

Tracking myeloid malignancies by targeted analysis of successive DNA methylation at neighboring CpG dinucleotides

Myeloid neoplasms are reflected by aberrant DNA methylation, which can be used for disease stratification.^{1,2} So far, these methods have been based primarily on comparisons of the levels of DNA methylation at individual CpG dinucleotides (CpG), subsets of CpG, or the genome-wide variability of DNA methylation profiles.^{3,4} It has also been shown that the analysis of DNA methylation levels of adjacent CpG can provide information on disease progression in patients with myeloid malignancies.⁵ In this study, we built on the assumption that malignant transformation entails patient-specific epigenetic patterns at successive CpG sites that can be identified by targeted sequencing approaches. We propose two

alternative scoring systems based on variation of DNA methylation levels at neighboring CpG within a genomic region: the delta score (d-score) and the neighborhood score (n-score). The d-score represents the sum of residuals of DNA methylation levels as compared to those of age-adjusted controls, whereas the n-score is the sum of absolute differences in DNA methylation between neighboring CpG. We focused on seven neighboring CpG within an age-associated region of the phosphodiesterase 4C gene (*PDE4C*). This region was analyzed by pyrosequencing in 91 samples taken from patients with acute myeloid leukemia (AML) either at first diagnosis or upon treatment, in 126 peripheral blood samples from patients with myelodysplastic syndromes (MDS), and in 155 samples from healthy control. Our results demonstrate that both the d-score and the n-score can be applied to a short amplicon to reliably discriminate healthy and malignant samples.

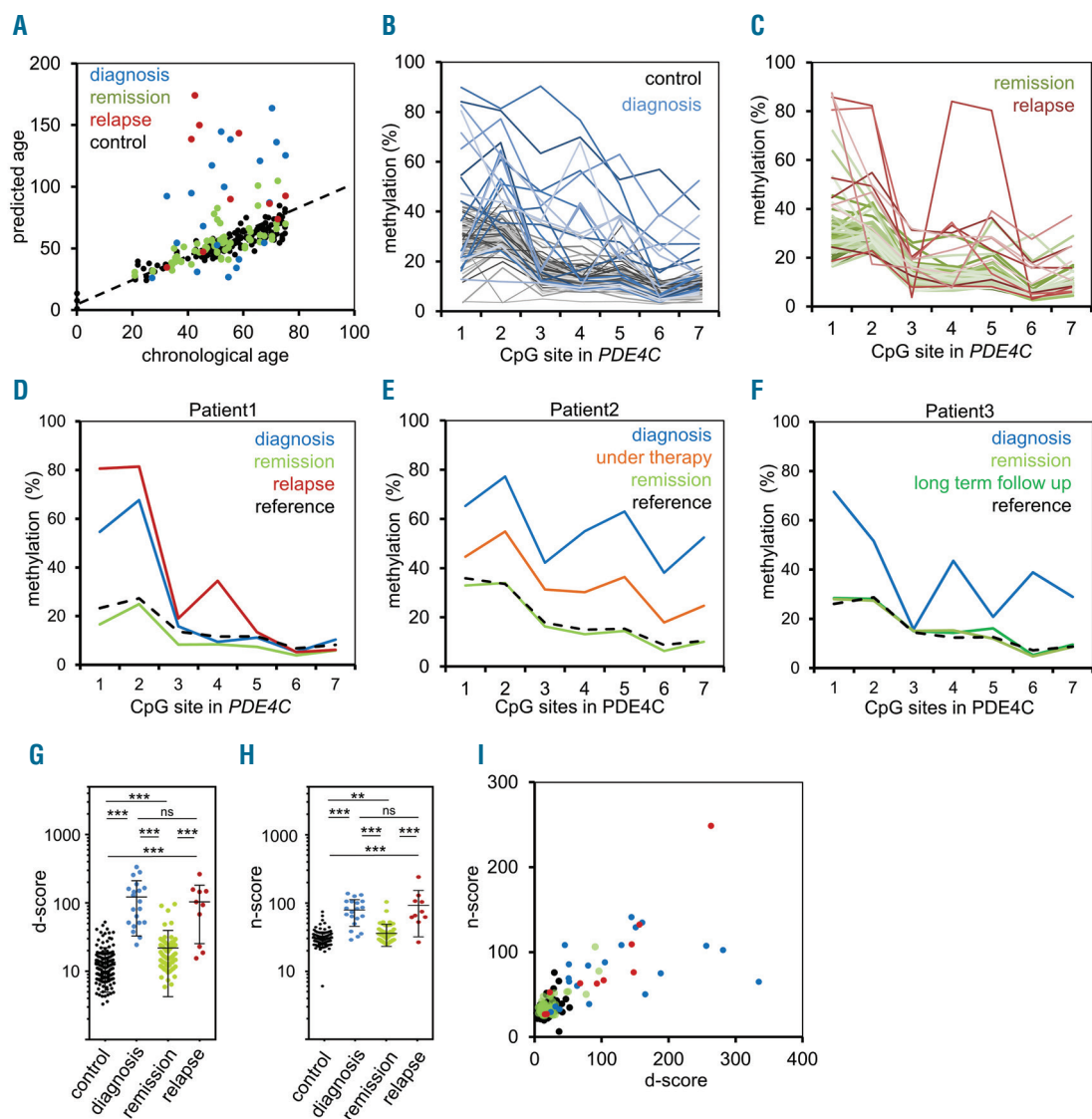


Figure 1. Aberrant methylation patterns of neighboring CpG sites are patient-specific. (A) Epigenetic age predictions were made based on three CpG sites (in the *ASP1*, *ITGA2B* and *PDE4C* genes) in samples taken from 20 patients with acute myeloid leukemia (AML) at first diagnosis, ten patients with persistent or relapsed AML, and 61 AML patients in complete remission. For comparison we analyzed blood from 155 healthy controls. (B, C) Methylation of the seven neighboring CpG sites of the pyrosequencing assay for *PDE4C* are shown for samples from healthy controls and patients at different stages of disease progression. (D-F) Representative analysis of methylation patterns for three AML patients at different stages of disease. The dashed lines indicate age-adjusted reference values. (G, H) Comparison of the d-scores and n-scores for healthy controls and AML patients at first diagnosis, in complete remission, and at relapse. (I) The association of the n-scores and d-scores. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns: not significant.

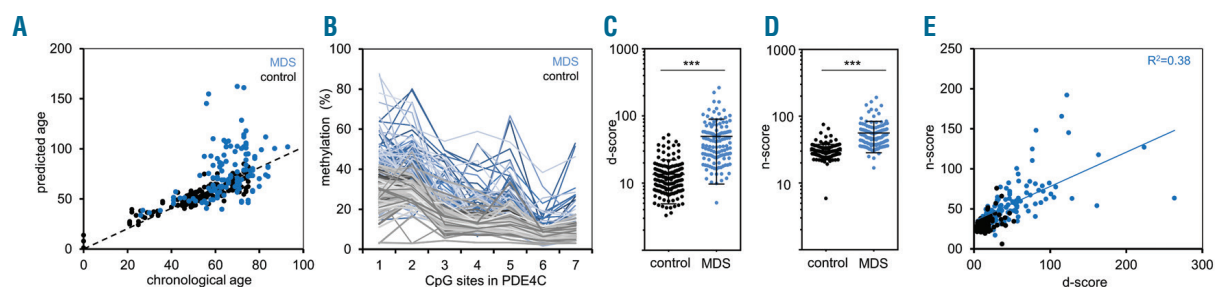


Figure 2. Variability scores are indicative of life-expectancy in patients with myelodysplastic syndromes. (A) Epigenetic age predictions were based on three CpG sites (in the *ASPA*, *ITGA2B* and *PDE4C* genes) for 126 patients with myelodysplastic syndromes (MDS) and 155 healthy controls. (B) DNA methylation of the seven CpG sites of *PDE4C* (pyrosequencing). (C,D) The d-scores and n-scores are shown for healthy controls and MDS patients at first diagnosis. (E) The association of the n-scores and d-scores is illustrated for healthy controls and MDS patients. *** $P < 0.001$.

Aging is reflected by highly reproducible DNA methylation changes at specific sites in the genome that can be used to estimate donor age.⁶ In blood samples from healthy individuals the mean DNA methylation levels at these dynamic regions reveal a cross-section of multiple, different hematopoietic clones. In contrast, in malignancies, age-associated regions usually reveal aberrant patterns due to the prevailing DNA methylation pattern of the malignant clone.⁷ In fact, age predictions based on three CpG sites associated with the genes aspartoacylase (*ASPA*), integrin alpha 2b (*ITGA2B*) and *PDE4C*⁶ revealed strong correlation with chronological age in 155 peripheral blood samples from healthy controls [$R^2 = 0.82$; mean absolute deviation (MAD) = 5.1 years], whereas epigenetic age predictions revealed a high offset in 20 AML samples taken at first diagnosis ($R^2 = 0.2$; MAD = 41.1 years). Upon treatment, age predictions of 61 AML patients in complete remission correlated with chronological age ($R^2 = 0.51$; MAD = 8.2 years), while the discrepancy remained in ten patients with persistent disease or relapse ($R^2 = 0.01$; MAD = 49.13 years) (*Online Supplementary Table S1*). These findings indicate that it was indeed the malignant clone that falsified the age predictions (Figure 1A).

We subsequently analyzed whether aberrant DNA methylation levels are co-regulated at the seven neighboring CpG sites of the *PDE4C* pyrosequencing amplicon. The amplicons of *ASPA* and *ITGA2B* could not be considered as they only included two and three CpG, respectively. In healthy blood samples the DNA methylation levels at the neighboring CpG of the *PDE4C* amplicon followed a reproducible pattern. In contrast, samples from patients with AML displayed marked variation, whereas samples taken in complete remission again followed the pattern observed in healthy donors (Figure 1B,C). We reasoned that the outgrowth of the malignant, epigenetically altered clone is reflected by patient-specific variations in DNA methylation between neighboring sites, rather than by absolute DNA methylation levels at individual CpG. In fact, we observed that distinct DNA methylation patterns at the neighboring CpG sites seemed to be patient-specific and recurred upon relapse of the disease, indicating that the aberrant pattern was preserved in the residual malignant cells during treatment (Figure 1D-F). Consequently, we aimed to quantify the variations of DNA methylation pattern by the sum of the absolute residuals between observed and age-adjusted DNA methylation levels at individual CpG sites (d-score). Alternatively, we did not take the reference DNA methy-

lation levels into account, but only the sum of absolute differences of DNA methylation between neighboring CpG sites (n-score). Both scores were significantly increased in samples taken from patients at first diagnosis ($n=20$) and at relapse or with persistent disease ($n=10$), whereas they normalized at complete remission ($n=61$) (Figure 1G,H). There was a tendency for samples with higher blast counts to have higher d-scores and n-scores (*Online Supplementary Figure S1*). In addition, we validated the scoring systems using barcoded bisulfite amplicon next-generation sequencing.⁸ This technique has the additional advantage that samples can be multiplexed and larger amplicons can be analyzed. In fact, healthy and malignant samples were discriminated more clearly (*Online Supplementary Figure S2*). On the other hand, the targeted pyrosequencing approach was faster and more cost-effective. There was a moderate correlation between d-score and n-score in AML samples, indicating that a combined approach might be advantageous for discriminating healthy and malignant samples (Figure 1I).

To test whether our scoring systems were also applicable to myeloid malignancies with lower clonal cell burden, we analyzed a cohort of 126 peripheral blood samples from patients before treatment of MDS. In analogy to AML patients, age in MDS patients was overestimated by a mean of 10 years (MAD=16.7 years) (Figure 2A). In comparison to the methylation pattern of controls, that of MDS patients revealed higher variability in the *PDE4C* amplicon (Figure 2B) as well as significantly increased d-scores and n-scores ($P < 0.0001$) (Figure 2C,D). The correlation of both scores supports the notion that they are of complementary use for identifying clonal diseases, including MDS (Figure 2E). The blast counts in peripheral blood were not associated with the n-score or d-score (*Online Supplementary Figure S3*), which might be due to the fact that not all derivatives of the malignant clone are classified as blasts. Notably, Cox regression analysis indicated that patients with higher d-scores (hazard ratio: 1.01; 95% confidence interval: 1.004-1.017; $P=0.001$) and n-scores (hazard ratio: 1.02; 95% confidence interval: 1.007-1.027; $P < 0.001$) had significant shorter overall survival.

Other authors have suggested that variability of methylation at a given locus, not the mean methylation value, is the best predictor of outcome.^{5,9,10} Teschendorff *et al.* analyzed variations in genome-wide inter-CpG correlations in DNA methylation during malignant transformation of cervical carcinomas. They identified subsets of genomic loci that exhibit large changes in covariance of

DNA methylation prior to malignant transformation. These loci are indicative of a high risk of neoplasia.¹¹ Li *et al.* examined epigenetic heterogeneity as assessed by DNA methylation within defined genomic loci in AML patients and observed that variation in DNA methylation is linked to inferior outcome.⁵ Their identification of epigenetically shifted loci (eloci) was also based on adjacent CpG sites, but they used a different algorithm to apply to genome-wide DNA methylation profiles.¹² Our data indicate that meaningful insight into variability of DNA methylation can also be derived from targeted approaches, which are much faster, less costly, and less labor-intensive.

In conclusion, we demonstrate here that the DNA methylation patterns at successive neighboring CpG in clonal myeloid disorders such as MDS and AML are patient-specific and recur with the reappearance of the malignant clone during relapse. The variability of DNA methylation within short amplicons can clearly discriminate healthy and malignant samples. This may provide a new perspective to track myeloid malignancies for early detection, risk stratification and evaluation of minimal residual disease after treatment.

Monika Eipel,¹ Tanja Božić,¹ Anna Mies,² Fabian Beier,³ Edgar Jost,³ Tim H. Brümmendorf,³ Uwe Platzbecker⁴ and Wolfgang Wagner¹

¹Helmholtz-Institute for Biomedical Engineering, Stem Cell Biology and Cellular Engineering, RWTH Aachen University, Aachen;

²Universitätsklinikum Carl Gustav Carus der TU Dresden, Medizinische Klinik und Poliklinik I, Dresden; ³Department of Hematology, Oncology, Hemostaseology and Stem Cell Transplantation, Medical Faculty University Hospital Aachen, RWTH Aachen University, Aachen and ⁴Universitätsklinikum Leipzig, Universität Leipzig, Medizinische Klinik und Poliklinik I, Hämatologie und Zelltherapie, Leipzig, Germany

Funding: this work was funded by the Interdisciplinary Center for Clinical Research within the Faculty of Medicine at the RWTH Aachen University (O1-3); by the Else Kröner-Fresenius-Stiftung (2014_A193); by the Deutsche Forschungsgemeinschaft (DFG; WA 1706/8-1 and WA1706/11 1), and by Deutsche Krebshilfe (TRACK-AML).

Correspondence: WOLFGANG WAGNER.

wagner@ukaachen.de

doi:10.3324/haematol.2018.209734

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Figueroa ME, Lugthart S, Li Y, et al. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell*. 2010;17(1):13-27.
2. Gebhard C, Glatz D, Schwarzfischer L, et al. Profiling of aberrant DNA methylation in acute myeloid leukemia reveals subclasses of CG-rich regions with epigenetic or genetic association. *Leukemia*. 2011;33(1):26-36.
3. Jost E, Lin Q, Ingrid WC, et al. Epimutations mimic genomic mutations of DNMT3A in acute myeloid leukemia. *Leukemia*. 2014;28(6):1227-1234.
4. Bozic T, Lin Q, Froebel J, et al. DNA-methylation in C1R is a prognostic biomarker for acute myeloid leukemia. *Clin Epigenetics*. 2015;7:116.
5. Li S, Garrett-Bakelman FE, Chung SS, et al. Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia. *Nat Med*. 2016;22(7):792-799.
6. Weidner CI, Lin Q, Koch CM, et al. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol*. 2014 2/3/2014;15(2):R24.
7. Lin Q, Wagner W. Epigenetic aging signatures are coherently modified in cancer. *PLoS Genet*. 2015;11(6):e1005334.
8. Franzen J, Zirkel A, Blake J, et al. Senescence-associated DNA methylation is stochastically acquired in subpopulations of mesenchymal stem cells. *Aging Cell*. 2017;16(1):183-191.
9. Hansen KD, Timp W, Bravo HC, et al. Increased methylation variation in epigenetic domains across cancer types. *Nat Genet*. 2011;43(8):768-775.
10. Itzykson R, Fenaux P. Epigenetics of myelodysplastic syndromes. *Leukemia*. 2014;28(3):497-506.
11. Teschendorff AE, Liu X, Caren H, et al. The dynamics of DNA methylation covariation patterns in carcinogenesis. *PLoS Comput Biol*. 2014;10(7):e1003709.
12. Li S, Garrett-Bakelman F, Perl AE, et al. Dynamic evolution of clonal epialleles revealed by methclone. *Genome Biol*. 2014;15(9):472.