Cancer Cell
Article

DNA Methylation Signatures Identify Biologically Distinct Subtypes in Acute Myeloid Leukemia

Maria E. Figueroa,[1] Sanne Lugthart,[5] Yushan Li,[1] Claudia Erpelinck-Verschueren,[5] Xutao Deng,[2] Paul J. Christos,[3] Elizabeth Schifano,[7] James Booth,[7] Wim van Putten,[6] Lucy Skrabanek,[2,4] Fabien Campagne,[2,4] Madhu Mazumdar,[3] John M. Greally,[8] Peter J.M. Valk,[5] Bob Löwenberg,[5] Ruud Delwel,[5,*] and Ari Melnick[1,*]
[1]Department of Medicine, Hematology Oncology Division
[2]HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine
[3]Department of Biostatistics and Epidemiology
[4]Department of Physiology and Biophysics
Weill Cornell Medical College, New York, NY 10065, USA
[5]Department of Hematology
[6]Department of Trials and Statistics
Erasmus University Medical Center, Rotterdam, 3000 CA, the Netherlands
[7]Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853, USA
[8]Department of Genetics, Albert Einstein College of Medicine, New York, NY 10461, USA
*Correspondence: h.delwel@erasmusmc.nl (R.D.), amm2014@med.cornell.edu (A.M.)
DOI 10.1016/j.ccr.2009.11.020

SUMMARY

We hypothesized that DNA methylation distributes into specific patterns in cancer cells, which reflect critical biological differences. We therefore examined the methylation profiles of 344 patients with acute myeloid leukemia (AML). Clustering of these patients by methylation data segregated patients into 16 groups. Five of these groups defined new AML subtypes that shared no other known feature. In addition, DNA methylation profiles segregated patients with *CEBPA* aberrations from other subtypes of leukemia, defined four epigenetically distinct forms of AML with *NPM1* mutations, and showed that established AML1-ETO, CBFb-MYH11, and PML-RARA leukemia entities are associated with specific methylation profiles. We report a 15 gene methylation classifier predictive of overall survival in an independent patient cohort ($p < 0.001$, adjusted for known covariates).

INTRODUCTION

Acute myeloid leukemia (AML) is a highly heterogeneous disease from the biological and clinical standpoint. This remains a significant barrier toward the development of accurate clinical classification, risk stratification, and targeted therapy of this disease. Epigenetic control of gene expression has been suggested to play a pivotal role in determining the biological behavior of cells. One such epigenetic mechanism is DNA cytosine methylation, which can alter gene expression by creating new binding sites for methylation-dependent repressor proteins (Jones et al., 1998; Nan et al., 1998), or by disrupting the ability of transcription factors to bind to their target sequences (Kanduri et al., 2000; Watt and Molloy, 1988). In normal development, the proper distribution of DNA methylation plays a critical role in tissue differentiation and homeostasis (Li et al., 1992; Okano et al., 1999). Disruption of normal DNA methylation distribution is a hallmark of cancer and can play critical roles in initiation, progression, and maintenance of the malignant phenotype. For example, aberrant hypermethylation and silencing of certain tumor suppressor genes such as $p15^{CDKN2B}$ has been widely reported in leukemias and other myeloid neoplasms (Cameron et al., 1999; Christiansen et al., 2003; Shimamoto et al., 2005; Toyota et al., 2001). We recently showed that hypermethylation

SIGNIFICANCE

We show that large-scale genome-wide DNA methylation profiling reveals the existence of distinct DNA methylation patterns in AML and identifies novel, biologically and clinically relevant defined AML subgroups. Additionally, we demonstrate that despite these distinct patterns, a set of genes can be identified that is consistently aberrantly methylated and silenced in AML versus normal controls, indicating their likely involvement as a common epigenetic pathway in the leukemic transformation process. Finally, we describe a 15 gene DNA methylation classifier capable of predicting overall survival in an independent cohort of patients and validated as an independent risk factor in a multivariate analysis, demonstrating the potential of epigenetic markers for use even in patients for whom clinical biomarkers are not currently available.

Table 1. Patient Characteristics

Gender	Total (%)
Male	188 (54)
Female	156 (46)
Age	Total (%)
< 60 years	294 (85%)
> 60 years	50 (15%)
Median years (range)	48 (15-77)
FAB	Total (%)
M0	12 (3.5%)
M1	75 (21.8%)
M2	82 (23.8%)
M3	9 (2.6%)
M4	65 (18.9%)
M5	70 (20.3%)
M6	3 (0.87%)
NA*	28 (8.1%)
Cytogenetics	Total (%)
inv(16)/t(16;16)	30 (9%)
t(8;21)	24 (7%)
t(15;17)	10 (3%)
t(9;22)	2 (0.6%)
t(6;9)	3 (0.9%)
t(*v*;11q23)	13 (3.8%)
3q abnormalities	2 (0.6%)
del5(q)/del7(q)	19 (5.5%)
Trisomy 8	14 (4%)
del9q	8 (2.3%)
Complex	8 (2.3%)
Normal	152 (44%)
Other	43 (12.5%)
NA[a]/Failure	13 (3.8%)
Cytogenetic risk	Total (%)
Favorable	53 (15%)
Intermediate	231 (67%)
Unfavorable	47 (14%)
NA[a]	14 (4%)
CEBPA abnormalities	Total (%)
Double mutation	24 (7%)
Single mutation	11 (3.1%)
Silenced	8 (2.4%)
Wild-type	301 (87.5%)
NPM1 mutation	Total (%)
Negative	239 (69.5%)
Positive	105 (30.5%)
FLT3-ITD	Total (%)
Negative	248 (72%)
Positive	96 (28%)
EVI1 abnormalities	Total (%)
Negative	317 (92%)
Positive	27 (8%)

For more patient details, please see Table S1.
[a] NA, not available.

and silencing of the master regulatory transcription factor *CEBPA* was associated with a leukemia entity with T cell/myeloid features, hypermethylation of a number of additional transcriptional regulators, and distinctive biological features (Figueroa et al., 2009b; Wouters et al., 2007).

Based on these data, we hypothesized that DNA methylation distributes into specific patterns in cancer, and that these methylation profiles impose and reflect critical biological differences with practical clinical and therapeutic implications. In order to test this hypothesis, we performed a comprehensive exploration of DNA patterning in human disease, focusing on a well-characterized cohort of 344 patients with AML.

RESULTS

AML Is Composed of Epigenetically Distinct Diseases

Because the molecular heterogeneity of AML remains only partially resolved, the first goal of our study was to determine whether DNA methylation profiling could identify new clinically and biologically relevant disease subtypes. For that purpose, blast cells of 344 newly diagnosed AML patients were subjected to DNA methylation profiling of over 50,000 CpG dinucleotides contained within ~14,000 unique gene loci using the HELP (*H*paII tiny fragment *e*nrichment by *l*igation-mediated *P*CR) method (Figueroa et al., 2009a; Khulan et al., 2006). Table 1 summarizes patients' characteristics. DNA methylation measured by HELP was highly concordant with a quantitative single locus DNA methylation validation assay (correlation coefficient r = −0.88) in these AML patients (see Figure S1A available online). An unsupervised analysis using hierarchical clustering (1 - Pearson correlation distance and Ward's clustering method) showed that leukemias could be distinctly grouped according to their methylation profiles. A cut-off of 16 clusters was selected for further analysis since this segregation most accurately overlapped with the currently known molecular subtypes of AML while at the same time revealing the existence of additional epigenetic differences among the remaining patients. The stability of these clusters was verified by performing comparison of multiple cluster analyses using a decreasing number of probe sets (based on alternative cutoffs of across-patient standard deviation, Figures S1B–S1E). Table 2 shows the clinical, cytogenetic, and molecular features of each of the 16 clusters. Three of these patient clusters correspond to AML subtypes defined by the World Health Organization classification (WHO, 2008) (Figure 1), another eight clusters were enriched for cases harboring specific genetic or epigenetic lesions, and the remaining five clusters could not be explained by any known morphologic, cytogenetic, or molecular feature. Each of these DNA methylation-defined AML subtypes displayed a unique epigenetic signature when compared with normal bone marrow CD34+ cells (Figure 2 and Tables S3A–S3P). Taken together, these data indicate that DNA methylation is not randomly distributed in AML blasts but rather is organized into highly coordinated and well-defined