Supplementary Materials and Methods

Patient characteristics and molecular analyses

CEBPA mutations were assessed in a cohort of 598 cases of *de novo* AML. Detailed clinical and molecular characteristics were available for 524/598 cases (Table S4). These 524 patients were enrolled in the Dutch-Belgian Hemato-Oncology Cooperative Group (HOVON)–04, -10, -12, -29, -32, -42, or -43 protocols (available at http://www.hovon.nl). Reverse-transcription polymerase chain reaction (RT-PCR) and sequence analyses for *FLT3*-ITD, *FLT3*-TKD, *NPM1*, N-*RAS*, and K-*RAS*, mutations were performed as described previously.[1-3]

Detection of CEBPA mutations

Complementary DNA (cDNA) was generated from 1 μg of mRNA using SuperScript reverse transcriptase (Invitrogen, Carlsbad, CA). The *CEBPA* coding region was divided into three overlapping amplicons (Figure 1A). Primers for the three fragments (A, B and C) are indicated below (table). PCR amplification for all three fragments was carried out using 2 μl of cDNA in mixes containing 0.5 mM dNTPs, 10% DMSO, 2 mM $MgCl_2$, 0.4 μM of forward and reverse primer, 1X PCR buffer and 2.5 units of Taq polymerase (Invitrogen), in a total volume of 50 μl. Thermal cycling conditions for the three reactions were equal, i.e. denaturation at 94°C for 5 minutes, followed by 35 cycles of 94°C for 1 minute, 56°C for 1 minute and 72°C for 1 minute, and a final 5-minute elongation step of 72°C. After PCR amplification, 10 μl of PCR product was mixed with 10 μl of corresponding PCR product obtained from NB4 cell line cDNA. Heteroduplexes were allowed to form in an Applied Biosystems (Foster City, CA) GeneAmp PCR System 9700 (2 cycles of 95°C for 3 minutes, cooled to

20°C with a ramp of 5%, and maintained at 20°C for 5 minutes). The samples were then subjected to denaturing high-performance liquid chromatography (dHPLC) analysis on a Transgenomics (Omaha, NE) WAVE device, using temperatures of 65.4°C, 66.4°C and 65.5°C, respectively. Data were analyzed using Transgenomics software, and aberrant peaks were independently scored by two investigators. Samples with aberrant peaks were subjected to direct nucleotide sequencing on an Applied Biosystems 3100 device using the forward and reverse primers. In case a mutation was found, a second analysis on new input material was performed to rule out PCR-induced artifacts.

In AML cases for which dHPLC had revealed one single heterozygous mutation, the *CEBPA* coding region was fully sequenced to exclude the possibility that a second mutation had gone unnoticed. In three cases with an N-terminal mutation (#4336, #5362 and #5364), this extra analysis revealed an additional mutation, two of which were point mutations in the bZIP region.

Cases that appeared negative by dHPLC were additionally screened as follows. The *CEBPA* N-terminal part was nucleotide sequenced using previously described primers 2 and 10.[4] Insertions or deletions in the basic leucine zipper domain were detected using a previously described ethidium bromide agarose gel electrophoresis approach and subsequent nucleotide sequencing (primers 4 and 8) in cases with apparent abnormalities. [4]

Table: Primer sequences

Primer name	Location relative to XM_009180.3 (relative to main translational start site)	Sequence (5' to 3')
A fw	122 – 141 (-28 – -9)	CGCCATGCCGGGAGAACTCT
A rev	420 – 400 (270 – 250)	CTTCTCCTGCTGCCGGCTGT
B fw	361 – 380 (211 – 230)	GCCGCCTTCAACGACGAGTT
B rev	663 – 644 (513 – 494)	CTTGGCTTCATCCTCCTCGC
C fw	616 – 633 (466 – 483)	CGGCCGCTGGTGATCAAG
C rev	1254 – 1236 (1104 – 1086)	CCCAGGGCGGTCCCACAGC

Statistical analysis

Statistical analyses were performed in Statistical Package for the Social Sciences (SPSS, Chicago, IL) software, version 16.0. All patients received induction therapy and were included in the survival analysis. Actuarial probabilities of overall survival (OS, with death due to any cause) and event-free survival (EFS, with failure in case of no complete remission at day 1 [CR1] or relapse or death) were estimated by the method of Kaplan and Meier, and significance was assessed with the log rank test. Detailed definitions of all clinical endpoints for the various HOVON studies may be found in the study protocols which are accessible through http://www.hovon.nl. In brief, CR was defined as a bone marrow blast percentage less than 5% with peripheral blood recovery, without signs of extramedullary disease. Relapse was defined as recurrence of more than 5% blasts in the bone marrow (excluding increased blasts in the context of regenerating marrow) and recurrence of leukemia in peripheral blood/at extramedullary sites. Cox's proportional hazards models were fitted for multivariable analysis. The choice of variables to include was based on their a priori presumed prognostic impact. No further forward or backward selection procedure was applied,

so all variables were included in the final models. The proportional hazard assumption for *CEBPA* status was visually tested by plotting –ln(-ln(survival probability)) adjusted for the other covariates versus ln(analysis time). The three *CEBPA* curves were roughly parallel and there was no indication of non-proportionality.

Cytogenetic risk groups (favorable, intermediate, or poor) were defined as described.[1] Briefly, patients with inv(16)/t(16;16), t(8;21), and t(15;17) abnormalities, irrespective of the presence of additional cytogenetic aberrations, were considered as being in the favorable-risk category. These included a small number of cases in which the abnormality had been identified by RQ-PCR, despite normal cytogenetics. The poor-risk category was defined by the presence of –5/del(5q), –7del(7q), t(6;9), t(9;22), 3q26 abnormality, or complex karyotype (more than 3 abnormalities) in the absence of good risk cytogenetic characteristics. All other patients were classified as intermediate risk.

All tests were 2 tailed, and a *P* value of less than 0.05 was considered statistically significant.

Gene expression profiling analysis

Gene expression profiles of 524 cases of AML were derived using Affymetrix (Santa Clara, CA) HGU133Plus2.0 GeneChips. Sample processing and quality control were carried out as described previously.[5] Raw microarray data were processed using Affymetrix Microarray Suite 5 (MAS5) to target intensity values of 100. Intensity values lower than 30 were set at 30, and subsequently all data were log2 transformed. Gene expression profiling data are available at the NCBI Gene Expression Omnibus (accession number GSE14468). Gene expression classifiers for $CEBPA^{mut}$ and $CEBPA^{double\text{-}mut}$ were derived using Prediction Analysis for Microarrays (PAM)[6]

version 1.28 in R version 2.1.0. The method of the nearest shrunken centroids identifies a subgroup of genes that best characterizes a predefined class. In accordance with good practice guidelines[7,8], all available data were used for classifier construction, and predictive performances were estimated based on cross-validation as follows. PAM was first used to train a classifier based on the entire data set of 524 AML cases. Next, selection of a shrinkage factor (in order to only use the most informative genes) as well as estimation of classifier performance were carried out using 10-fold cross-validation, involving a random split of the data into 10 folds which was balanced with respect to mutation status. Each fold was once used as an independent validation set for a classifier that has been trained on the remaining 9 folds. The minimum number of misclassified cases was subsequently determined, and the corresponding shrinkage threshold was recorded. Furthermore, sensitivity and specificity were calculated. This entire procedure of 10-fold random cross-validation was repeated 100 times. Reported final classifiers represent the probe sets that remained after shrinkage using the median threshold over the 100 rounds of cross-validation. Reported final sensitivities and specificities represent the averages over the 100 rounds of cross-validation. Criterion for the *CEBPA*^{mut} classifier was minimum total misclassification rate (i.e. minimum false positives + false negatives). Criterion for the reported *CEBPA*$^{double\text{-}mut}$ classifier was minimum misclassification of double mutant specimens (i.e. minimum false negatives). We also assessed the possibility to derive a *CEBPA*^{mut} classifier that minimized the number of false negatives. However, it was not possible to find such a classifier that correctly predicted all mutant specimens. Furthermore, at the minimum misclassification rate (average of 6/38 misclassified; average sensitivity 84%), an unacceptably high number of false positives was found (i.e. an average of 85 cases; average specificity 83%).

Principal component analysis was performed using Spotfire Decision Site (Spotfire, Inc., Somerville, MA). Before the analysis, data for all probe sets were mean-centered.

References

1. Verhaak RG, Goudswaard CS, van Putten W, et al. Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. Blood. 2005;106:3747-3754.
2. Valk PJ, Bowen DT, Frew ME, Goodeve AC, Lowenberg B, Reilly JT. Second hit mutations in the RTK/RAS signaling pathway in acute myeloid leukemia with inv(16). Haematologica. 2004;89:106.
3. Care RS, Valk PJ, Goodeve AC, et al. Incidence and prognosis of c-KIT and FLT3 mutations in core binding factor (CBF) acute myeloid leukaemias. Br J Haematol. 2003;121:775-777.
4. Barjesteh van Waalwijk van Doorn-Khosrovani S, Erpelinck C, Meijer J, et al. Biallelic mutations in the CEBPA gene and low CEBPA expression levels as prognostic markers in intermediate-risk AML. Hematol J. 2003;4:31-40.
5. Valk PJ, Verhaak RG, Beijen MA, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med. 2004;350:1617-1628.
6. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A. 2002;99:6567-6572.
7. Simon R. Roadmap for developing and validating therapeutically relevant genomic classifiers. J Clin Oncol. 2005;23:7332-7341.
8. Dupuy A, Simon RM. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. J Natl Cancer Inst. 2007;99:147-157.