Epigenetic and functional changes imposed by NUP98-HOXA9 in a genetically engineered model of chronic myeloid leukemia progression

NUP98-HOXA9 ("NA9") is a fusion gene commonly found, together with other mutations, in human acute myeloid leukemia cells, and occasionally in blast-phase (BP) cells of patients with chronic myeloid leukemia (CML) in whom disease progression has occurred.[1] These findings underscore interest in elucidating the role of NA9 in contributing to an aggressive leukemic program. Previous studies addressing this issue reported biological and molecular consequences of NA9 alone in human cells, including normal human hematopoietic cells,[2,3] and evidence of a co-operative role of NA9 with BCR-ABL1 in experimentally-induced models of leukemia in mouse cells.[4-6] We now report the results of assessing the biological and molecular consequences of lentivirally-introduced expression of NA9 in freshly isolated CD34$^+$ CML cells from patients with chronic phase (CP) disease (for a general outline and technical details, see *Online Supplementary Figure S1A-E*). Samples from patients with CP CML are known to contain variable ratios of normal and *BCR-ABL1*$^+$/Philadelphia chromosome-positive cells, particularly in the most primitive cell compartments.[7] Therefore, for this study, we selected samples from three CP-CML patients for whom we had previously shown the content of any type of normal CD34$^+$ cells to be very low based on assessments of their progeny genotypes produced for >30 weeks (*Online Supplementary Table S1*).[8]

Initial *in vitro* experiments with these three CML patients' CP cells showed that *NA9*-transduced cells (compared to control-transduced cells) produced particularly large expansions of granulopoietic colony-forming cells in growth factor-supplemented liquid suspension cultures, with somewhat reduced effects on expanding the outputs of these progenitors in 6-week long-term cultures on human growth factor-producing stroma, or in serial replating assays of the transduced cells cultured in semi-solid medium (*Online Supplementary Figure S1F-H*). However, in all cases, the progeny of the *NA9*-transduced CML cells ultimately underwent terminal differentiation (*Online Supplementary Figure S1I*), thus recapitulating results reported previously for *NA9*-transduced normal human CD34$^+$ hematopoietic cells.[2,9]

Outputs of *NA9*- and control-transduced CD34$^+$ cells from the same three CP-CML patients' samples were also assessed in sublethally irradiated NOD-*Prkdc*^{scid}*IL2Rγc*$^{-/-}$ mice engineered to produce human interleukin 3, granu-

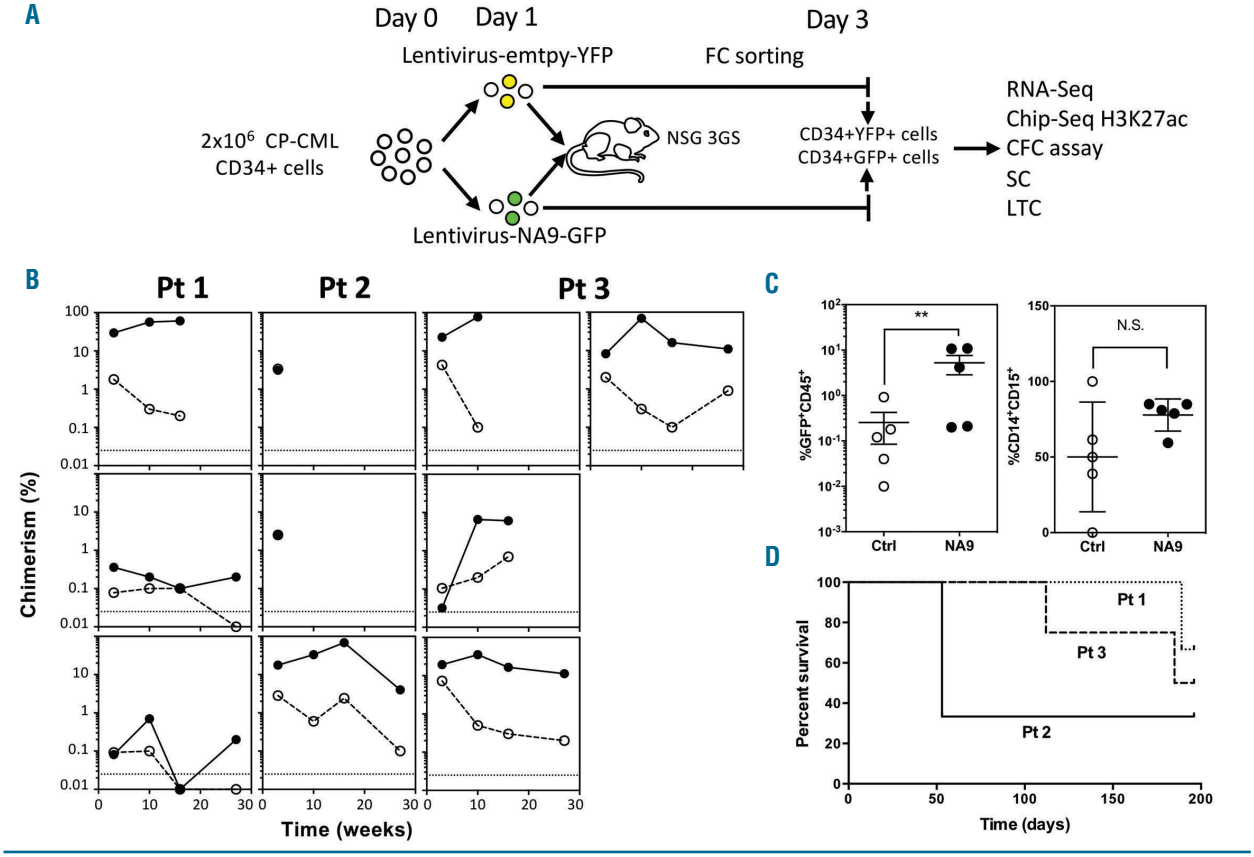

Figure 1. NA9 promotes disease progression in xenografts of transduced CD34$^+$ chronic phase chronic myeloid leukemia cells. (A) Experimental design: 8-week old NSG-3GS mice were sublethally irradiated with 3.15 Gy of ^{137}Cs γ-rays. 2 x 10^6 cells of a mixture containing equal proportions of chronic phase (CP) chronic myeloid leukemia (CML) CD34$^+$ cells exposed to control and NA9 virus were injected intravenously into ten mice (3 with cells from patient #1, 3 from patient #2, and 4 from patient #3). (B) Kinetics of % NA9$^+$ (GFP$^+$) cells and control (YFP$^+$) cells in the blood (at week 3) and bone marrow (all other time points) of NSG-3GS mice transplanted with transduced CP CML CD34$^+$ cells (left, patient #1; middle, patient #2; right patient #3. (C) Differences in bone marrow chimerism (left; paired *t*-test, *P*=0.002) and the proportion of mature granulocyte-macrophage (CD14$^+$CD15$^+$) cells (not significant, paired *t*-test, *P*=0.18) between control (YFP$^+$) and NA9$^+$ (GFP$^+$) human cells 28 weeks after transplantation. (D) Survival curves for recipients of different patients' samples (n=10 mice). (E) Representative flow cytometric analysis of two mouse bone marrow samples assessed 16 weeks after transplantation. The results indicate a predominant population derived from NA9-transduced (GFP$^+$) cells (left histograms) of either the granulocyte (upper panels) or monocyte lineages (lower panel). YFP: yellow fluorescent protein, GFP: green fluorescent protein; FC: flow cytometry; RNA-Seq: RNA-sequencing; Chip-Seq: chromatin immunoprecipitation sequencing; CFC: colony-forming cells; SC: suspension culture; LTC: long-term culture; Pt: patient.

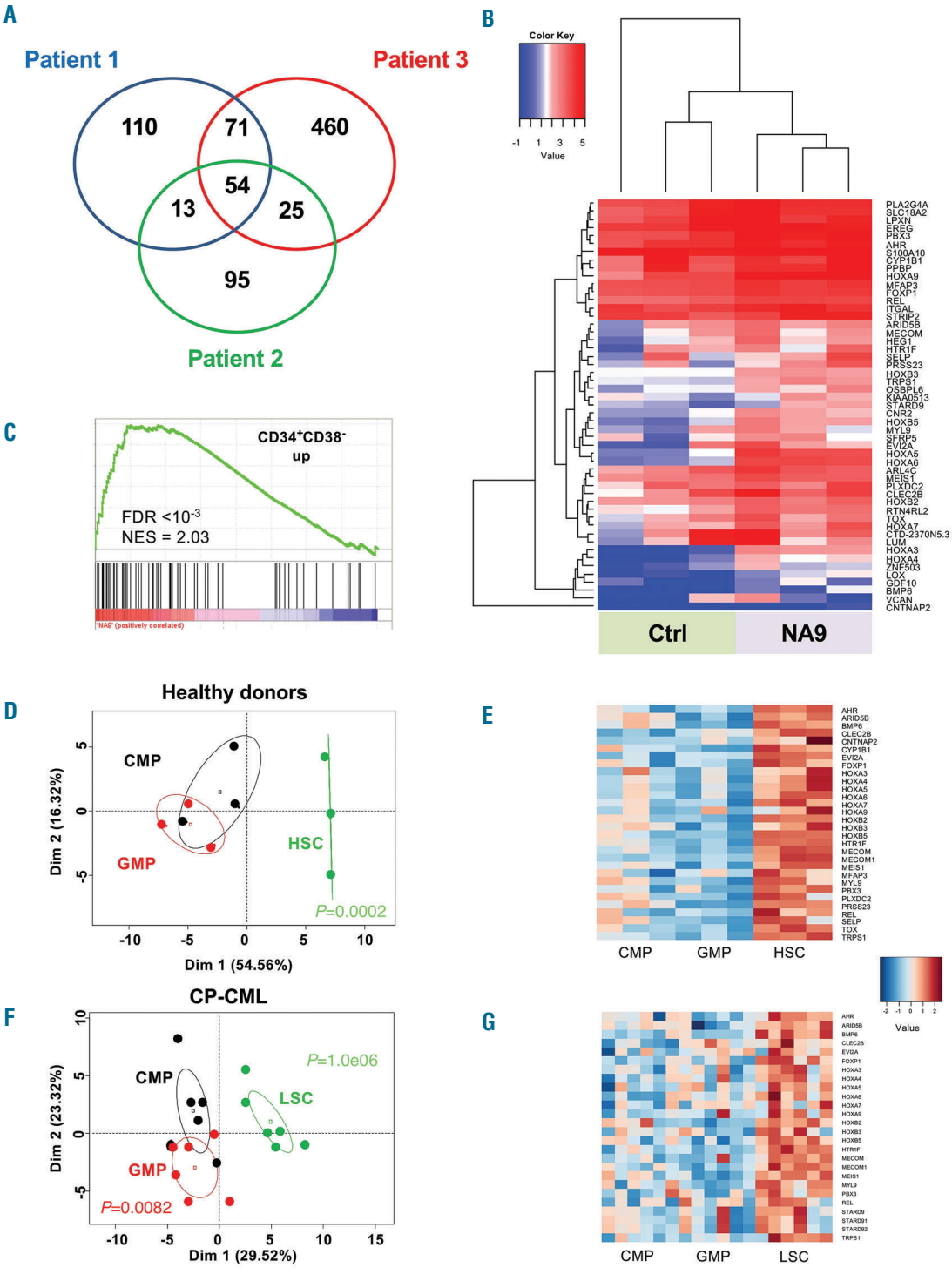

Figure 2. RNA-sequencing analysis of NA9$^+$ chronic myeloid leukemia CD34$^+$ cells. Libraries were prepared from 10 ng aliquots of total RNA extracted from FACS-sorted GFP$^+$ (NA9$^+$) and YFP$^+$ (control-transduced) CD34$^+$ cells cultured for 2 days after transduction in serum-free medium plus five growth factors. Differential expression analysis was performed with DEfine v.1 (unpublished), with a false discovery rate of 0.05 and a minimum reads per kilobase per million mapped reads of 0.01. (A) Venn diagram showing the distribution of upregulated genes in the three tested *NA9*-transduced samples from patients (Venny 2.1.0). (B) Heatmap of genes upregulated in all *NA9*-transduced CD34$^+$ chronic myeloid leukemia cell samples compared to control-transduced cells. An unsupervised clustering analysis using Euclidean distance was performed on $\log_2$-transformed expression data for genes referenced in HUGO with increased transcript levels in all three samples (using Gplots package in R version 3.2.3). (C) Gene set enrichment analysis of whole transcriptomes showing significant enrichment of a previously reported set of transcripts found to be present at increased levels in lin$^-$CD34$^+$CD38$^-$ cells compared with lin$^+$CD34$^+$CD38$^+$ cells from normal cord blood, adult bone marrow and mobilized peripheral blood cells.[12] (D) Principal component analysis (PCA) of the *NA9* gene signature identified in Figure 2B examined in normal lin$^-$CD34$^+$CD38$^-$CD45RA$^-$CD90$^+$ cells (HSC), lin$^-$CD34$^+$CD38$^+$CD123$^+$CD45RA$^-$(CMP) and lin$^-$CD34$^+$CD38$^+$CD123$^+$CD45RA$^+$ (GMP) (GSE47927.). (E) PCA x-axis coordinates of transcripts present at levels that distinguish HSC from GMP and CMP from healthy donors. (F) PCA of the NA9 gene signature in lin$^-$CD34$^+$CD38$^-$CD45RA$^-$CD90$^+$ leukemic stem cells (LSC), and CMP and GMP from CML patients (GSE47927). (G) PCA x-axis coordinates of transcript levels separating LSC from GMP and CMP isolated from CML patients. NES: normalized enrichment score; FDR: false discovery rate; Ctrl: control; CMP: common myeloid progenitors; GFP: green fluorescent protein; GMP: granulocyte-monocyte progenitors; HSC: hematopoietic stem cells; NA9: NUP98-HOXA9; YFP: yellow fluorescent protein.

locyte-macrophage colony-stimulating factor (CM-CSF) and stem cell factor constitutively (NSG-3GS mice). In these experiments, equal amounts of the *NA9*- and control-transduced cells were combined immediately after transduction and injected together into ten of these mice. Follow-up of these mice for 7 months showed that all the human cells subsequently detected in them were exclusively myeloid ($CD33^+$) (*Online Supplementary Figure S3*), as is typical of CP-CML xenografts.[10] In nine of the ten mice, bone marrow aspirates showed competitive overgrowth by the *NA9*-transduced (GFP^+) cells already evident within 3 weeks after transplantation and increasing thereafter. In five mice, this led to death or symptoms requiring euthanasia (Figure 1A-D, *Online Supplementary Table S2*).

RNA-sequencing analysis of extracts of *NA9*$^+$ (GFP^+)- and control (YFP^+)-transduced cells obtained 2 days after exposure to virus revealed transcripts of 54 coding genes to be already upregulated in all three $NA9^+$ $CD34^+$ cell samples (Figure 2A, *Online Supplementary Datasets S1-5*). These upregulated transcripts allowed the $NA9^+$ CML cells to be separated from the control-transduced cells in an unsupervised clustering analysis (Figure 2B). Functional annotation of this NA9 signature showed an overrepresentation of transcription factors (TF) with DNA binding and RNA pol II distal enhancer activity. Many of these are the same as those found to be rapidly deregulated in normal $CD34^+$ cells transduced with a *NA9*-encoding retroviral vector,[2,9] consistent with the similarity in the early *in vitro* biological responses exhibited by *NA9*-transduced CP CML and normal $CD34^+$ cells (*Online Supplementary Figure S3A-C*).[2,9] However, comparison of our transcriptome data for *NA9*-transduced CP CML cells with those published for *NA9*-transduced $CD34^+$ cells isolated from normal mobilized blood samples also indicated major differences, with 60% of the transcripts found to be upregulated in the CML NA9 signature and not similarly affected in $NA9^+$ normal cells (*Online Supplementary Dataset S6*).

We also performed an unsupervised clustering analysis of the GSE4170 microarray dataset for cells from 42 $CD34^+$ CP-CML and 28 BP-CML patients using the 54 transcripts that constituted the NA9-CP CML signature identified here. This analysis showed that many of the transcripts in the NA9 signature were among the transcripts with the most highly upregulated expression in the leukemic $CD34^+$ cells from BP patients compared with those from patients with accelerated phase or CP disease (including *ARID5B, STARD9, AHR, TOX* and *FOXP1*, Fisher one-way analysis of variance; $P<5x10^{-9}$) (*Online Supplementary Figures S4 and S5*). These results suggest that increased expression of NA9 targets may be a frequent hallmark, and hence an important element, in the creation of a BP disease program. It was therefore of interest that the transcripts comprising the NA9 signature were also deregulated in the transcriptomes of intermediate and high-risk acute myeloid leukemia patients carrying a variety of driver mutations (*Online Supplementary Figure S6* and *S7*). Notably, the NA9 signature was also predictive of a significantly reduced overall survival in The Cancer Genome Atlas acute myeloid leukemia cohort[11] (*Online Supplementary Figure S8*).

Gene set enrichment analysis (GSEA) of the NA9-upregulated transcripts in the transduced CML samples also showed a significant enrichment in transcripts found at higher levels in the $CD34^+CD38^-$ fraction of normal bone marrow, cord blood or mobilized blood cells in comparison to the more predominant, and more differentiated, $CD34^+CD38^+$ cells present in the same normal samples[12] (Figure 2C). Specific examples include genes in the *HOXA* and *HOXB* clusters and their cofactors, *MEIS-1* and *PBX3, MECOM, AHR, BMP6* and *GDF10*. GSEA also revealed a high enrichment in the *NA9*-transduced CML cells of previously reported upregulated NA9 targets, particularly TF transcripts. These findings also extended to differential transcriptome data derived from a comparison of $CD34^+CD38^-CD90^+$ cells CP-CML cells and their more differentiated progeny as well as their normal counterparts (Figure 2D-G) (GSE47927). Taken together, these results all point to a rapid and selective upregulation of a "stem cell" program in *NA9*-transduced $CD34^+$ CP-CML cells.

We also compared the frequency and sites of H3K27ac marks in the same paired sets of *NA9*- and control-transduced cells. *NA9*-specific *de novo* H3K27ac marked regions were found near (±2 kb) the transcriptional start sites, or within intragenic regions of *HOXA9, HOXA7, MEIS1* and *PBX3*, but not in *HOXC* or *HOXD* clusters, or in *PBX1* (Figure 3A, *Online Supplementary Dataset S7*). These NA9-specific *de novo* H3K27ac regions overlap with previously identified NA9 binding sites in the HEK293FT cell line (GSE62587).[3] Additional NA9-specific *de novo* H3K27ac marked regions that were identified are known to drive expression of genes previously reported as upregulated in normal $CD34^+$ cells expressing a *NA9* transgene, in the GM-CSF signaling pathway, and in myeloid cells exposed to lipopolysaccharide (Figure 3 B and C, *Online Supplementary Dataset S8*). *De novo* H3K27ac marks in $NA9^+$ CML $CD34^+$ cells also showed a significant enrichment of TF binding sites in *GATA1, RELA, MEF2A* and *IKZF1*, previously mapped by chromatin immunoprecipitation sequencing experiments (adjusted *P*-value <0.05) (Figure 3D). *REL* is of interest as it belongs to the set of genes (i.e., *GATA1, RELA, MEF2A* and *IKZF1*) whose expression was altered by forced NA9 expression in the CML $CD34^+$ cells with concomitant acquisition of H3K27ac marks that overlap with previously identified NA9 binding sites (GSE62587) (*Online Supplementary Figure S9*).[3]

We also identified ten super enhancers consistently present in the *NA9*-transduced $CD34^+$ CP CML cells and absent in the control cells (Figure 3A, E and F). These super enhancers are predicted to regulate ten genes, including *PBX3, ANGPT1* and *MBNL1* with the *PBX3* super enhancer again overlapping extensively with NA9 binding sites previously identified (GSE62587).[3] The *MBNL1* super enhancers are part of a previously identified short list of common binding sites of NA9, NUP98-HOXD13 and MLL-AF9 fusions and *ANGPT1* is one of the rare genes with NA9 binding sites in its enhancers and increased expression in NA9-expressing cells.[13] Another notable finding was an increase in H3K27ac marks forming an intronic super enhancer 193 kb from the *FOXP1* transcriptional start site enriched in NA9-binding sites and an associated upregulated expression of *FOXP1* (*Online Supplementary Figure 9D*). FOXP1 has recently been found to enhance normal and leukemic stem/progenitor cell expansion by downregulating expression of $p21^{-CIP1}$ and $p27^{-KIP1}$, two cell-cycle inhibitors.[14]

Taken together, these findings suggest that in spite of the somewhat variable biological changes induced by NA9 expression in the total $CD34^+$ CP CML populations from different patients, the transcriptional program is coherently altered to activate key elements of an enhanced self-renewal program. Given previous evidence that the blasts in BP CML have features of self-renewing

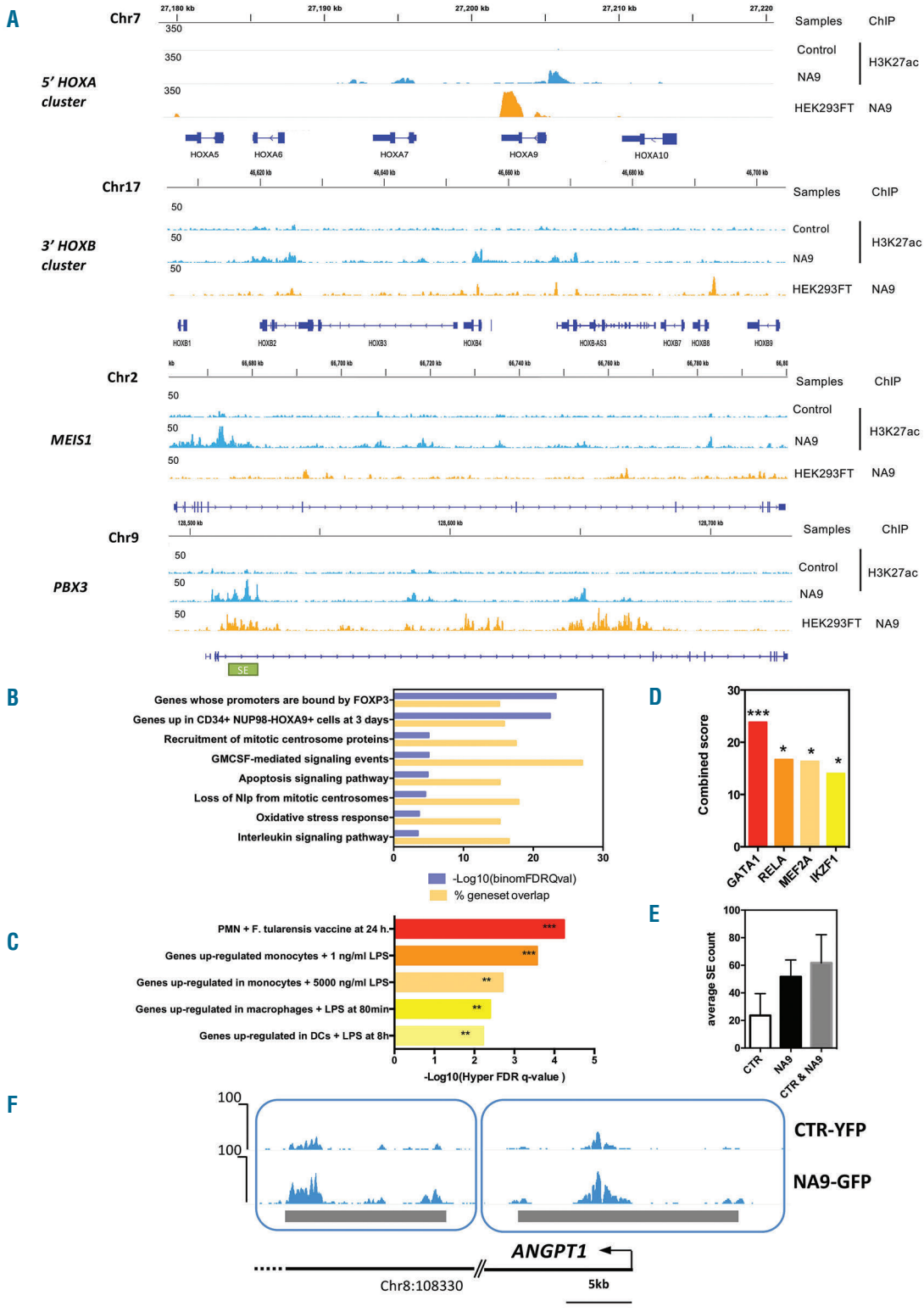

Figure 3. NA9 induces a rapid modification of H3K27ac marks in multiple enhancers. (A) H3K27ac peaks on indicated hg19 loci in control- (top) and *NA9* (middle)-transduced CD34$^+$ cells, in blue, as compared with NA9 peaks identified in the HEK293FT cell line, in orange (GSE62587).[3] A green bar indicates an 11.7 kb super enhancer found 10.7 kb downstream of the *PBX3* transcriptional start site (*Online Supplementary Dataset S9*). (B) Most significant gene sets from the MySigDB perturbation, MySigDB pathway and Panther pathway ranked by a binomial false discovery rate q-value (blue). Percentage overlap between the gene sets and the genes with proximal H3K27ac marks identified in NA9$^+$ (GFP$^+$) chronic phase (CP) chronic myeloid leukemia (CML) CD34$^+$ cells and absent in control (YFP$^+$) CP CML CD34$^+$ cells are also indicated (orange). (C) H3K27ac peaks present in NA9$^+$ (GFP$^+$) CD34$^+$ cells and absent in control (YFP$^+$) cells overlap with enhancers of genes upregulated in monocytes, macrophages, neutrophils and dendritic cells following exposure to lipopolysaccharides. (D) Transcription factors showing enriched binding sites in gene enhancers with proximal H3K27ac marks identified in NA9$^+$ (GFP$^+$) CP CML CD34$^+$ cells and absent in control (YFP+) CP CD34$^+$ cells. (E) Average number of super enhancers found in control (YFP$^+$) CP CML CD34$^+$ cells only (CTR), or NA9$^+$ (GFP$^+$) CP CML CD34$^+$ cells only (NA9), or in both NA9$^+$ and control cells (CTR & NA9). Differences were not statistically significant (Wilcoxon matched-pairs signed ranked test). (F) Location of two super enhnacers (gray bars) found within the *ANGPT1* gene in NA9$^+$ (GFP$^+$) CP CML CD34$^+$ cells and not in control (YFP$^+$) CP CML CD34$^+$ cells.

granulocyte-macrophage colony-forming units,[15] our results suggest that this also reflects the establishment of a program that can be induced by NA9 expression. The significant enrichment of H3K27ac marks in enhancers of genes of the GM-CSF signaling and inflammatory pathways adds further support to this concept. In addition, we note that NA9 appears to induce many features of progression to BP disease but without accumulation of blasts, thus distinguishing it from results obtained in mouse models.[4-6] Nevertheless, our findings strongly point to the NA9 signature identified here as one that is common to the BP populations that emerge spontaneously in CML patients whose CP disease becomes uncontrolled.

Ivan Sloma,[1-3°] Philip Beer,[1] Christophe Desterke,[4] Elizabeth Bulaeva,[1] Misha Bilenky,[5] Annaïck Carles,[6] Michelle Moksa,[6] Kamini Raghuram,[1] Cedric Brimacombe,[1] Karen Lambie,[1] Ali G. Turhan,[7] Orianne Wagner-Ballon,[2,3] Philippe Gaulard,[2,8] Keith Humphries,[1] Martin Hirst[5,6] and Connie J. Eaves[1,9]

[1]Terry Fox Laboratory, British Columbia Cancer Agency, Vancouver, British Columbia, Canada; [2]Université Paris Est Creteil, INSERM, IMRB, Creteil, France; [3]Assistance Publique des Hôpitaux de Paris (AP-HP), Hopital Henri Mondor, Departement d'Hematologie et Immunologie, Creteil, France; [4]Université Paris-Sud, Faculté de Médecine Kremlin Bicêtre, INSERM UMS 33 Villejuif, France; [5]Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia, Canada; [6]Department of Microbiology and Immunology, Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada; [7]Service d'Hématologie Biologique, Hôpital Universitaire Paris-Saclay, INSERM U935, AP-HP, Kremlin Bicêtre, France; [8]INSERM UMRS955, Pathology Department, Hôpital Henri Mondor, AP-HP, Faculté de Médecine de Créteil, Université Paris-Est Créteil, Créteil, France and [9]University of British Columbia, Department of Medical Genetics, Vancouver, British Columbia, Canada

°Present address: Université Paris Est Creteil, INSERM, IMRB and AP-HP, Hopital Henri Mondor, Departement d'Hematologie et Immunologie, Creteil, France

Correspondence: IVAN SLOMA - ivan.sloma@aphp.fr

doi:10.3324/haematol.2020.249243

Disclosures: no conflicts of interest to disclose.

Contributions: IS and CJE designed the study. IS prepared the vectors with assistance from KR and CB. IS isolated the CD34$^+$ cells, and with EB performed the transductions and with KR, CB and PB carried out the in vitro experiments. IS and PB performed the in vivo experiments and, with assistance from PG and OWB, the necropsy analyses. MH supervised the RNA-sequencing (RNA-seq) and chromatin immunoprecipitation sequencing (ChIP-seq) studies. MM prepared the RNA-seq libraries. MB analyzed the primary RNA-seq data and differential expression. CD performed variant calling, in silico transcriptomics analysis and integration of RNA-seq and ChiP-seq data. IS and CD contributed to the tertiary RNA-seq analysis. MM performed ChiP-seq experiments. AC did the primary and secondary analyses of ChiP-seq experiments while IS contributed to secondary and tertiary analyses. IS and CJE wrote the manuscript with input from PB, CD, MB and MM and approval from all authors.

Funding: IS received a grant from Cent pour Sang La Vie and AGT received a grant from the Oncostem program of the French national IPSC Infrastructure "Investissements d'Avenir" INGESTEM. CE received grants supported by funds from the Terry Fox Run and from CIHR. MH received grants supported by funds from the Terry Fox Run and the Terry Fox Research Institute (TFF-122869).

References

1. Takeda A, Yaseen NR. Nucleoporins and nucleocytoplasmic transport in hematologic malignancies. Semin Cancer Biol. 2014;27:3-10.
2. Takeda A, Goolsby C, Yaseen NR. NUP98-HOXA9 induces long-term proliferation and blocks differentiation of primary human CD34+ hematopoietic cells. Cancer Res. 2006;66(13):6628-6637.
3. Rio-Machin A, Gómez-López G, Muñoz J, et al. The molecular pathogenesis of the NUP98-HOXA9 fusion protein in acute myeloid leukemia. Leukemia. 2017;31(9):2000-2005.
4. Mayotte N, Roy DC, Yao J, Kroon E, Sauvageau G. Oncogenic interaction between BCR-ABL and NUP98-HOXA9 demonstrated by the use of an in vitro purging culture system. Blood. 2002;100(12):4177-4184.
5. Ito T, Kwon HY, Zimdahl B, et al. Regulation of myeloid leukaemia by the cell-fate determinant Musashi. Nature. 2010;466(7307):765-768.
6. Ashton JM, Balys M, Neering SJ, et al. Gene sets identified with oncogene cooperativity analysis regulate in vivo growth and survival of leukemia stem cells. Cell Stem Cell. 2012;11(3):359-372.
7. Sloma I, Jiang X, Eaves AC, Eaves CJ. Insights into the stem cells of chronic myeloid leukemia. Leukemia. 2010;24(11):1823-1833.
8. Sloma I, Imren S, Beer PA, et al. Ex vivo expansion of normal and chronic myeloid leukemic stem cells without functional alteration using a NUP98HOXA10homeodomain fusion gene. Leukemia. 2013;27(1):159-169.
9. Chung KY, Morrone G, Schuringa JJ, et al. Enforced expression of NUP98-HOXA9 in human CD34(+) cells enhances stem cell proliferation. Cancer Res. 2006;66(24):11781-11791.
10. Sloma I, Beer PA, Saw KM, et al. Genotypic and functional diversity of phenotypically defined primitive hematopoietic cells in patients with chronic myeloid leukemia. Exp Hematol. 2013;41(10):837-847.
11. Network CGAR. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059-2074.
12. Georgantas RW, Tanadve V, Malehorn M, et al. Microarray and serial analysis of gene expression analyses identify known and novel transcripts overexpressed in hematopoietic stem cells. Cancer Res. 2004;64(13):4434-4441.
13. Xu H, Valerio DG, Eisold ME, et al. NUP98 fusion proteins interact with the NSL and MLL1 complexes to drive leukemogenesis. Cancer Cell. 2016;30(6):863-878.
14. Naudin C, Hattabi A, Michelet F, et al. PUMILIO/FOXP1 signaling drives expansion of hematopoietic stem/progenitor and leukemia cells. Blood. 2017;129(18):2493-2506.
15. Jamieson CH, Ailles LE, Dylla SJ, et al. Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML. N Engl J Med. 2004;351(7):657-667.