Bioinformatic analysis of the effects and mechanisms of decitabine and cytarabine on acute myeloid leukemia

SHIYONG ZHOU*, PENGFEI LIU* and HUILAI ZHANG

Department of Lymphoma, Sino-US Center of Lymphoma and Leukemia, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin’s Clinical Research Center for Cancer, Tianjin 300060, P.R. China

Received April 12, 2016; Accepted March 14, 2017

DOI: 10.3892/mmr.2017.6581

Abstract. Acute myeloid leukemia (AML) is a frequently occurring malignant disease of the blood and may result from a variety of genetic disorders. The present study aimed to identify the underlying mechanisms associated with the therapeutic effects of decitabine and cytarabine on AML, using microarray analysis. The microarray datasets GSE40442 and GSE40870 were downloaded from the Gene Expression Omnibus database. Differentially expressed genes (DEGs) and differentially methylated sites were identified in AML cells treated with decitabine compared with those treated with cytarabine via the Linear Models for Microarray Data package, following data pre-processing. Gene Ontology (GO) analysis of DEGs was performed using the Database for Annotation, Visualization and Integrated Analysis Discovery. Genes corresponding to the differentially methylated sites were obtained using the annotation package of the methylation microarray platform. The overlapping genes were identified, which exhibited the opposite variation trend between gene expression and DNA methylation. Important transcription factor (TF)-gene pairs were screened out, and a regulated network subsequently constructed. A total of 190 DEGs and 540 differentially methylated sites were identified in AML cells treated with decitabine compared with those treated with cytarabine. A total of 36 GO terms of DEGs were enriched, including nucleosomes, protein-DNA complexes and the nucleosome assembly. The 540 differentially methylated sites were located on 240 genes, including the acid-repeat containing protein (ACRC) gene that was additionally differentially expressed. In addition, 60 TF pairs and overlapped methylated sites, and 140 TF-pairs and DEGs were screened out. The regulated network included 68 nodes and 140 TF-gene pairs. The present study identified various genes including ACRC and proliferating cell nuclear antigen, in addition to various TFs, including TATA-box binding protein associated factor 1 and CCCTC-binding factor, which may be potential therapeutic targets of AML.

Introduction

Acute myeloid leukemia (AML) is one of the most frequently occurring malignant diseases of the blood and patients of all ages may present with symptoms. It has previously been reported that AML in children accounts for 25% of pediatric leukemia cases and affects ~180 patients annually in Japan (1). A total of 19,000 cases of AML are diagnosed each year, with ~10,000 of these in the United States (2). Outcomes have improved in younger patients, with a 40-50% 5-year overall survival rate (3). However, the majority of AML cases occur in adults, and in these cases the mortality rate remains high. It has been demonstrated that only 10-20% of patients aged >60 years survive to 5 years; 80% of patients are incurable as a result of primary refractoriness, relapse or treatment-associated mortality (4,5). AML has several subtypes and treatment and prognosis varies among them. AML is treated traditionally with chemotherapy and recent genetic research has provided more personalized treatment options. Clinicians can now predict which drug or drugs may work best for a particular person, and how long that person is likely to survive. Furthermore, numerous studies have reported that various genetic abnormalities in the following genes: Nucleophosmin 1, runt related transcription factor 1, Tet methylcytosine dioxygenase 2 and isocitrate dehydrogenase [NADP(+)] 1 cytosolic, are associated with the occurrence, progression and recurrence of AML and may therefore be used to predict prognosis and guide future therapeutic research (6-9). Döhner et al (4) summarized the frequency and clinical significance of various important mutated genes. The primary first line therapeutic for the treatment of AML is combined chemotherapy with anthracycline
and cytarabine (10). Further therapeutic options include the hypomethylating agents decitabine and azacitidine, low-dose cytarabine, investigational agents, and supportive care with hydroxyurea and transfusions (11). Decitabine is a deoxynucleoside analogue of cytidine that selectively inhibits DNA methyltransferases. It is considered an effective and well-tolerated alternative treatment to cytarabine or supportive care in older patients with AML (12). To improve the efficacy and structure of decitabine, the present study examined the mechanism underlying the effects of decitabine and cytarabine on AML, via microarray analysis. Although some progress has been made in targeted therapy of AML, the diagnosis and treatment of it remain challenging. The present study identified additional biomarkers associated with the therapeutic effects of drugs, in order to explore the corresponding mechanisms.

Materials and methods

Microarray data. The microarray datasets GSE40442 (13) and GSE40870 (13) were downloaded from the Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo/). The expression profile of GSE40442 contained 67 primary AML samples cultured with medium only for 1 day, followed by 3 days treatment with decitabine (the case group; n=17), cytarabine (the control group; n=16), dimethyl sulfoxide (DMSO; n=17) or untreated (n=17). These data were identified via the GPL5188 [HuEx-1_0-st] Affymetrix Human Exon 1.0 ST Array [probe set (exon) version] platform. The GSE40870 profile presented the methylation data of AML cell samples treated with decitabine (the case group; n=16), cytarabine (the control group; n=16) or DMSO (n=16). Detection of the methylation data was performed via GPL13534 Illumina HumanMethylation450 BeadChip (HumanMethylation 450_15017482).

Data preprocessing. To create the expression profile, the original data were converted into a recognizable format in R, and the affy (14) package (bioconductor.org/packages/release/bioc/html/affy.html) was used for background correction and normalization, followed by conversion from the probe symbol to the gene symbol with the biomaRt (15) package of R (bioconductor.org/packages/release/bioc/html/biomaRt.html). The β-value of every methylated site in all samples was extracted via GenomeStudio software version 2.0 (Illumina, Inc., San Diego, CA, USA) to create a methylation profile.

Identification of differentially expressed genes (DEGs) and differentially methylated sites. For GSE40442, the Linear Models for Microarray Data (16) package of R (bioconductor.org/packages/release/bioc/html/limma.html) was used to identify the DEGs in AML cells treated with decitabine compared with those treated with cytarabine. The DEGs were identified according to the criteria P<0.05 and log(fold-change)>0.5. The heatmap of DEGs in every sample of the control and the case group was constructed. For GSE40870, the differentially methylated sites were identified in AML cells treated with decitabine compared with cytarabine via the Illumina Methylation Analyzer (17) package of R (ima.r-forge.r-project.org/), and were screened out with the criteria P<0.05 and log(fold-change)>0.2.

Functional enrichment analysis. Gene Ontology (GO) enrichment analysis of DEGs was performed via the Database for Annotation, Visualization and Integrated Discovery (david.abcc.ncifcrf.gov/) (18) with the threshold of P<0.05.

Screening of important genes and methylated sites. Genes corresponding to the differentially methylated sites were obtained by the annotation package of the methylation microarray platform. The genes that exhibited an overlap compared with DEGs were selected, and those that exhibited the opposite trend in the methylation variation compared with their expression were screened out.

Identification and analysis of important transcription factor (TF)-gene pairs and establishment of TF-gene regulated network. Methylation in the gene promoter region may affect the binding of TFs to genes and result in the variation of gene expression. Firstly, chromosomal locations of the methylation sites were identified using the annotation package of the methylation microarray platform. Following this, chromosomal locations of all the known and predicted TF binding sites were downloaded from the University of California Santa Cruz (UCSC) database (19) (genome.ucsc.edu/). The methylation sites were considered to affect the binding of TFs and genes when the chromosomal location of the methylation sites overlapped with the region of the TF binding site. Furthermore, all the known and predicted TF-gene pairs were downloaded from UCSC and the TF-gene pairs were screened out. The TF-gene regulated network was established via Cytoscape version 3.11 (www.cytoscape.org/).

Results

DEGs and differentially methylated sites. A total of 190 DEGs (102 up- and 88 downregulated) and 540 differentially methylated sites were identified in AML cells treated with decitabine compared with cytarabine, and all the identified differentially methylated sites were hypomethylated. The top 30 DEGs and the top 30 differentially methylated sites are presented in Tables I and II, respectively, and the heatmap of DEGs is presented in Fig. 1.

Enriched GO terms of the DEGs. A total of 36 enriched GO terms of DEGs, including nucleosome, protein-DNA complex, and nucleosome, chromatin and protein-DNA complex assemblies, were obtained and are presented in Table III.

Important genes and methylated sites. A total of 240 genes were screened, in which 540 differentially methylated sites were identified. These 240 genes were compared with the 190 DEGs, and the acid-repeat containing protein (ACRC) exhibited an overlap. Furthermore, ACRC corresponded to the methylated site of cg26924445 and demonstrated the opposite trend in the methylation variation compared with gene expression.
Important TF-gene pairs and the TF-gene regulated network. A total of 60 TF-gene pairs and overlapped methylated sites were screened out, including cg22475974-SET domain bifurcated (SETDB)1, cg14063817-estrogen receptor (ER)α, cg22475974-ERα, cg05835309-SETDB1 and cg00334293-signal transducer and activator of transcription-3. In addition, 140 pairs of TFs and DEGs were identified, including CCAAT/enhancer binding protein β (CEBPB)-cysteine rich secretory protein 3, CEBPB-C-X-C motif chemokine ligand 2, CEBPB-fanconi anemia complementation group I, CEBPB-histone cluster 1 H1 family member C and CEBPB-microRNA 378e. In addition, the 60 pairs of TFs and overlapped methylated sites contained 20 TFs and 51 methylated sites. The TF-gene regulated network was established according to the 140 TF-gene pairs (Fig. 2). A total of 68 nodes and 140 pairs were involved in the regulated network. Furthermore, the 68 nodes contained 18 TFs (presented as triangles) and 50 genes (presented as circles). The top 20 nodes according to connections with other nodes in the network are presented in Table IV.

Discussion
Decitabine (Dacogen®; 5-aza-2'-deoxycytidine) has been extensively used for the treatment of AML as an inhibitor of DNA methylation, which triggers demethylation leading to consecutive reactivation of epigenetically silenced tumor suppressor genes (20). When administered at low doses, decitabine may reduce genomic DNA methylation as a consequence of irreversible binding to DNA methyltransferases following incorporation into newly synthesized DNA (21). Cytarabine inhibits DNA synthesis by suppressing DNA polymerase activity; however, it additionally inhibits the elongation of the polynucleotide chain and interferes with
the physiological function of DNA, which is important for the treatment of hematological malignancies (22,23). In the present study, all the identified differentially methylated sites were hypomethylated in the primary AML samples treated with low-dose decitabine compared with cytarabine, which is consistent with differing underlying mechanisms of decitabine and cytarabine in AML. The AML cell samples treated with decitabine differed from those treated with cytarabine in the heatmap of DEGs.

In the present study, a total of 36 GO terms enriched in DEGs were obtained. They were primarily associated with the combination of protein and DNA, (protein-DNA complex and protein-DNA complex assembly), chromosome conformation (chromatin assembly or disassembly, chromosomal part and chromatin organization), and biological processes associated with the assembly of macromolecular complexes (nucleosome assembly, protein-DNA complex assembly, macromolecular complex assembly and cellular macromolecular complex assembly). It has previously been demonstrated that DNA methylation is important in the biological processes of genomic imprinting, X-chromosome inactivation, suppression of transposable elements and carcinogenesis (24-28). DNA methylation is considered a potent epigenetic modification and may inhibit TF recruitment, resulting in suppression of transcription (24,29), and closely associates with health and disease in humans (30,31). Furthermore, it has been
previously reported that DNA methylation is an epigenetic activity that affects the structure of chromosomes, but not the sequence of genes (32-34). Therefore, the aforementioned data demonstrated that decitabine may affect AML via gene methylation.

Of the identified DEGs, ACRC was the only one to additionally contain a differentially methylated site, cg26924445, and demonstrated an opposite trend in methylation variation compared with expression. Nestheide et al (35) suggested that ACRC is an important biomarker of Ewing sarcoma and concludes that epigenetic dysregulation may contribute to the pathogenesis of angiosarcoma, via analysis of expression and methylation profiles. The results of the present study demonstrated that decitabine can alter the methylation status of cg26924445, and that as in their study, increasing expression of ACRC was conducive to treating AML. Therefore, it was suspected that decitabine might treat AML through altering the methylation status of cg26924445 in ACRC.

Table III. Enriched GO terms of differentially expressed genes.

| Category   | GO ID   | GO name                                           | P-value      |
|------------|---------|---------------------------------------------------|--------------|
| CC         | 0000786 | Nucleosome                                        | 1.03x10⁻⁷    |
| CC         | 0032993 | Protein-DNA complex                                | 6.69x10⁻⁷    |
| BP         | 0006334 | Nucleosome assembly                                | 8.70x10⁻⁷    |
| BP         | 0031497 | Chromatin assembly                                 | 1.07x10⁻⁶    |
| BP         | 0065004 | Protein-DNA complex assembly                       | 1.40x10⁻⁶    |
| BP         | 0034728 | Nucleosome organization                            | 1.59x10⁻⁶    |
| BP         | 0006323 | DNA packaging                                      | 6.06x10⁻⁶    |
| CC         | 0000785 | Chromatin                                         | 7.65x10⁻⁶    |
| BP         | 0006333 | Chromatin assembly or disassembly                  | 9.73x10⁻⁶    |
| CC         | 0005694 | Chromosome                                        | 3.77x10⁻⁵    |
| CC         | 0044427 | Chromosomal part                                   | 7.21x10⁻⁵    |
| BP         | 0016584 | Nucleosome positioning                             | 2.24x10⁻⁴    |
| BP         | 0065003 | Macromolecular complex assembly                    | 0.001008     |
| BP         | 0034622 | Cellular macromolecular complex assembly           | 0.001458     |
| CC         | 0031012 | Extracellular matrix                               | 0.001562     |
| BP         | 0043933 | Macromolecular complex subunit organization       | 0.001593     |
| BP         | 0034621 | Cellular macromolecular complex subunit organization | 0.002613   |
| BP         | 0006259 | DNA metabolic process                              | 0.003411     |
| BP         | 0006325 | Chromatin organization                             | 0.003467     |
| CC         | 0005654 | Nucleoplasm                                       | 0.004301     |
| CC         | 0005578 | Proteinaceous extracellular matrix                 | 0.006181     |
| BP         | 0006260 | DNA replication                                    | 0.006457     |
| CC         | 0044421 | Extracellular region part                          | 0.007453     |
| BP         | 0051276 | Chromosome organization                            | 0.011355     |
| MF         | 0003677 | DNA binding                                       | 0.012103     |
| BP         | 0006974 | Response to DNA damage stimulus                    | 0.015069     |
| BP         | 0030162 | Regulation of proteolysis                          | 0.018269     |
| BP         | 0033554 | Cellular response to stress                        | 0.022660     |
| BP         | 0006281 | DNA repair                                         | 0.024911     |
| MF         | 0005125 | Cytokine activity                                  | 0.024942     |
| BP         | 0032026 | Response to magnesium ion                          | 0.030924     |
| CC         | 0031981 | Nuclear lumen                                      | 0.035330     |
| BP         | 0046685 | Response to arsenic                                | 0.038507     |
| CC         | 0000307 | Cyclin-dependent protein kinase holoenzyme complex | 0.039727     |
| BP         | 0051726 | Regulation of cell cycle                           | 0.040353     |
| MF         | 0004984 | Olfactory receptor activity                        | 0.049749     |

GO, Gene Ontology; CC, cellular component; BP, biological process; MF, molecular function.
ZHOU et al: BIOINFORMATIC ANALYSIS OF DECITABINE AND CYTARABINE IN AML

reported that the SET-NUP214 (TAF1/CAN) fusion gene is an important influencing factor in the survival rate of AML. Therefore, TAF1 may be a potential novel target gene in decitabine-treated AML. The CCCTC-binding factor (CTCF) was another TF that regulated numerous genes or TFs. Manodoro et al (37) demonstrated that in AML, the methylation of CTCF binding sites may result in loss of imprinting at 14q32. Furthermore, the present study demonstrated that proliferating cell nuclear antigen (PCNA) was the gene regulated by the greatest number of TFs, and Buchi et al (38) reported that the expression of PCNA was altered in AML treated with decitabine or cytarabine.

In conclusion, the results of the present study suggested that decitabine suppresses the function of certain antitumor genes via methylation, in its role as a therapeutic agent for the treatment of AML, and that this underlying mechanism of action differs to that of cytarabine. The present study provides information regarding potential drug targets, which may improve the efficacy of decitabine in the treatment of AML.

Acknowledgements

The present study was supported by the Health Bureau Science and Technology Foundation of Tianjin (grant no. 2012KZ063) and the Municipal Science and Technology Commission of Tianjin (grant no. 15ZLZLF00440).

Figure 2. TF-gene regulated network. Triangles represent TF nodes; circles represent gene nodes. TF, transcription factor.

Table IV . Top 20 most significant nodes according to degree.

| Node     | Degree |
|----------|--------|
| TAF1     | 36     |
| CTCF     | 25     |
| C-MYC    | 14     |
| FOXA1    | 10     |
| RAD21    | 9      |
| CEBPB    | 8      |
| PCNA     | 8      |
| HEY1     | 6      |
| NRSF     | 6      |
| FANCl    | 6      |
| PU.1     | 5      |
| STAT3    | 5      |
| CXCL2    | 5      |
| THBS1    | 5      |
| HIST1H4H | 5      |
| NAMPT    | 5      |
| RPL36AL  | 5      |
| HIST1H1E | 5      |
| GABP     | 4      |
| CCNE2    | 4      |

Degree, connection with other nodes.
References

1. Taga T, Tomizawa D, Takahashi H and Adachi S: Acute myeloid leukemia in children: Current status and future directions. Pediart Int 58: 71-80, 2016.

2. Kadia TM, Ravandi F, Cortes J and Kantarjian H: New drugs in acute myeloid leukemia. Ann Oncol 27: 770-778, 2016.

3. Döhner H, Estey EH, Amadori S, Appelbaum FR, Büchner T, Burnett AK, Dombret H, Fenaux P, Grimwade D, Larson RA, et al: Diagnosis and management of acute myeloid leukemia in adults: Recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood 115: 453-474, 2010.

4. Döhner H, Weisdorf DJ and Bloomfield CD: Acute myeloid leukemia. N Engl J Med 373: 1136-1152, 2015.

5. Burnett A, Wetzler M and Löwenberg B: Therapeutic advances in acute myeloid leukemia. J Clin Oncol 29: 487-494, 2011.

6. Marcucci G, Haerlach T and Döhner H: Molecular genetics of adult acute myeloid leukemia: Prognostic and therapeutic implications. J Clin Oncol 24: 475-486, 2011.

7. Meyer SC and Levine RL: Translational implications of somatic genomics in acute myeloid leukemia. Lancet Oncol 15: e382-e394, 2014.

8. Cancer Genome Atlas Research Network; Ley TJ, Miller CA, Koboldt DC, Welch JS, Karchin R, et al: The catalog of somatic cell-type-specific mutations in 14,000 cancer samples. Nature 526: 386-390, 2015.

9. Guo Y, Bour S, Liu D, Ferrari G and Santini V: Loss of imprinting at the 14q32 domain is associated with partially expressed proteins in AML. Pediatr Blood Cancer 60: 1437-1446, 2013.

10. Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosoone CB, Johnson CS, Smiraglia DJ and Liu S: IMMA: An R package for high-throughput analysis of Illumina's 450k Infinium methylation data. Bioinformatics 28: 729-730, 2012.

11. Sherman BT, Huang da W, Tan Q, Guo Y, Bour S, Liu D, Stephens R, Baseler MW, Lane HC and Lempicki RA: DAVID Knowledgebase: A gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. BMC Bioinformatics 8: 426, 2007.

12. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM and Haussler D: The human genome browser at UCSC. Genome Res 12: 996-1002, 2002.

13. Daskalakis M, Blagitko-Dorfis N and Hackanson B: Decitabine. In: Small Molecules in Oncology. Springer, Berlin Heidelberg, pp131-157, 2010.

14. Flotho C, Claus R, Batz C, Schneider M, Sandrock I, Ihde S, Plass C, Niemeyer CM and Lübbert M: The DNA methyltransferase inhibitors azacitidine, decitabine and zebularine exert differential effects on cancer gene expression in acute myeloid leukemia cells. Leukemia 23: 1019-1028, 2009.

15. Löwenberg B: Sense and nonsense of high-dose cytarabine for acute myeloid leukemia. Blood 121: 26-28, 2013.

16. Momparler RL: Optimization of cytarabine (ARA-C) therapy for acute myeloid leukemia. Exp Hematol Oncol 2: 20, 2013.

17. Su S, Wan J, Su Y, Song Q, Zeng Y, Nguyen HN, Shin J, Cox E, Rho HS, Woodard C, et al: DNA methylation presents distinct binding sites for human transcription factors. Elife 2: e00726, 2013.

18. Egger G, Liang G, Aparicio A and Jones PA: Epigenetics in human disease and prospects for epigenetic therapy. Nature 429: 457-463, 2004.

19. Robertson KD: DNA methylation and human disease. Nat Rev Genet 6: 597-610, 2005.

20. Feinberg AP: Phenotypic plasticity and the epigenetics of human disease. Nature 447: 433-440, 2007.

21. Reik W: Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447: 425-432, 2007.

22. Weaver IC, Hellstrom IC, Brown SE, Andrews SD, Dymov S, Driorio J, Zhang FY, Szyf M and Meaney MJ: The methylated-DNA binding protein MB2 enhances NGFI-A (egr-1)-mediated transcriptional activation of the glucocorticoid receptor. Philos Trans R Soc Lond B Biol Sci 369: pii: 20130513, 2014.

23. Robertson KD and Wolfle AP: DNA methylation in health and disease. Nat Rev Genet 1: 11-19, 2000.

24. Hu Y, Liu F, Lin JY, Gao GF and Zhu B: Dissemination of the mcr-1 colistin resistance gene. Lancet Infect Dis 16: 313-317, 2016.

25. Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F and Bird A: Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell 69: 905-914, 1992.

26. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL and Schübel D: Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37: 853-862, 2005.

27. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M and Jacobsen SE: Shotgun bisulfite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452: 215-219, 2008.

28. Nestheide S, Bridge JA, Barnes M, Frayer R and Sumeji J: Pharmacologic inhibition of epigenetic modification reveals targets of aberrant promoter methylation in Ewing sarcoma. Pediatr Blood Cancer 60: 1437-1446, 2013.

29. Ben Abdelali R, Roggery A, Leguay T, Cieslak A, Rennaveille A, Touzart A, Banoos A, Randriamalala E, Caillot D, Lioure B, et al: SET-NUP214 is a recurrent y0 lineage-specific fusion transcript associated with corticosteroid/chemotherapy resistance in adult T-ALL. Blood: 123: 1860-1863, 2014.

30. Manodorof F, Marzoc J, Chaplin T, Mirakj-Moud F, Moravesk E, Jovanovic JV, Wang J, Iqbal S, Taussig D, Grimwade D, et al: Loss of imprinting at the 14q32 domain is associated with microRNA overexpression in acute promyelocytic leukemia. Blood 123: 2066-2074, 2014.

31. Buchi F, Spinelli E, Masala E, Gozzini A, Sanna A, Bosi A, Ferrari G and Santini V: Proteomic analysis identifies differentially expressed proteins in AML/ETO acute myeloid leukemia cells treated with DNMT inhibitors azacitidine and decitabine. Leuk Res 36: 607-618, 2012.