storti Foundation 2009; **94**: 131–134.

14 Mulloy JC, Cammenga J, MacKenzie KL, Berguido FJ, Moore MAS, Nimer SD. The AML1-ETO fusion protein promotes the expansion of human hematopoietic stem cells. *Blood* 2002; **99**: 15–23.

15 Wei J, Wunderlich M, Fox C, Alvarez S, Cigudosa JC, Wilhelm JS *et al.* Microenvironment determines lineage fate in a human model of MLL-AF9 leukemia. *Cancer Cell* 2008; **13**: 483–495.

16 Rapin N, Bagger FO, Jendholm J, Mora-Jensen H, Krogh A, Kohlmann A *et al.* Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. *Blood* 2014; **123**: 894–904.

17 Zeisig BB, Milne T, García-Cuéllar M-P, Schreiner S, Martin M-E, Fuchs U *et al.* Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Mol Cell Biol* 2004; **24**: 617–628.

18 Kroon E, Krosl J, Thorsteinsdottir U, Baban S, Buchberg AM, Sauvageau G. Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J* 1998; **17**: 3714–3725.

19 Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature*. Nature Publishing Group 2008; **454**: 436–444.

20 Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; **144**: 646–674.

21 De Marzo AM, Platz EA, Sutcliffe S, Xu J, Grönberg H, Drake CG *et al.* Inflammation in prostate carcinogenesis. *Nat Rev Cancer* 2007; **7**: 256–269.

22 Abdul-Nabi AM, Yassin ER, Varghese N, Deshmukh H, Yaseen NR. *In vitro* transformation of primary human CD34+ cells by AML fusion oncogenes: early gene expression profiling reveals possible drug target in AML. *PLoS ONE*. Public Library of Science 2010; **5**: e12464.

23 Bagger FO, Rapin N, Theilgaard-Mönch K, Kaczkowski B, Thoren LA, Jendholm J *et al.* HemaExplorer: a database of mRNA expression profiles in normal and malignant haematopoiesis. *Nucleic Acids Res* 2012; **41**(Database issue): D1034–D1039.

24 Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004 3 Article3.

25 Ross ME, Mahfouz R, Onciu M, Liu H-C, Zhou X, Song G *et al.* Gene expression profiling of pediatric acute myelogenous leukemia. *Blood*. American Society of Hematology 2004; **104**: 3679–3687.

26 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005; **102**: 15545–15550.