Research Paper

Prognostic impact of mitochondrial DNA D-loop variations in pediatric acute myeloid leukemia

Anudishi Tyagi[1], Raja Pramanik[1], Sreenivas Vishnubhatla[2], Radhika Bakhshi[3] and Sameer Bakhshi[1]

[1]Department of Medical Oncology, Dr. B. R. A. Institute Rotary Cancer Hospital, All India Institute of Medical Sciences, New Delhi, India

[2]Department of Biostatistics, All India Institute of Medical Sciences, New Delhi, India

[3]Department of Biomedical Sciences, Shaheed Rajguru College of Applied Sciences, University of Delhi, New Delhi, India

Correspondence to: Sameer Bakhshi, *email*: sambakh@hotmail.com

ABSTRACT

The role of mitochondrial DNA (mt-DNA) changes, especially those in the regulatory D-loop region in Acute Myeloid Leukemia (AML) remains investigational. Consecutive 151 *de novo* pediatric AML patients, (≤18 yr) were prospectively enrolled from June 2013-August 2016, to assess the prognostic impact of mt-DNA D-loop variations (somatic/germline) on survival. For each patient, D-loop region was sequenced on baseline bone marrow and buccal swab, and mother's blood sample. In 151 AML subjects, 1490 variations were found at 237 positions; 80.9% were germline and 19.1% somatic. The mean number of variations per position was 6.3. Variations with frequency ≥6 were analyzed for their impact on survival and 4 categories were created, namely "somatic-protective", "somatic-hazardous", "germline-protective" and "germline- hazardous". Although, somatic-protective could not predict event free survival (EFS) or overall survival (OS), somatic-hazardous [(OS) HR = 2.33, p = 0.06] and germline-hazardous [(OS) HR = 2.85, p < 0.01] significantly predicted OS and EFS. Notably, the germline-protective, could significantly predict EFS (HR = 0.31, p = 0.03) and OS (HR = 0.19, p < 0.01), only when variations at ≥2 positions were present. On multivariate analysis, three positions namely 16111, 16126, 16362 and karyotype were found to be predictive of EFS. A prognostic index (PI) was developed using nomogram PI = (0.8*karyotype) + (1.0*c16111) + (0.7*t16362) + (1.2*t16126). Hazard ratio for EFS increased significantly with increasing PI reaching to a maximum of 3.3 (p < 0.01). In conclusion, the impact of mt-DNA D-loop variations on outcomes in pediatric AML depends on their nature (germline/somatic), position and mutational burden, highlighting their potential role as evolving prognostic biomarkers.

INTRODUCTION

Mitochondria are a unique organelle and so is the mitochondrial DNA (mt-DNA). The human mitochondrial genome is 16,569 base pair (bp) in length, with double-stranded circular DNA molecules containing 37 genes [1]. A High turnover, lack of histones protection and poor proof reading ability of mt-DNA polymerase gamma render it highly susceptible to damage. Further, its close proximity to the electron transport chain, which generates reactive oxygen species (ROS), makes it even more prone to damage [2]. Lack of introns in mt-DNA ensures that most mutations occur in the coding sequence leading to biological consequences. Role of mitochondria in various human malignancies including leukemia has been long proposed and explored with varying outcomes [3–5]. This genome also includes a non-coding displacement region (D-loop) which consists of 1122 bp (16024 – 577 bp) of mitochondrial DNA. It acts as a promoter region for both the heavy and light strands of the mt-DNA, and contains essential transcription and replication elements [6]. It contains three hyper

variable regions (HV1 at positions 16024–16383, HV2 at positions 57–372 and HV3 at positions 438–574). The D-loop is a hot spot region for mt-DNA variations. Genetic variability in the D-loop region has been suggested to affect the function of the respiration chain, leading to high ROS levels and instability in the mt-DNA [7, 8]. Thus, it is not surprising that mitochondrial dysfunction has been linked to human degenerative diseases and cancers, including leukemia [9, 10].

AML is a heterogeneous disease characterized by different cytogenetic aberrations, acquired mutations and impaired gene expression. Zhou *et al.*, 2017 recently reported that polymorphism T152C in the Dloop region was associated with particularly AML-M3 subtype [11]. A previous study from our group reported that some of the mitochondrial D-loop variations were significantly associated with inferior survival in pediatric AML [12].

Our group recently published a descriptive analysis of mt-DNA D-loop variations among Indian children with AML [13]. However, the prognostic impact of somatic and germline mt-DNA D-loop variations in large cohort of AML subjects have never been reported. In view of these lacunae, the present study was conducted with the primary objective of assessing the impact of mitochondrial D-loop variations (somatic/germline) on outcomes in pediatric AML.

RESULTS

A total of 200 patients were registered at the centre during the study period. Out of 200 subjects, 49 patients were excluded from the study, (5 patients were acute promyelocytic leukemia, 15 patients had only one visit, 4 patients had an unsatisfactory buccal swab, 15 patients had failed sequencing and for 10 patient's the mother sample was not available). Therefore, 151 patients were eligible for the study.

In this sample of 151 eligible patients, median age was 10 years (0.7 to 18 years); male: female ratio was 2.5:1. Cytogenetics was evaluable in 125 patients (82.8%); 48.8% of the patients had good risk cytogenetics while intermediate and poor risk cytogenetics was present in 39.2% and 12.0% subjects respectively (Supplementary Table 1).

CR with induction therapy was achieved by 123 (81.5%) patients and 85 patients subsequently relapsed/ died after achieving remission. Median follow up was 33.8 months. The 2-year EFS (±SE) and OS was $32.3 \pm 3.9\%$ and $44.1 \pm 4.2\%$ respectively.

Relationship of baseline patients characteristics with EFS and OS

In univariate analysis, adverse cytogenetics risk group was significantly associated with inferior outcome whereas presence of chloromas was found to be significantly associated with better outcome (Table 1).

Mitochondrial D-Loop variations

All patients had one or more mitochondrial D-loop variations. A total of 1490 variations were identified at 237 positions in the D-Loop; 855 positions did not have variations. Thus, the mean number of variations per mutated position for the entire D-loop region was 6.3. Of these 1490 variations; 1206 (80.9%) variations were germline and 284 (19.1%) variations were somatic. One hundred and four, out of 237 positions in mtDNA had not been previously reported when compared with available databases updated to 17.12.17. All the variations have been submitted to the bankit (NCBI) at https://www.ncbi.nlm.nih.gov/WebSub/?tool=genbank vide accession number; GenBank "MG816363-MG816455".

Relationship of D-loop variations with EFS

Of the 237 positions with variations, 40 positions were affected with variations in ≥6 patients (mean frequency of variation in the D-loop region was 6.3). Univariate analysis revealed that among somatic variations, 8 positions (16111, 16126, 16189, 16209, 16278, 16304, 151 and 204) were found to be significantly associated with inferior EFS. Among germline variations, position 16126 was significantly associated with inferior EFS. On combining both somatic and germline variations (any variations), three positions (16111, 16126 and 482) were significantly associated with inferior EFS (Supplementary Table 2 and Supplementary Table 3).

D-loop variations, thus identified, were categorized into four different groups namely somatic protective, somatic hazardous, germline protective, germline hazardous (Table 2). Somatic protective group did not have any association with EFS; whereas the other three categories predicted EFS. Notably, in germline protective group the association was significantly seen when variations at ≥2 positions were observed (Table 2 and Supplementary Figure 1).

Relationship of D-loop variations with OS

Univariate analysis revealed that somatic variations at 7 positions viz. 16111, 16126, 16209, 16278, 16304, 151 and 204 had significantly inferior OS while germline variations at two sites, namely 16126 and 146, were significantly associated with inferior OS. When somatic and germline variations were combined (any variation), four positions viz 16111, 16126, 146 and 482 were significantly associated with inferior OS whereas variations at position 16318 was significantly associated with better OS (Supplementary Table 4 and Supplementary Table 5).

D-loop variations, thus identified, were categorized into four different groups namely somatic protective, somatic hazardous, germline protective, germline hazardous (Table 3). Somatic protective group did not have any association with OS; whereas the other three

Table 1: Association of baseline patients' characteristics with survival outcome (months)

Parameters (n = 151)	Event free survival			Overall Survival		
	Median	HR, (95%CI)	*P*	Median	HR, (95%CI)	*P*
Age (Years)						
≤10 (n = 77)	9.4	1.00		18.9	1.00	
≥10 (n = 74)	12.2	0.8 (0.6–1.3)	0.42	24.1	1.6 (0.9–2.7)	0.06
Sex						
Male (n = 108)	11.2	1.00		20.1	1.00	
Female (n = 43)	9.4	1.3 (0.8–1.9)	0.29	14.2	1.2 (0.7–1.8)	0.55
Hemoglobin (g/dl)						
≤8 (n = 82)	8.4	1.00		13.8	1.00	
≥8 (n = 69)	12.7	1.4 (0.9–2.1)	0.09	21.7	1.0 (0.6–1.7)	0.95
TLC (/mm³)						
≤11,000 (n = 55)	11.4	1.00		21.3	1.00	
≥11,000 (n = 96)	9.7	1.2 (0.7–1.8)	0.43	16.3	1.1 (0.7–1.8)	0.57
Platelets (/mm³)						
≤50,000 (n = 48)	10.9	1.00		21.3	1.00	
≥50,000 (n = 103)	10.2	1.1 (0.7–1.7)	0.65	19.9	1.0 (0.7–1.7)	0.84
CSF (n = 105)						
Negative (n = 92)	12.1	1.00		21.7	1.00	
Positive (n = 13)	9.3	1.5 (0.8–2.9)	0.18	18.9	1.3 (0.6–2.7)	0.44
Chloromas						
Negative (n = 126)	9.4	1.00		13.3	1.00	
Positive (n = 25)	24.0	0.4 (0.2–0.8)	0.01	24.0	0.4 (0.2–0.8)	0.01
Cytogenetic risk (n = 125)						
Favorable (n = 61)	12.7	1.00		24.0	1.00	
Intermediate (n = 49)	11.2	1.2 (0.7–1.9)	0.02	20.3	1.3 (0.8–2.3)	0.002
Adverse (n = 15)	5.9	2.2 (1.2–4.3)		8.6	3.1 (1.6–6.2)	
FLT3-ITD (n = 127)						
Negative (n = 117)	10.9	1.00		18.9	1.00	
Positive (n = 10)	5.5	1.4 (0.6–3.0)	0.39	6.4	1.8 (0.8–3.9)	0.16
NPM1 (n = 122)						
Negative (n = 113)	10.9	1.00		20.1	1.00	
Positive (n = 9)	8.2	1.4 (0.6–3.0)	0.41	11.7	1.4 (0.6–3.3)	0.43
*AML-ETO (n = 125)						
Negative (n = 76)	10.9	1.00		20.1	1.00	
Positive (n = 49)	11.4	0.9 (0.6–1.4)	0.68	21.8	0.8 (0.5–1.4)	0.48

Abbreviations: CI, Confidence interval; HR, Hazard ratio; TLC, Total leukocyte count; CSF, Cerebrospinal fluid; FLT3-ITD, *FMS* like tyrosine kinase-3 internal tandem duplication; NPM1, Nucleophosmin1; AML-ETO, Acute myeloid leukemia eight twenty one.
*Done by cytogenetics.

categories predicted OS. Notably, in germline protective group the association was significantly seen when variation at ≥2 positions were observed (Table 3 and Figure 1).

Multivariate analysis for outcome

On multivariate analysis, karyotype and any variation at positions 16111, 16126 and 16362 emerged as independent prognostic factors for poor EFS whereas in OS, karyotype along with any variations at positions 16111 and 16126 emerged as independent prognostic factors (Table 4).

Development of the prognostic model

The hazard coefficient for each of the above variables independently predictive of EFS was calculated. On the basis of these hazard coefficients, the prognostic index (PI) for each individual patient was calculated

using the following formula: PI = (0.8*karyotype) + (1.0*c16111) + (1.2*t16126). (0.7*t16362) (Table 4).

Based on this, median EFS of 3 different PI categories (PI = 0, PI ≥ 0 to ≤1 and PI ≥ 1) is shown in Figure 2. The median survival curves separate significantly ($p^* = 0.004$ and $p^{**} < 0.01$) with better EFS for patients with PI score ≤0 and the worst EFS for those with the PI score ≥ 1 (Figure 2). The model could not accurately predict the OS.

Impact of change in nucleotide

We analyzed our data in term of the quantum of change from one type of nucleotide to the other both in the somatic and germline categories separately and its impact on survival. Any somatic variation that involves a change into T from other nucleotide was noted in 64 patients and was associated with inferior OS ($p = 0.02$) (Supplementary

Table 6). On analyzing the quantitative burden of such changes, we observed that a higher burden (>2 positions) of such variations was significantly associated with inferior overall survival ($p = 0.02$) (Supplementary Table 6).

DISCUSSION

Variations in the mitochondrial DNA D-loop region have been observed in different solid malignancies and their incidence varies from 21% to 80% depending on the cancer type [13, 14–21]. Although, several studies have reported alterations in mt-DNA in hematological malignancies, studies elucidating the role of D-loop variations in AML and their prognostic significance are only a few [11, 12]. Studies focusing exclusively on AML suggest that these variations may play a role in the development, response to treatment and prognosis of AML [11, 12, 22].
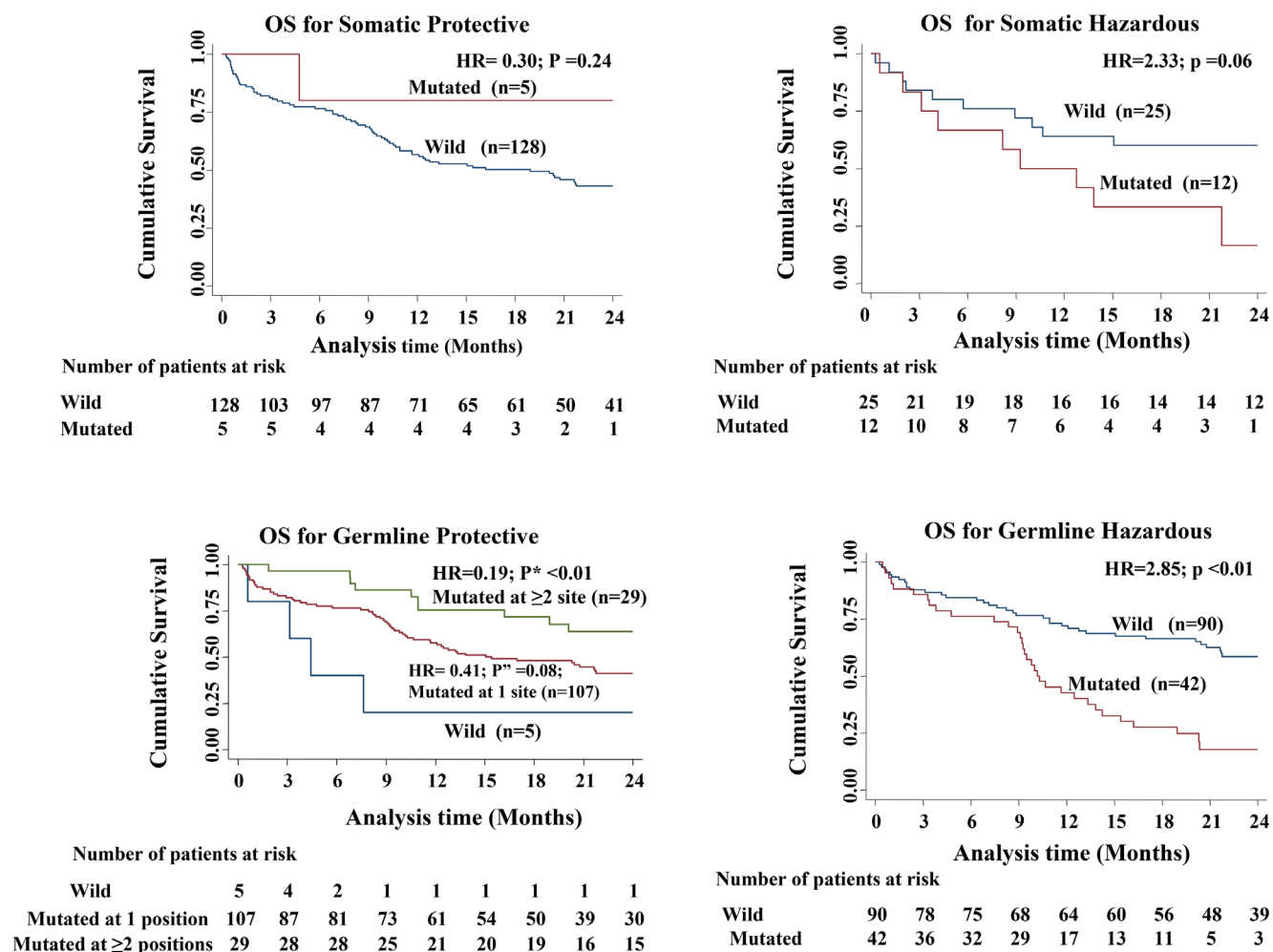

Figure 1: Kaplan–Meier curves comparing the overall survival among the four different categories of variations (*p* value shows the log rank of comparison between the survival curves of wild type patients versus those with variation and *p"* value shows the log rank of comparison between the survival curves of wild type patients versus those with variation at 1 position; *P** value shows the log rank of comparison between the survival curves of wild type patients versus those with variation at 2 positions).

Table 2: Event free survival of patients with variations into four categories of prognostic groups

Type of variations (prognostic groups)	Specific positions of significance	N	Median Survival Months	HR	P
Somatic Protective	16184, 150, 198, 489				
Wild		57	10.9	1.00	--
Variations		10	19.3	0.65	0.36
Somatic Hazardous	16051, 16093, 16111, 16126, 16189,				
Wild	16209, 16278, 16304, 16311, 16362,	26	15.4	1.00	--
Variations	16390, 151, 152, 204, 482	12	8.2	2.46	0.02
Germline Protective	16278, 16318, 73				
Wild		6	3.9	1.00	--
Variation at 1 position		118	10.3	0.59	0.25
Variation at ≥2 positions		18	15.4	0.31	0.03
Germline Hazardous	16126, 16172, 16327, 16390, 482				
Wild		113	12.7	1.00	--
Variations		29	5.9	2.27	≤0.01

Abbreviations: HR, Hazard ratio; N, Number of patients.

In each prognostic groups, EFS of patients with variations have been compared to wild type.

Table 3: Overall survival of patients with variations into four categories of prognostic groups

Type of variations (prognostic groups)	Specific positions of significance	N	Median Survival Months	HR	P
Somatic Protective	16184, 150, 198				
Wild		128	16.2	1.00	--
Variations		5	--	0.30	0.24
Somatic Hazardous	16051, 16093, 16111, 16126, 16192,				
Wild	16209, 16278, 16304, 16311, 16327,	25	--	1.00	--
Variations	16362, 16390, 151, 152, 204, 482	12	9.3	2.33	0.06
Germline Protective	16209, 16278, 16304, 16318, 73				
Wild		5	3.1	1.00	--
Variation at 1 position		107	15.1	0.41	0.08
Variation at ≥2 positions		29	--	0.19	≤0.01
Germline Hazardous	16126, 16172, 16327, 16390, 482				
Wild		90	--	1.00	--
Variations		42	10.2	2.85	≤0.01

Abbreviations: HR, Hazard ratio; N, Number of patients.

In each prognostic groups: OS of patients with variations have been compared to wild type.

Table 4: Factors associated with survival based on multivariate analysis

Variables in model	Event free survival			Overall survival		
	Hazard coefficient	HR (95%CI)	P	Hazard coefficient	HR (95%CI)	P
Cytogenetic risk (*others Vs adverse)	0.8	2.2 (1.2–4.2)	0.01	1.1	2.9 (1.5–5.7)	0.001
16111	1.0	2.8 (1.2–6.5)	0.02	1.4	4.1 (1.6–9.8)	0.002
16126	1.2	3.2 (1.6–6.7)	0.002	0.9	2.6 (1.2–5.7)	0.01
16362	0.7	1.9 (1.1–3.6)	0.02	0.4	1.6 (0.8–3.2)	0.18

Abbreviations: CI, Confidence interval; HR, Hazard ratio.

*Others - Favourable and intermediate ELN risk group.

The previous study from our group showed for the first time the impact of D-loop variations on EFS on pediatric AML; in that study, we identified three variations (16126, 16224 and 16311) in the hyper variable (HV1) region that predicted inferior EFS [12]. In the present study we went further ahead, classified the variations from a different set of 151 new subjects according to their germline and somatic status, and tried to identify unique prognostic groups.

In the present study, we could identify variations in 100% of patients as compared to 79% ($n = 44$) by Surender K. Sharawat et al., 2010 [12] and 60% ($n = 18$) by Yao et al., 2007 [5]. Our data, however, concurs with Zhou et al., 2017 [11] and Silkjaer et al., 2013 [22] who also reported variations in 100% of their AML subjects. Interestingly, Zhou et al., 2017 had sequenced only the D-loop region while Silkjaer et al., 2013 sequenced the whole mtDNA [11, 22]. Strikingly, none of these studies separately identified these variations as germline or somatic. Thus, the classification of variations as germline or somatic constitutes a novel information in this field.
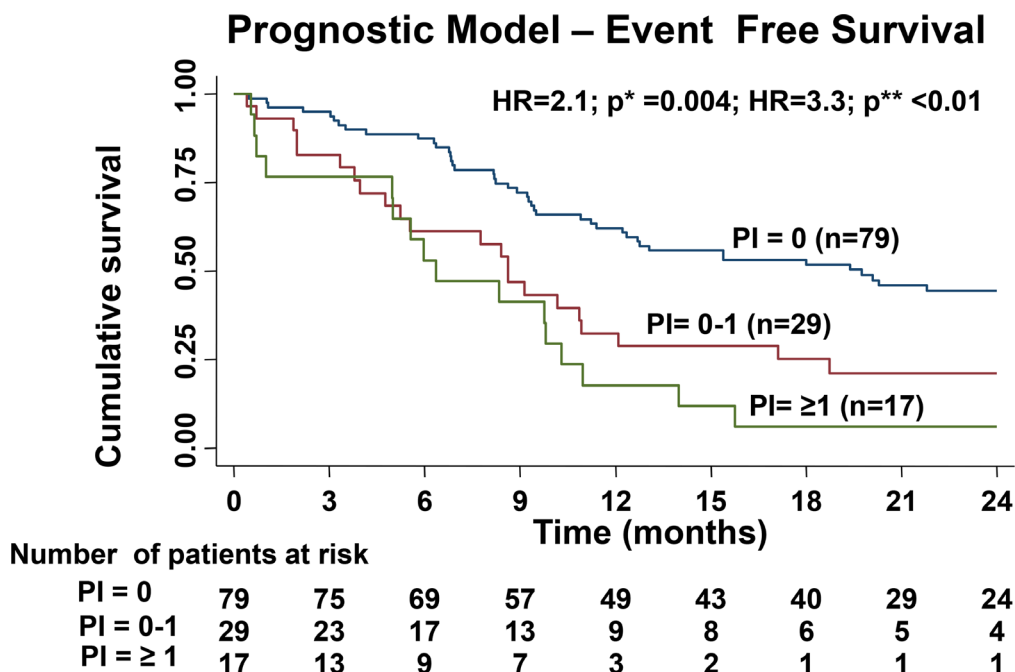
We observed that variations at certain positions, depending on their germline status, predicted a better or worse survival. It was even more interesting to note that the effect of germline protective variations was dependent on the number of variation positions in the individual. It appears as if both the characteristics and the burden of mt-DNA variations influence the phenotype of the disease and ultimately its biological behavior and survival outcomes. In the current era of precision medicine, tumor mutational burden (TMB) is being used as biomarker predicting response to immune-therapies in solid malignancies. In the recent phase-3 trial, immunotherapy was found to be significantly better than chemotherapy in the frontline treatment of metastatic non small cell lung cancer (NSCLC) with high TMB (>10 mutation per mega base) [23]. In this context, our observation that the burden of mt-DNA variations influences outcomes in AML appears highly significant and thought provoking.

We developed a prognostic model using the variations at three positions viz. 16111, 16126, 16362 and karyotype as prognostic variables and could demonstrate significant survival differences between patients with different prognostic indices. Although this predictive model needs to be validated in a larger cohort, nevertheless it is an important first attempt to integrate the novel biomarker in the prognostication of AML. Although, variations at 16111 and 16362 have not been reported to affect AML pathogenesis and prognosis in literature previously, their role has been reported in prostate cancer risk [24].

T16311C was reported to predict an inferior EFS by Surender K. Sharawat et al., 2010 [11] and OS (univariate analysis only) by Silkjaer et al., 2013 [22]. However, it was later found to be significantly associated with the favorable prognostic group comprising of APML, inv16 or t (8; 21) [22]. In our study, we could not demonstrate any significant association of these variations either with EFS or OS.

One of the limitations of our study is that we sequenced only D-loop region of the mitochondria in our subjects; studies also need to be done on the whole mitochondrial genome as other part of the mt-DNA

Prognostic Model – Event Free Survival

HR=2.1; $p^* =0.004$; HR=3.3; $p^{**} <0.01$

PI = 0 (n=79)

PI= 0-1 (n=29)

PI= ≥1 (n=17)

Number of patients at risk

	0	3	6	9	12	15	18	21	24
PI = 0	79	75	69	57	49	43	40	29	24
PI = 0-1	29	23	17	13	9	8	6	5	4
PI = ≥ 1	17	13	9	7	3	2	1	1	1

Figure 2: Difference in EFS of the three cohorts based on the prognostic index created by hazard coefficient of all the significant predictors for EFS in multivariable analysis (p^* value shows the comparison the survival curves prognostic index (PI) PI = 0 versus PI = 0–1 and p^{**} value shows the comparison between the survival curves of PI = 0 versus PI ≥ 1).

may also have a role in its pathogenesis. In term of methodology, we used PCR for amplification prior to sequencing which has its own inherent error. In order to increase the specificity of the reaction, we performed nested PCR and PCR products were sequenced in both forward and reverse direction. Further the prognostic model, which we developed using the variations at positions (16111, 16126 and 16362), needs to be validated in a larger independent cohort.

The strength of our study lies in the uniformity of the treatment protocol used and use of corresponding DNA from mothers to identify germline and somatic variations. Instead of bluntly separating out patients with or without variations, we categorized variations into four prognostically relevant groups namely somatic protective, somatic hazardous, germline protective and germline hazardous. The present paper is the first to report a prognostic model in AML integrating mt-DNA D-loop variations as prognostic biomarkers. This may be a useful tool to refine the prognostic stratification of AML, an area of unmet need.

MATERIALS AND METHODS

Patient's sample

This was a single centre, prospective cohort study conducted on *de novo* pediatric (≤18 year) patients with AML at the tertiary cancer centre at our institute, who were registered between June 2013 to August 2016. Study was approved by the institute ethics committee. Informed consent was taken from guardian of patients and assent form from those subject's ≤8 years of age. We excluded patients with acute promyelocytic leukemia (APML) and myelodysplastic syndromes (MDS) related AML. We sampled bone marrow (BM) and buccal swab from each patient and collected blood samples from the mother's of the patients. Patients whose mother's sample was not available were excluded from the study.

DNA extraction

Total DNA was isolated from 200 µl of patient's BM sample and patient's mother sample using the QI-Aamp DNA Mini-Kit (Qiagen, Hilden, Germany). DNA from buccal swab of the patients was isolated using G-biosciences DNA extraction kit. The quantity and quality of the DNA was assessed by spectrophotometer and agarose gel electrophoresis respectively.

Nested polymerase chain reaction (PCR)

The D-loop region was amplified by nested PCR to increase the specificity. All the primers were high performance liquid chromatography (HPLC) purified and details of the primers sequences are given below.

Primer	Sequence
Outer Primer Forward	5′-TCCACCATTAGCACCCAAAG-3′
Outer Primer Reverse	5′-GGGGATGCTTGCATGTGTA-3′
D1F	5′-TCCACCATTAGCACCCAAAG-3′
D1R	5′-GCTGTGCAGACATTCAATTGTT-3′
D2F	5′-GAGCTCTCCATGCATTTGGT-3′
D2R	5′-GGGGATGCTTGCATGTGTA-3′

PCR conditions for amplification for the whole D-loop region were: initial denaturation at 94ºC for 3 min; 35 cycles of denaturation at 94ºC for 45 s; annealing at 58ºC for 1 min; elongation at 74ºC for 30 s The PCR master mix contained 1.25 mmol/l of each dNTP (Fermentas, Glen Burnie, MD, USA), 20 pmol of each primer, 10 mmol/l Tris–HCl (pH 9), 50 mmol/l KCl, 1.5 mmol/l MgCl2 and Taq DNA polymerase (5 u/µl) (New England Biolabs, Ipswich, MA) in a total volume of 50 µl. PCR conditions for internal fragments of the D-loop were as follows: initial denaturation at 94ºC for 3 min; 35 cycles of denaturation at 94ºC for 30 s; annealing at 56ºC for 30 s; elongation at 72ºC for 45 s with a final extension at 72ºC for 8 min (Verti, Applied Biosystem; USA). PCR products were checked by resolving on a 2.0% agarose gel.

Nucleotide base sequencing and identification of variations

PCR product was eluted and purified from the gel using a column based PCR purification kit (Qiagen, Hilden, Germany). The purified product was quantified and sequenced using Sanger's sequencing (ABI 3730 XL; Applied Biosystems Instruments, Foster City, CA, USA). Obtained sequences and chromatograms were examined using DNA star software (Laser gene 14.1; USA) with reference sequence (NC_012920.1). Furthermore, all sequences were compared with Cambridge mitomap database (http://www.mitomap.org/bin/view/MITOMAP; updated on 17 Jan 2018) for already known polymorphisms and mutations. The available databases also include the polymorphisms and mutations observed in healthy individual in India that overlaps with those observed in other countries; however, there is no separate database for mitochondrial D-loop polymorphisms in Indian subjects [25, 26]. Therefore, we preferred to designate the D-loop sequence alterations from the database in our subjects as variations instead of mutations or polymorphisms.

Baseline cytogenetics and molecular analysis

Pretreatment samples from all patients underwent cytogenetics analysis at NABL accredited laboratory and chromosomal abnormalities were described according

to the International System for Human Cytogenetic Nomenclature (ISCN) [27]. Molecular analysis for FLT3-ITD and NPM1 were done using PCR [28, 29]. Based on this information, the European Leukemia Network (ELN) classification was used to categorize the patients into three different prognostic risk groups; good, intermediate and adverse risk [30].

Treatment

All patients were treated on a uniform protocol. They were induced with 3+7 regimen (daunorubicin 60 mg/m^2 for 3 days and cytosine arabinoside 100 mg/m^2 as a 24 hour continuous infusion for 7 days). Patients, who did not achieve complete remission (CR) after 1st induction, were given ADE (Cytarabine; 100 mg/m^2 BID day1–10 Daunorubicin; 50 mg/m^2 day1–3 and Etoposide; 100 mg/m^2 day1–3) as a 2nd induction [31]. CR was defined as bone marrow blast <5%, absolute neutrophil count >1000/uL, platelet count >100000/uL, no residual evidence of extramedullary disease and the child being independent of transfusion [32]. After achieving complete remission (CR), the patients received three cycles of high dose cytosine arabinoside at 18 g/m^2. Twenty eight patients received consolidation at 12g/m^2 as a part of a randomized controlled trial. Relapse following CR was defined as reappearance of leukemic blast in peripheral blood or the finding of >5% blasts in the bone marrow, not attributable to another cause [32]. Salvage chemotherapy was attempted for relapse and patients received allogeneic transplantation in CR2, if a matched sibling was available.

Sample size calculation

Our earlier work indicated that subjects with atleast 3 variations had higher mortality (HR = 2.03) although, statistically not significant (p = 0.20) [11]. To declare this excess of hazards of mortality significant in a 2-sided log rank test with 80% power and 5% alpha error, we required 45 subjects with atleast 3 D-loop variations and 45 cases with <3 variations. During the course of the study we realized that minimum variations per patient in our study were 4. Further, all variations did not behave similarly (some protective versus some hazardous), which prompted us to arbitrarily add 60 more patients.

Statistical analysis

Baseline characteristics for all the patients were summarized using median and range. Primary end point was overall survival (OS). OS was measured as the duration from the date of enrollment to death from any cause. Event free survival (EFS) was measured as the duration from date of enrollment to date of relapse or death due to disease. The data was censored at the date (22-December-2017) on last follow up for alive patients. The Kaplan–Meier statistics were used to estimate OS and EFS and log-rank test was used to compare differences between survival curves. P value ≤ 0.05 was considered significant. All statistical analysis was done using Stata 11.2.

Classification and categorization of variations

Germline variations were defined as those which were detected in all the three samples (patient's BM, buccal swab and mother's blood) whereas somatic variations were those which were present in patient's BM sample. Further to check the impact of these variations on EFS and OS, whether these variations are protective or hazardous, we performed the Cox proportional hazard analysis. We selected only those positions where variations were present in ≥6 patients as the mean frequency of variation in the D-loop was 6.3. We got 40 such positions and calculated the hazard ratio for EFS for each individual positions using Cox proportional hazard model and restricted mean survival time (RMST). All positions with hazard ratio ≤0.8 and p value ≤0.25 were considered protective. All positions with hazard ratio ≥1.2 and p value ≤0.25 were considered hazardous. Those with hazard ratio between ≤0.8 and ≥1.2 were not considered for analysis. The above values were taken because we wanted to include all position that showed a trend towards significance. We wanted to be inclusive at this stage and therefore relaxed the criteria. We then examined the effect of these four categories on survival. This time we used the Cox proportional hazard model and kept a strict threshold of p value ≤0.05 for determining statistical significance. Similar analysis was done on the overall survival.

Multivariate analysis

Variations in the D-Loop region of mt-DNA were compared with the wild type using Cox proportional hazards model for EFS and OS. All factors which were significant in the univariate analysis (p < 0.05) were considered for multivariate analysis. Stepwise multivariate Cox regression method was employed to evaluate the independent prognostic factors. Prognostic index (PI) of individual variations was calculated with hazard coefficient. Using this index, nomogram for predicting EFS was developed.

Author contributions

A.T. and S.B. designed the study. A.T., R.B., contribution to acquisition of data, interpretation of data. S.V. was the statistician, analyzed and interpreted the results. A.T., R.P., S.B. wrote the paper. All authors reviewed and gave the final approval for the paper.

REFERENCES

1. Copeland WC, Young MJ. Human mitochondrial DNA replication machinery and disease. Curr Opin Genet Dev. 2016; 38:52–62. https://doi.org/10.1016/j.gde.2016.03.005.

2. Alexeyev MF. Is there more to aging than mitochondrial DNA and reactive oxygen species? FEBS J. 2009; 276:5768–87. https://doi.org/10.1111/j.1742-4658.2009.07269.x.

3. Damm F, Bunke T, Thol F, Markus B, Wagner K, Göhring G, Schlegelberger B, Heil G, Reuter CW, Püllmann K, Schlenk RF, Döhner K, Heuser M, et al. Prognostic implications and molecular associations of NADH dehydrogenase subunit 4 (ND4) mutations in acute myeloid leukemia. Leukemia. 2011; 26:289–95. https://doi.org/10.1038/leu.2011.200.

4. He L, Luo L, Proctor SJ, Middleton PG, Blakely EL, Taylor RW, Turnbull DM. Somatic mitochondrial DNA mutations in adult-onset leukaemia. Leukemia. 2003; 17:2487–91. https://doi.org/10.1038/sj.leu.2403146.

5. Yao YG, Ogasawara Y, Kajigaya S, Molldrem JJ, Falcão RP, Pintão MC, McCoy JP Jr, Rizzatti EG, Young NS. Mitochondrial DNA sequence variation in single cells from leukemia patients. Leukemia. 2007; 109:756–62. https://doi.org/10.1182/blood-2006-01-011007.

6. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, et al. Sequence and organization of the human mitochondrial genome. Nature. 1981; 290:457–65. https://doi.org/10.1038/290457a0.

7. Gille P. Cell culture models for oxidative stress: superoxide and hydrogen peroxide versus normobaric hyperoxia. Mutat Res. 1992; 275:405–14. https://doi.org/10.1016/0921-8734(92)90043-O.

8. Lièvre A, Chapusot C, Bouvier A, Zinzindohoué F, Roignot P, Arnould L, Beaune P, Faivre J, Laurent-Puig P. Clinical Value of Mitochondrial Mutations in Colorectal Cancer. J Clin Oncol. 2019; 23:3517–25. https://doi.org/10.1200/JCO.2005.07.044.

9. Księżakowska-Łakoma K, Zyła M, Wilczyński JR. Mitochondrial dysfunction in cancer. Prz Menopauzalny. 2014; 18:136–44. https://doi.org/10.5114/pm.2014.42717.

10. Piccoli C, Ripoli M, Scrima R, Stanziale P, Di Ianni M, Moretti L, Biscottini B, Carella M, Boffoli D, Tabilio A, Capitanio N. MtDNA mutation associated with mitochondrial dysfunction in megakaryoblastic leukaemic cells. Leukemia. 2008; 22:1938–41. https://doi.org/10.1038/leu.2008.69.

11. Zhou J, Gou H, Ye Y, Zhou Y, Lu X, Ying B. Sequence variations of mitochondrial DNA D-loop region in patients with acute myeloid leukemia. Oncol Lett. 2017; 14:6269–6276. https://doi.org/10.3892/ol.2017.6988.

12. Sharawat SK, Bakhshi R, Vishnubhatla S, Bakhshi S. Mitochondrial D-loop variations in paediatric acute myeloid leukaemia: A potential prognostic marker. Br J Haematol. 2010; 149:391–8. https://doi.org/10.1111/j.1365-2141.2010.08084.x.

13. Tyagi A, Pramanik R, Vishnubhatla S, Ali S, Bakhshi R, Chopra A, Singh A, Bakhshi S. Pattern of mitochondrial D-loop variations and their relation with mitochondrial encoded genes in pediatric acute myeloid leukemia. Mutat Res. 2018; 810:13–8. https://doi.org/10.1016/j.mrfmmm.2018.05.002.

14. Fliss MS, Usadel H, Cobarello OL. Facile detection of mitochondrial DNA mutations in tumours and bodily fluids. Science. 2000; 287:2017–9. https://doi.org/10.1126/science.287.5460.2017.

15. Ivanova R, Lepage V, Loste MN, Schachter F, Wijnen E, Busson M, Cayuela JM, Sigaux F, Charron D. Mitochondrial DNA sequence variation in human leukemic cells. Int J Cancer. 1998; 76:495–8. https://doi.org/10.1002/(SICI)1097-0215(19980518)76:4<495::AID-IJC9>3.0.CO;2-M.

16. Guo XG, Guo QN. Mutations in the mitochondrial DNA D-Loop region occur frequently in human osteosarcoma. Cancer Lett. 2006; 239:151–5. https://doi.org/10.1016/j.canlet.2005.08.008.

17. Wulfert M, Küpper A, Tapprich C, Bottomley SS, Bowen D, Germing U, Haas R, Gattermann N. Analysis of mitochondrial DNA in 104 patients with myelodysplastic syndromes. Exp Hematol. 2008; 36:577–86. https://doi.org/10.1016/j.exphem.2008.01.004.

18. Tan DJ, Chang J, Liu LL, Bai RK, Wang YF, Yeh KT, Wong LJ. Significance of somatic mutations and content alteration of mitochondrial DNA in esophageal cancer. BMC Cancer. 2006; 6:93. https://doi.org/10.1186/1471-2407-6-93.

19. Nomoto S, Yamashita K, Koshikawa K, Nakao A. Mitochondrial D-loop mutations as clonal markers in multicentric hepatocellular carcinoma and plasma. Clin Cancer Res. 2002; 8:481–487. http://clincancerres.aacrjournals.org/content/8/2/481.short.

20. Grist S, Lu X, Morley A. Mitochondrial mutations in acute leukaemia. Leukemia. 2004; 18:1313–6. https://doi.org/10.1038/sj.leu.2403380.

21. Alonso A, Martin P, Albarran C, Aguilera B, Garcia O, Guzman A, Oliva H, Sancho M. Detection of somatic mutations in the mitochondrial DNA control region of

colorectal and gastric tumors by heteroduplex and single-strand conformation analysis. Electrophoresis. 1997; 18:682–5. https://doi.org/10.1002/elps.1150180504.

22. Silkjaer T, Nørgaard JM, Aggerholm A, Ebbesen LH, Kjeldsen E, Hokland P, Nyvold CG. Characterization and prognostic significance of mitochondrial DNA variations in acute myeloid leukemia. Eur J Haematol. 2013; 90:385–96. https://doi.org/10.1111/ejh.12090.

23. Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, Minenza E, Linardou H, Burgers S, Salman P, Borghaei H, Ramalingam SS, Brahmer J, et al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. N Engl J Med. 2018; 378:2093–104. https://doi.org/10.1056/NEJMoa1801946.

24. Navaglia F, Basso D, Fogar P, Sperti C, Greco E, Zambon CF, Stranges A, Falda A, Pizzi S, Parenti A, Pedrazzoli S, Plebani M. Mitochondrial DNA D-loop in pancreatic cancer: Somatic mutations are epiphenomena while the germline 16519T variant worsens metabolism and outcome. Am J Clin Pathol. 2006; 126:593–601. https://doi.org/10.1309/GQFCCJMH5KHNVX73.

25. Palanichamy MG, Sun C, Agrawal S, Bandelt HJ, Kong QP, Khan F, Wang CY, Chaudhuri TK, Palla V, Zhang YP. Phylogeny of Mitochondrial DNA Macrohaplogroup N in India, Based on Complete Sequencing: Implications for the Peopling of South Asia. Am J Hum Genet. 2004; 75:966–78. https://doi.org/10.1086/425871.

26. Thangaraj K, Chaubey G, Kivisild T, Selvi Rani D, Singh VK, Ismail T, Carvalho-Silva D, Metspalu M, Bhaskar LV, Reddy AG, Chandra S, Pande V, Prathap Naidu B, et al. Maternal footprints of Southeast Asians in North India. Hum Hered. 2008; 66:1–9. https://doi.org/10.1159/000114160.

27. Shaffer LG, McGowan-Jordan J, Schmid M. ISCN 2013: An International System for Human Cytogenetic Nomenclature. 2013. ISBN 978-331-8022-537.

28. Sharawat SK, Raina V, Kumar L, Sharma A, Bakhshi R, Vishnubhatla S, Gupta R, Bakhshi S. High fms-like tyrosine kinase-3 (FLT3) receptor surface expression predicts poor outcome in FLT3 internal tandem duplication (ITD) negative patients in adult acute myeloid leukaemia: A prospective pilot study from India. Indian J Med Res. 2016; 143:S11–6. https://doi.org/10.4103/0971-5916.191740.

29. Chopra A, Soni S, Pati H, Kumar D, Diwedi R, Verma D, Vishwakama G, Bakhshi S, Kumar S, Gogia A, Kumar R. Nucleophosmin mutation analysis in acute myeloid leukaemia: Immunohistochemistry as a surrogate for molecular techniques. Indian J Med Res. 2016; 143:763–8. https://doi.org/10.4103/0971-5916.192027.

30. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, Dombret H, Ebert BL, Fenaux P, Larson RA, Levine RL, Lo-Coco F, Naoe T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017; 129:424–48. https://doi.org/10.1182/blood-2016-08-733196.

31. Burnett AK, Russell NH, Hills RK, Kell J, Cavenagh J, Kjeldsen L, Mcmullin M, Cahalin P, Dennis M, Friis L, Thomas IF, Milligan D, Clark RE. A randomized comparison of daunorubicin 90 mg/m^2 vs 60 mg/m^2 in AML induction: results from the UK NCRI AML17 trial in 1206 patients. Blood. 2016; 125:3878–86. https://doi.org/10.1182/blood-2015-01-623447.

32. O'Donnell MR, Tallman MS, Abboud CN, Altman JK, Appelbaum FR, Arber DA, Bhatt V, Bixby D, Blum W, Gore SD, Hall AC, Kropf P, Lancet J, et al. Acute Myeloid Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2017; 15:926–57. https://doi.org/10.6004/jnccn.2017.0116.