

ORIGINAL ARTICLE

# Cytidine deaminase polymorphisms and worse treatment response in normal karyotype AML

Lyoung Hyo Kim<sup>1,2</sup>, Hyun Sub Cheong<sup>2,3</sup>, Youngil Koh<sup>4</sup>, Kwang-Sung Ahn<sup>5</sup>, Chansu Lee<sup>3</sup>, Hyung-Lae Kim<sup>6</sup>, Hyoung Doo Shin<sup>1,2</sup> and Sung-Soo Yoon<sup>3,4,7</sup>

The *cytidine deaminase (CDA)* catalyzes the irreversible hydrolytic deamination of the cytarabine (AraC) into a 1- $\beta$ -D-arabinofuranosyluracil (AraU), an inactive metabolite that plays a crucial role in lowering the amount of AraC, a key chemotherapeutic drug, in the treatment of patients with acute myeloid leukemia (AML). In this study, we hypothesized that *CDA* polymorphisms were associated with the AraC metabolism for AML treatment and/or related clinical phenotypes. We analyzed 16 polymorphisms of *CDA* among 50 normal karyotype AML (NK-AML) patients, 45 abnormal karyotype AML (AK-AML) patients and 241 normal controls (NC). Several polymorphisms and haplotypes, rs532545, rs2072671, rs471760, rs4655226, rs818194 and CDA-ht3, were found to have a strong correlation with NK-AML compared with NC and these polymorphisms also revealed strong linkage disequilibrium with each other. Among them, rs2072671 (79A>C), which is located in a coding region and the resultant amino acid change K27Q, showed significant associations with NK-AML compared with NC ( $P=0.009$  and odds ratio = 2.44 in the dominant model). The AC and CC genotypes of rs2072671 (79A>C) were significantly correlated with shorter overall survival rates ( $P=0.03$ , hazard ratio = 1.84) and first complete remission duration ( $P=0.007$ , hazard ratio = 3.24) compared with the AA genotype in the NK-AML patients. Our results indicate that rs2072671 in *CDA* may be an important prognostic marker in NK-AML patients.

*Journal of Human Genetics* (2015) 60, 749–754; doi:10.1038/jhg.2015.105; published online 10 September 2015

## INTRODUCTION

Acute myeloid leukemia (AML) is the most common acute leukemia affecting adults. It arises from the accumulation of abnormal myeloid cells in the bone marrow that disrupts the production of normal blood cells.<sup>1</sup> Cytogenetic characterization of blast cells is the most important prognostic factor for predicting remission rate, relapse and overall survival (OS). Based on cytogenetic findings, AML patients can be categorized into three different risk subgroups: favorable, intermediate and poor.<sup>2,3</sup> Approximately 60% of patients with AML harbor abnormal karyotypes (AK) that are non-random, somatically acquired chromosomal translocations and inversions.<sup>3,4</sup> Among the remaining, as many as 40% of patients with AML have no cytogenetic abnormality, representing normal karyotype (NK), and are classified into the intermediate-risk group.<sup>3,5</sup> For patients in the favorable and poor cytogenetic risk groups, treatment choice is well defined. However, for patients in the intermediate-risk group having no cytogenetic abnormalities, the most optimal treatment is controversial because this group is very heterogeneous about treatment response, survival and risk of relapse.<sup>6,7</sup> The *FLT3-ITD*, *NPM1* and *CEBPA*

mutations have been reported to predict OS and first complete remission duration (CR1D) in NK-AML.<sup>8–11</sup> The inherited genetic polymorphisms are also reported in *SULT1C2*, *XPA* and *MDR1* in intermediate-risk group AML patients. These genes encode drug-metabolism enzyme, drug transport protein and DNA repair protein.<sup>12</sup> Whether these genetic markers classify a clinically relevant subgroup in NK-AML patients remains uncertain. For individualizing treatment, there is still need for additional genetic markers in NK-AML patients.

The nucleoside analog cytosine arabinoside (AraC) is a major component of the most common treatment for AML patients. It is metabolized by a number of pharmacogenetic enzymes (for example, cytidine deaminase, deoxycytidine kinase, dCMP deaminase and 5'-nucleotidase).<sup>13–15</sup> Among these drug-metabolizing enzymes, cytidine deaminase (CDA) is related with pyrimidine salvaging and catalyzes the irreversible hydrolytic deamination of cytidine to uridine nucleosides.<sup>16</sup> Therefore, CDA is responsible for the inactivation of AraC during chemotherapy of AML. One nonsynonymous single-nucleotide polymorphism (SNP), rs2072671 (79A>C), in *CDA* changes lysine to glutamine, resulting in decreased enzyme activity,

<sup>1</sup>Department of Life Science, Sogang University, Seoul, Republic of Korea; <sup>2</sup>Department of Genetic Epidemiology, SNP Genetics, Seoul, Republic of Korea; <sup>3</sup>Cancer Research Institute, Seoul National University Hospital, Seoul, Republic of Korea; <sup>4</sup>Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea; <sup>5</sup>Functional Genome Institute, PDxEn Biosystem, Seoul, Republic of Korea; <sup>6</sup>Department of Biochemistry, School of Medicine, Ewha Womans University, Seoul, Republic of Korea and <sup>7</sup>Biomedical Research Institute, Seoul National University Hospital, Seoul, Republic of Korea

Correspondence: Dr HD Shin, Department of Life Science, Sogang University, 1 Shinsu-dong, Mapo-gu, Seoul 121-742, Republic of Korea.

E-mail: hdshin@sogang.ac.kr

or Dr S-S Yoon, Department of Internal Medicine, Seoul National University Hospital, Seoul National University College of Medicine, 101 Daehang-no, Jongno-gu, Seoul 110-744, Korea.

E-mail: ssysmc@snu.ac.kr

Received 12 June 2015; revised 20 July 2015; accepted 30 July 2015; published online 10 September 2015

reported several *in vitro* studies.<sup>17,18</sup> Previous Chinese AML association study reported that rs2072671 was associated with decreased sensitivity to AraC used in chemotherapy of childhood leukemias.<sup>19</sup> More recently, it was reported that the 79 C/C and -451T/T genotypes in *CDA* were associated with decreased global DNA methylation level in cells and shorter OS in FLT3-ITD-positive/NPM1-positive patients in NK-AML.<sup>20</sup>

The goal of this study was to investigate the potential roles of common genetic polymorphisms in *CDA* gene as prognostic factors in AML through the comparative analysis of 95 AMLs (50 NK-AMLs and 45 AK-AMLs) and 241 normal controls (NC) in Korean population. Our findings indicate that polymorphisms of *CDA* could be the potential diagnostic and prognostic markers in the NK-AML patients.

## MATERIALS AND METHODS

### Patient characteristics and treatment protocol

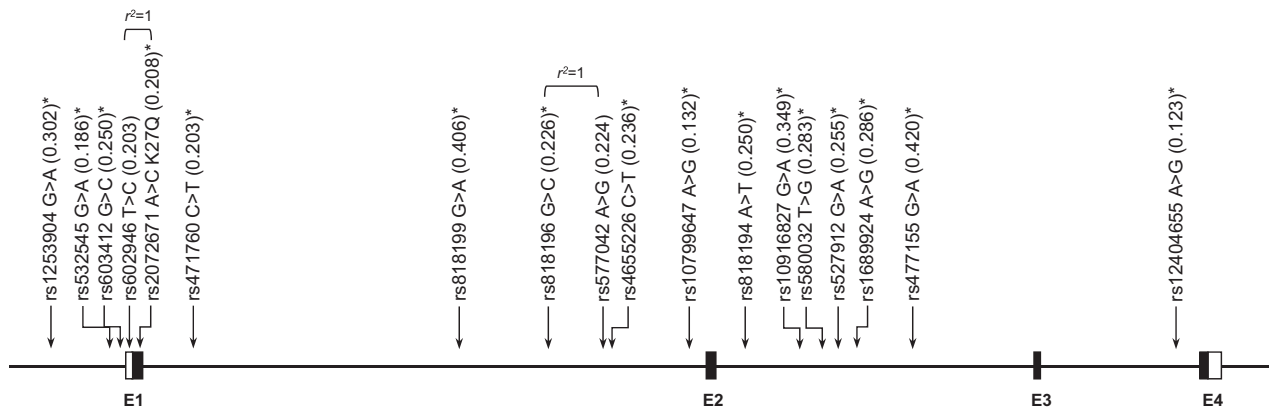
A total of 95 newly diagnosed AML patients (56 males and 39 females) who had marrow samples available for DNA isolation were enrolled in the study from Seoul National University Hospital, Seoul, Korea between December 2001 and September 2009. The patient characteristics were as follows: the median age was 49 years (range 16–76 years). The FAB (French–American–British) classification showed the following distribution of subtypes: M0, 2 patients (2%); M1, 13 (14%); M2, 43 (45%); M4, 30 (32%); M5, 5 (5%); and M7, 2 patients (2%).

Patients diagnosed with acute promyelocytic leukemia (M3 FAB subtype) were excluded from the study. The criteria used to describe a cytogenetic clone and karyotype followed the recommendations of the International System for Human Cytogenetic Nomenclature.<sup>21</sup> The cytogenetic risk groups were classified according to the MRC10 criteria, as described previously<sup>22</sup> (favorable risk: AML associated with t(8;21), t(15;17), or inv(16); unfavorable risk: the presence of a complex karyotype, -5, del(5q), -7 or abnormalities of 3q; intermediate risk: the remaining group of patients, including those with 11q23 abnormalities, +8, +21, +22, del(9q), del(7q) or other miscellaneous structural or numerical defects not included in the favorable or unfavorable risk groups). Information of first induction chemotherapy outcome was available in 90 patients. All patients were treated with standard induction chemotherapy (idarubicin 12 mg m<sup>-2</sup> per day intravenous for 3 consecutive days and AraC 200 mg m<sup>-2</sup> per day intravenous for 7 consecutive days). In patients over 60 years of age, the dose of idarubicin was modified to 10 mg m<sup>-2</sup> per day. Consolidation therapies were performed based on two more cycles of high-dose AraC (3 g m<sup>-2</sup> per day intravenous twice a day on days 1, 3 and 5). The institutional review board of Seoul National University Hospital approved the study protocol, and all subjects provided informed consent.

### SNP genotyping

Among 46 HapMap SNPs (release #27) detected in *CDA*, 18 tagging SNPs were selected based on their minor allele frequencies (>5%) and linkage disequilibrium in the HapMap database. Genotyping was performed at a multiplex level

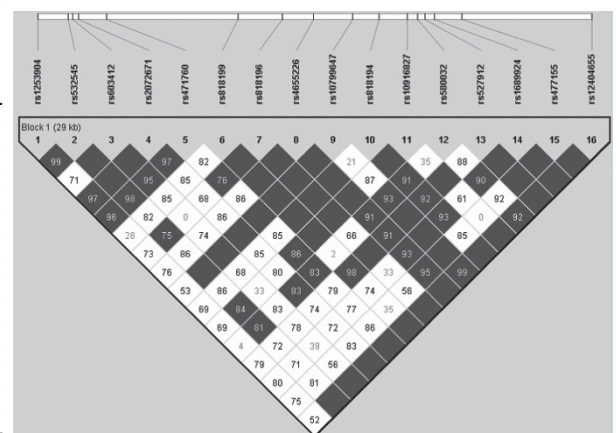
### a Map of *CDA* (cytidine deaminase) on chromosome 1p36.2-p35 (32 kb)



### b Haplotypes in *CDA*

Hap.	rs1253904	rs532545	rs603412	rs2072671	rs471760	rs818199	rs818196	rs4655226	rs10799647	rs818194	rs10916827	rs580032	rs527912	rs1689924	rs477155	rs12404655	Freq.
ht1	G	G	G	A	C	A	G	C	A	A	G	T	A	G	A	A	0.272
ht2	G	G	G	A	C	G	C	C	A	A	A	G	G	A	G	A	0.161
ht3	A	A	C	C	T	G	G	T	A	T	G	T	G	A	G	A	0.131
ht4	G	G	G	A	C	G	G	C	G	A	A	T	G	A	A	G	0.110
ht5	A	G	G	A	C	A	G	C	A	A	G	G	G	A	G	A	0.073
ht6	G	G	G	A	C	G	G	T	A	T	G	T	G	A	G	A	0.035
ht7	G	G	G	A	C	A	G	C	A	A	G	T	G	A	G	A	0.033
ht8	G	G	G	A	C	A	G	C	A	A	G	G	G	A	G	A	0.033
ht9	G	G	C	A	C	G	C	C	A	A	T	G	G	A	A	A	0.031
ht10	A	G	G	A	C	G	C	C	G	A	A	T	G	A	A	G	0.017
ht11	G	G	G	A	C	G	G	C	G	T	G	T	G	A	A	A	0.013
ht12	A	G	C	C	T	G	G	T	A	T	G	T	G	A	G	A	0.013
ht13	A	G	G	A	C	G	C	C	A	A	A	T	G	G	A	A	0.013
ht14	A	G	G	A	C	G	C	C	A	A	A	G	G	A	G	A	0.010
Others	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.054

### c LD coefficient ( $r^2$ ) among *CDA* SNPs



**Figure 1** Gene map of *cytidine deaminase* (*CDA*). (a) Map of *CDA* on chromosome 1p36.2–p35. Coding exons are marked by black blocks and 5' and 3' untranslated regions (UTRs) by white blocks (Ref. NM#. NM\_001785.2). The frequencies in parentheses were based on the Korean acute myeloid leukemia (AML) patients ( $n=95$ ). Asterisks indicate single-nucleotide polymorphisms (SNPs) that were used in statistical analysis. (b) Haplotypes in *CDA*. Sixteen *CDA* polymorphisms were used for haplotype construction. Five haplotypes with over 5% of frequency were used in statistical analyses. (c) Linkage disequilibrium (LD) coefficient ( $r^2$ ) among *CDA* polymorphisms. Scores in the boxes mean  $r^2$ .

**Table 1 Allele test of 16 CDA SNPs and haplotypes with risk of NK-AML compared with NCs**

Rs#	Allele	Position (AA)	NK-AML	NCs	Codominant			Dominant			Recessive		
			(n = 50)	(n = 241)	OR (95% CI)	P*	P**	OR (95% CI)	P*	P**	OR (95% CI)	P*	P**
rs1253904	G>A	Promoter	0.34	0.25	1.6 (0.96–2.67)	0.07	–	1.69 (0.87–3.26)	0.12	–	2.17 (0.69–6.76)	0.18	–
rs532545	G>A	Promoter	0.24	0.14	1.93 (1.07–3.49)	<b>0.03</b>	NS	2.18 (1.11–4.27)	<b>0.02</b>	NS	1.83 (0.26–12.88)	0.55	–
rs603412	G>C	Promoter	0.28	0.19	1.64 (0.94–2.88)	0.08	–	1.9 (0.99–3.66)	0.06	–	1.14 (0.19–6.89)	0.88	–
rs2072671	A>C	Exon 1 (K27Q)	0.27	0.15	2.13 (1.19–3.82)	<b>0.01</b>	NS	2.44 (1.25–4.75)	<b>0.009</b>	NS	1.83 (0.26–12.93)	0.54	–
rs471760	C>T	Intron 1	0.27	0.16	2.07 (1.15–3.71)	<b>0.02</b>	NS	2.34 (1.21–4.56)	<b>0.01</b>	NS	1.83 (0.26–12.88)	0.55	–
rs818199	G>A	Intron 1	0.37	0.45	0.63 (0.39–1.01)	0.06	–	0.44 (0.23–0.87)	<b>0.02</b>	NS	0.77 (0.33–1.78)	0.54	–
rs577042	A>G	Intron 1	0.17	0.24	0.74 (0.41–1.35)	0.32	–	0.84 (0.43–1.66)	0.62	–	–	0.98	–
rs4655226	C>T	Intron 1	0.32	0.16	2.5 (1.47–4.24)	<b>0.0007</b>	<b>0.03</b>	3.02 (1.54–5.93)	<b>0.001</b>	NS	3.66 (1.05–12.72)	<b>0.04</b>	NS
rs10799647	A>G	Intron 1	0.14	0.15	0.97 (0.51–1.85)	0.92	–	0.99 (0.47–2.06)	0.97	–	0.76 (0.08–7.42)	0.81	–
rs818194	A>T	Intron 2	0.32	0.18	2.19 (1.3–3.69)	<b>0.003</b>	NS	2.49 (1.28–4.85)	<b>0.007</b>	NS	3.32 (0.98–11.26)	0.05	–
rs10916827	G>A	Intron 2	0.31	0.37	0.88 (0.54–1.43)	0.59	–	1 (0.52–1.96)	0.99	–	0.53 (0.17–1.67)	0.28	–
rs580032	T>G	Intron 2	0.21	0.31	0.67 (0.4–1.14)	0.14	–	0.71 (0.37–1.37)	0.31	–	0.29 (0.06–1.42)	0.13	–
rs527912	G>A	Intron 2	0.26	0.32	0.65 (0.39–1.08)	0.10	–	0.53 (0.27–1.04)	0.06	–	0.71 (0.23–2.19)	0.55	–
rs1689924	A>G	Intron 2	0.28	0.36	0.6 (0.37–0.99)	<b>0.04</b>	NS	0.55 (0.28–1.05)	0.07	–	0.43 (0.14–1.3)	0.13	–
rs477155	G>A	Intron 2	0.42	0.51	0.6 (0.38–0.96)	<b>0.03</b>	NS	0.6 (0.29–1.21)	0.15	–	0.4 (0.16–0.97)	<b>0.04</b>	NS
rs12404655	A>G	Intron 3	0.14	0.13	1.14 (0.59–2.2)	0.69	–	1.19 (0.57–2.5)	0.64	–	0.94 (0.09–9.97)	0.96	–
CDA-ht1	–	–	0.23	0.29	0.63 (0.36–1.1)	0.11	–	0.56 (0.28–1.09)	0.09	–	0.64 (0.16–2.54)	0.53	–
CDA-ht2	–	–	0.13	0.17	0.81 (0.42–1.56)	0.52	–	0.91 (0.44–1.9)	0.80	–	–	0.98	–
CDA-ht3	–	–	0.23	0.11	2.56 (1.36–4.8)	<b>0.003</b>	NS	2.78 (1.39–5.57)	<b>0.004</b>	NS	3.69 (0.39–34.9)	0.26	–
CDA-ht4	–	–	0.13	0.11	1.42 (0.72–2.83)	0.31	–	1.51 (0.7–3.24)	0.29	–	1.41 (0.11–17.95)	0.79	–
CDA-ht5	–	–	0.05	0.07	0.78 (0.29–2.1)	0.63	–	0.61 (0.2–1.89)	0.40	–	–	0.99	–

Abbreviations: AA, amino acid; CDA, cytidine deaminase; CI, confidence interval; MAF, minor allele frequency; NC, normal controls; NK-AML, normal karyotype acute myeloid leukemia; NS, not significant; OR, odds ratio; SNP, single-nucleotide polymorphism.

\**P*-value of logistic analysis by adjusting age and sex as a covariate.

\*\**P*<sub>corr</sub>-value after Bonferroni correction by multiplying the number of independent maker SNPs (*n* = 16) and three different models (*n* = 3).

Bold values indicate the statistical significance of *P* < 0.05.

using the Illumina GoldenGate genotyping system (Illumina Inc., San Diego, CA, USA).<sup>23</sup> The genotype quality score for retaining data was set to 0.25 and all SNPs were successfully genotyped.

### Statistics

To determine whether the individual variant was in equilibrium at each locus in the population (Hardy–Weinberg equilibrium),  $\chi^2$  tests were performed. We examined a widely used measure of linkage disequilibrium between all pairs of biallelic loci, *D'* (the correlation coefficient [ $\Delta$ ,  $|D'|$ ]) and *r*<sup>2</sup> using Haploview.<sup>24</sup> Association with AML was analyzed under a logistic model by adjusting for age and sex as a covariate using the SAS (Cary, NC, USA). In order to correct for multiple comparisons, Bonferroni correction was applied based on the number of independent SNPs (*n* = 16) and three different models (*n* = 3). To determine the ethnic difference of CDA polymorphisms,  $\chi^2$  test and Fisher's exact test were performed. Cox proportional hazard model was applied to determine the significance of difference in OS times and CRID. Haplotypes were inferred using PHASE algorithm ver. 2.0.<sup>25</sup> Subsequently, statistical analysis was performed using SAS version 9.1.

### RESULTS

In order to examine the relationship between common CDA polymorphisms and AML patients, 18 SNP markers were genotyped using the Illumina GoldenGate genotyping assay. Pairwise comparisons between the SNPs revealed two sets of absolute linkage disequilibriums ( $|D'| = 1$  and  $r^2 = 1$ ) (Figure 1c). The position and frequency of the genetic variants genotyped in the CDA gene are shown in Figure 1a. A total of 16 SNPs (3 SNPs in the promoter: rs1253904, rs532545 and rs603412; 1 nonsynonymous SNP: rs2072671; and 12 SNPs in intron: rs471760, rs818199, rs577042, rs4655226, rs10799647,

rs818194, rs10916827, rs580032, rs527912, rs1689924, rs477155 and rs12404655) were analyzed. Among the 14 haplotypes constructed, 5 common haplotypes (frequency > 5%) were used for further analysis (Figure 1b).

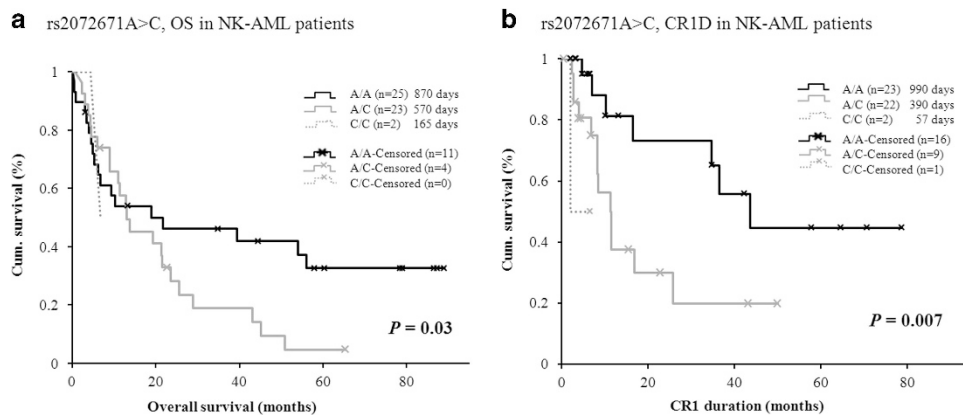
A total of 336 samples were used for the association study, consisting of 95 AMLs (50 NK-AML and 45 AK-AML) and 241 NC. The observed genotypes and allele frequencies for the CDA SNPs were in the Hardy–Weinberg equilibrium (*P* > 0.05). Several CDA SNPs and haplotypes manifested a marginal association with AML when the entire AMLs and NC were analyzed (*P* = 0.01–0.04; Supplementary Table 1). Four SNPs were also marginally associated with CRID after chemotherapy was applied to the AMLs (Supplementary Table 2). However, when the patients were stratified according to karyotype status, the most significantly associated variant was found between NK-AML and NC (Table 1). In contrast, none of SNPs and haplotypes manifested any association between AK-AML and NC (Supplementary Table 3). The rs532545, rs2072671, rs471760, rs4655226, rs818194 and ht3 were strongly associated with the risk of NK-AML (*P* = 0.03–0.0007; Table 1). The most significant association was discovered in the intronic SNP, rs4655226 (*P* = 0.0007, *P*<sub>corr</sub> = 0.03 and odds ratio = 2.5; Table 1). Although the marginal association of the nonsynonymous SNP, rs2072671 (K27Q), completely disappeared after Bonferroni correction, rs4655226 maintained the significance (*P*<sub>corr</sub> = 0.03) that was strongly linked with rs2072671 ( $r^2 = 0.86$ ). In addition, rs2072671 was reported as having significant associations with enzyme activity in previous studies.<sup>20,26</sup> Therefore, in subsequent statistical analysis, we selected the nonsynonymous SNP, rs2072671, as a most promising candidate.

**Table 2** Association analysis of rs2072671 with NK-AML compared with NC and AK-AML

Analyzing model	Genotype	NK-AML (n = 50)	NC (n = 241)	OR (95% CI)	P*
Codominant	A/A	25 (50%)	172 (71.3%)	2.13 (1.19–3.82)	<b>0.01</b>
	A/C	23 (46%)	65 (27%)		
	C/C	2 (4%)	4 (1.7%)		
Dominant	A/A	25 (50%)	172 (71.3%)	2.44 (1.25–4.75)	<b>0.009</b>
	A/C+C/C	25 (50%)	10 (28.6%)		
Recessive	A/A+A/C	48 (96%)	273 (98.3%)	1.83 (0.26–12.93)	0.54
	C/C	2 (4%)	4 (1.7%)		
Referent	A/A	25 (50%)	172 (71.3%)	1	–
	A/C	23 (46%)	65 (27%)		
	C/C	2 (4%)	4 (1.7%)		
				1.82 (0.75–4.36)	0.18

Abbreviations: AK-AML, abnormal karyotype AML; AML, acute myeloid leukemia; CI, confidence interval; NC, normal controls; NK-AML, normal karyotype AML; OR, odds ratio.

\*P-value of logistic analysis values of four alternative models (codominant, dominant, recessive and referent models) by adjusting age and sex as a covariate. Bold values indicate statistical significance of  $P < 0.05$ .



**Figure 2** Differences in overall survival (OS) and first complete remission duration (CR1D) depending on rs2072671 (79 A>C). In normal karyotype acute myeloid leukemia (AML) patients, heterozygous AC and rare homozygous CC genotypes were significantly associated with a shorter OS and CR1D. (a) The mean OS was 5.5 vs 19 vs 29 months for CC vs AC vs AA,  $P = 0.03$ . (b) The mean CR1D was 1.9 vs 13 vs 33 months for CC vs AC vs AA,  $P = 0.007$ .

In the logistic analysis, rs2072671 was found to be significantly associated with the risk of NK-AML compared with NC ( $P = 0.009$  and odds ratio = 2.44). The frequencies of the AC and CC genotypes were higher in the NK-AML than that of the NC genotype (0.50 vs 0.28 in NK-AML vs NC; Table 2). Taking the subjects carrying the AA genotype for rs2072671 as a reference, the subjects carrying genotype AC showed an increased risk of NK-AML ( $P = 0.008$  and odds ratio = 2.46).

In the Cox regression analysis, rs2072671 (K27Q) was shown to have significant correlation with clinical outcomes of NK-AML after the first course of chemotherapy (Figure 2). NK-AML patients with AC or CC genotypes had worse OS ( $P = 0.03$ , hazard ratio = 1.84) in Figure 2a and CR1D ( $P = 0.007$ , hazard ratio = 3.24) outcomes compared with patients with the common homozygous AA genotype in Figure 2b. In contrast, rs2072671 genotypes were not associated with treatment response in AK-AML patients (Supplementary Figure 1).

## DISCUSSION

CDA is an important enzyme responsible for catalyzing the irreversible hydrolytic deamination of cytidine into uridine nucleoside, one of the enzymes that regulate the homeostasis of the cellular pyrimidine pool.<sup>16</sup> Therefore, CDA plays a crucial role in the metabolism of AraC, the pyrimidine analog drug for AML treatments, and variants of CDA may influence subsequent treatments through the alteration of gene

expression and function. Our results showed several CDA SNPs and haplotypes have significant correlations with the increased risk of NK-AML.

The nonsynonymous SNP, rs2072671 (K27Q), is highly associated with the risk of NK-AML and has a strong relationship with the clinical outcomes of chemotherapy for NK-AML. The heterozygotes (AC) and minor allele homozygotes (CC) of rs2072671 A>C are more common in NK-AML than in AK-AML and NC. The AC and CC were also significantly associated with a shorter OS and CR1D in NK-AML patients. A study by Falk *et al.*<sup>20</sup> reported that there was no significant difference in rs2072671 and rs532545 genotype distribution between NK-AML and HapMap-CEU reference ( $P > 0.8$ ). Moreover, these two SNPs were not associated with the treatment response of the entire NK-AMLs in the Swedish population. This discrepancy between the Swedish study and this study may be attributed to the allele frequency differences between the Korean and Swedish populations (Supplementary Table 4). In order to identify the ethnic differences between the 16 CDA SNPs, we analyzed the genotype distribution used in Fisher's test among four ethnicities, including Korean, Japanese, Han Chinese and European-American populations, using Hapmap data (Supplementary Table 5). The distribution of rs532545, rs603412, rs2072671, rs471760, rs580032 and rs1689924 genotypes was significantly different between the Korean and European-American populations. These frequency differences between the CDA SNPs may reflect a distinct ethnic background between the

Asian and European-American patients, and thus might affect the clinical phenotype of AML.

Karyotype abnormalities are appeared to be associated with multi-drug resistance (MDR) gene regulation. MDR-positive phenotype is more often observed in patients with abnormal karyotype.<sup>27,28</sup> In AML patients with abnormal karyotypes, highly expressed MDR gene may be correlated with response to chemotherapy. In contrast, in NK-AML patients, without MDR overexpression, CDA gene may be one of major regulating factor.

A study by Vasile *et al.*<sup>29</sup> indicated that the AA genotype of rs2072671 was associated with an improved progression-free survival in non-small-cell lung cancer patients treated with gemcitabine, and this was similar to our results. In previous studies, the AA variant of the major allele homozygote in rs2072671 was found to increase CDA enzyme activity and *in vitro* deamination of AraC.<sup>17,18</sup> The samples with heterozygous (AC) and minor homozygote (CC) genotypes in rs2072671 had significantly reduced CDA enzyme activity compared with samples for the common homozygous (AA) genotype.<sup>30</sup> The decreased CDA activity may not maintain the cellular pyrimidine pool.<sup>31</sup> It could instead increase the chance of mismatch during DNA replication. As such, point mutations would occur more frequently in patients with the AC and CC genotypes. Although NK-AML patients have normal karyotypes at the chromosomal level, they may have increased point mutations at the DNA level. The whole-exome sequencing method would be useful in verifying this hypothesis. However, we were unable to apply exome sequencing in this study. On the other hand, variants of CDA might influence to reduce the degree of DNA methylation. Farrell *et al.*<sup>32</sup> suggested that pharmacological inhibition of CDA affected the lower level of DNA methylation. Other study group reported that NK-AML patients with C allele of rs2072671 and T allele of rs532545 showed a significantly less methylated pattern.<sup>20</sup> Together, variants of CDA may influence the decreased DNA methylation level and alter the normal gene expression and downstream signaling in NK-AML patients. Another possible mechanism explaining the poor prognosis in NK-AML by the CDA variant is that the chemotherapeutic agent may be remaining at high concentrations in the blood as a result of the decreased CDA activity, thereby inducing the apoptosis of normal blood cells and increasing the cytotoxicity.<sup>33</sup>

In our study, the insufficient number of NK-AML cases is a major limitation. Further study will be required to validate the association in an independent population. In conclusion, our results strongly support CDA SNP as a potential prognostic marker for NK-AML patients. Among the several possibilities, the precise functional mechanisms underlying the harmful effects of the heterozygotes and the rare allele homozygotes of CDA rs2072671 on the clinical outcome of NK-AML still remain to be elucidated.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

This work was supported by the SNUH Research Fund (03-2011-0010); Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare (HI14C0072); the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2011-0008846).

- 1 Lowenberg, B., Downing, J. R. & Burnett, A. Acute myeloid leukemia. *N. Engl. J. Med.* **341**, 1051–1062 (1999).
- 2 Wheatley, K., Burnett, A. K., Goldstone, A. H., Gray, R. G., Hann, I. M., Harrison, C. J. *et al.* A simple, robust, validated and highly predictive index for the determination of risk-directed therapy in acute myeloid leukaemia derived from the MRC AML 10 trial. United Kingdom Medical Research Council's Adult and Childhood Leukaemia Working Parties. *Br. J. Haematol.* **107**, 69–79 (1999).
- 3 Slovak, M. L., Kopecky, K. J., Cassileth, P. A., Harrington, D. H., Theil, K. S., Mohamed, A. *et al.* Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group Study. *Blood* **96**, 4075–4083 (2000).
- 4 Look, A. T. Oncogenic transcription factors in the human acute leukemias. *Science* **278**, 1029–1032 (1997).
- 5 Rausei-Mills, V., Chang, K. L., Gaal, K. K., Weiss, L. M. & Huang, Q. Aberrant expression of CD7 in myeloblasts is highly associated with de novo acute myeloid leukemias with FLT3/ITD mutation. *Am. J. Clin. Pathol.* **129**, 624–629 (2008).
- 6 Brunet, S., Esteve, J., Berlanga, J., Ribera, J. M., Bueno, J., Marti, J. M. *et al.* Treatment of primary acute myeloid leukemia: results of a prospective multicenter trial including high-dose cytarabine or stem cell transplantation as post-remission strategy. *Haematologica*. **89**, 940–949 (2004).
- 7 Lowenberg, B. Strategies in the treatment of acute myeloid leukemia. *Haematologica*. **89**, 1029–1032 (2004).
- 8 Du, Y. Z., Su, L., Li, W., Yu, P., Tan, Y. H., Lin, H. *et al.* [Clinical characteristics and therapeutic efficacy of normal karyotype AML patients with CEBPA mutation]. *Zhongguo Shi Yan Xue Ye Xue Za Zhi*. **22**, 16–19 (2014).
- 9 Schnittger, S., Schoch, C., Kern, W., Mecucci, C., Tschulik, C., Martelli, M. F. *et al.* Nucleophosmin gene mutations are predictors of favorable prognosis in acute myelogenous leukemia with a normal karyotype. *Blood* **106**, 3733–3739 (2005).
- 10 Frohling, S., Schlenk, R. F., Breitruck, J., Benner, A., Kreitmeier, S., Tobis, K. *et al.* Prognostic significance of activating FLT3 mutations in younger adults (16 to 60 years) with acute myeloid leukemia and normal cytogenetics: a study of the AML Study Group Ulm. *Blood* **100**, 4372–4380 (2002).
- 11 Tallman, M. S., Gilliland, D. G. & Rowe, J. M. Drug therapy for acute myeloid leukemia. *Blood* **106**, 1154–1163 (2005).
- 12 Mardis, E. R., Ding, L., Dooling, D. J., Larson, D. E., McLellan, M. D., Chen, K. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N. Engl. J. Med.* **361**, 1058–1066 (2009).
- 13 Monzo, M., Brunet, S., Urbano-Ispizua, A., Navarro, A., Perea, G., Esteve, J. *et al.* Genomic polymorphisms provide prognostic information in intermediate-risk acute myeloblastic leukemia. *Blood* **107**, 4871–4879 (2006).
- 14 Ellison, R. R., Holland, J. F., Weil, M., Jacquillat, C., Boiron, M., Bernard, J. *et al.* Arabinosyl cytosine: a useful agent in the treatment of acute leukemia in adults. *Blood* **32**, 507–523 (1968).
- 15 Keating, M. J., McCredie, K. B., Bodey, G. P., Smith, T. L., Gehan, E. & Freireich, E. J. Improved prognosis for long-term survival in adults with acute myelogenous leukemia. *JAMA* **248**, 2481–2486 (1982).
- 16 Betts, L., Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W. Jr. Cytidine deaminase. The 2.3A crystal structure of an enzyme: transition-state analog complex. *J. Mol. Biol.* **235**, 635–656 (1994).
- 17 Gilbert, J. A., Salavaggi, O. E., Ji, Y., Pelleymounter, L. L., Eckloff, B. W., Wieben, E. D. *et al.* Gemcitabine pharmacogenomics: cytidine deaminase and deoxycytidylate deaminase gene resequencing and functional genomics. *Clin. Cancer Res.* **12**, 1794–1803 (2006).
- 18 Kirch, H. C., Schroder, J., Hoppe, H., Esche, H., Seeber, S. & Schutte, J. Recombinant gene products of two natural variants of the human cytidine deaminase gene confer different deamination rates of cytarabine *in vitro*. *Exp. Hematol.* **26**, 421–425 (1998).
- 19 Bhatla, D., Gerbing, R. B., Alonzo, T. A., Conner, H., Ross, J. A., Meshinchi, S. *et al.* Cytidine deaminase genotype and toxicity of cytosine arabinoside therapy in children with acute myeloid leukemia. *Br. J. Haematol.* **144**, 388–394 (2009).
- 20 Falk, I. J., Fyrberg, A., Paul, E., Nahii, H., Hermanson, M., Rosenquist, R. *et al.* Decreased survival in normal karyotype AML with single-nucleotide polymorphisms in genes encoding the AraC metabolizing enzymes cytidine deaminase and 5'-nucleotidase. *Am. J. Hematol.* **88**, 1001–1006 (2013).
- 21 Mitelman, F. *An International System for Human Cytogenetic Nomenclature* (Karger, Basel, Switzerland, 1995).
- 22 Grimwade, D., Walker, H., Oliver, F., Wheatley, K., Harrison, C., Harrison, G. *et al.* The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood* **92**, 2322–2333 (1998).
- 23 Oliphant, A., Barker, D. L., Stuelpnagel, J. R. & Chee, M. S. BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques*. **56–58**(Suppl), 60–61 (2002).
- 24 Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
- 25 Stephens, M., Smith, N. J. & Donnelly, P. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* **68**, 978–989 (2001).
- 26 Abraham, A., Varatharajan, S., Abbas, S., Zhang, W., Shaji, R. V., Ahmed, R. *et al.* Cytidine deaminase genetic variants influence RNA expression and cytarabine cytotoxicity in acute myeloid leukemia. *Pharmacogenomics*. **13**, 269–282 (2012).
- 27 Goasguen, J. E., Dossot, J. M., Fardel, O., Le Mee, F., Le Gall, E., Leblay, R. *et al.* Expression of the multidrug resistance-associated P-glycoprotein (P-170) in 59 cases of

- de novo acute lymphoblastic leukemia: prognostic implications. *Blood* **81**, 2394–2398 (1993).
- 28 Leith, C. P., Kopecky, K. J., Chen, I. M., Eijdem, L., Slovak, M. L., McConnell, T. S. *et al*. Frequency and clinical significance of the expression of the multidrug resistance proteins MDR1/P-glycoprotein, MRP1, and LRP in acute myeloid leukemia: a Southwest Oncology Group Study. *Blood* **94**, 1086–1099 (1999).
- 29 Vasile, E., Giovannetti, E., Tibaldi, C., Mey, V., Nannizzi, S., Landi, L. *et al*. Analysis of single-nucleotide polymorphisms (SNPs) of cytidine deaminase (CDA) genes for the prediction of clinical response to gemcitabine and cisplatin in advanced non-small-cell lung cancer (NSCLC) patients. *J. Clin. Oncol.* **24**, abstract 7219 (2006).
- 30 Mahfouz, R. Z., Jankowska, A., Ebrahim, Q., Gu, X., Visconte, V., Tabaroki, A. *et al*. Increased CDA expression/activity in males contributes to decreased cytidine analog half-life and likely contributes to worse outcomes with 5-azacytidine or decitabine therapy. *Clin. Cancer Res.* **19**, 938–948 (2013).
- 31 Kuhn, K., Bertling, W. M. & Emmrich, F. Cloning of a functional cDNA for human cytidine deaminase (CDD) and its use as a marker of monocyte/macrophage differentiation. *Biochem. Biophys. Res. Commun.* **190**, 1–7 (1993).
- 32 Farrell, J. J., Bae, K., Wong, J., Guha, C., Dicker, A. P. & Elsaleh, H. Cytidine deaminase single-nucleotide polymorphism is predictive of toxicity from gemcitabine in patients with pancreatic cancer: RTOG 9704. *Pharmacogenomics J.* **12**, 395–403 (2012).
- 33 Lemaire, M., Momparler, L. F., Raynal, N. J., Bernstein, M. L. & Momparler, R. L. Inhibition of cytidine deaminase by zebularine enhances the antineoplastic action of 5-aza-2'-deoxycytidine. *Cancer Chemother. Pharmacol.* **63**, 411–416 (2009).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)