FMUC FACULDADE DE MEDICINA
UNIVERSIDADE DE COIMBRA

MESTRADO INTEGRADO EM MEDICINA – TRABALHO FINAL

DUARTE GIL PACHECO SILVA

ALDEHYDE DEHYDROGENASE POLYMORPHISMS: THE ROLE IN MYELODYSPLASTIC SYNDROMES AND ACUTE MYELOID LEUKEMIA

ARTIGO CIENTÍFICO

ÁREA CIENTÍFICA DE BIOLOGIA MOLECULAR APLICADA/HEMATOLOGIA

Trabalho realizado sob a orientação de:
PROFESSORA DOUTORA ANA CRISTINA PEREIRA GONÇALVES
PROFESSORA DOUTORA ANA BELA SARMENTO ANTUNES DA CRUZ RIBEIRO

ABRIL 2019

ALDEHYDE DEHYDROGENASE POLYMORPHISMS: THE ROLE IN MYELODYSPLASTIC SYNDROMES AND ACUTE MYELOID LEUKEMIA

Duarte Gil Pacheco Silva[1,2]; Raquel Fernanda da Silva Alves[2,3,4]; Joana Margarida Verdasca Jorge[2,3,4]; Ana Bela Sarmento Antunes da Cruz Ribeiro[2,3,4,5]; Ana Cristina Pereira Gonçalves[2,3,4]

1. Medical Student, Faculty of Medicine of University of Coimbra (FMUC), Coimbra, Portugal
2. Laboratory of Oncobiology and Hematology and University Clinic of Hematology/Faculty of Medicine of University of Coimbra (FMUC), Coimbra, Portugal
3. Coimbra Institute for Clinical and Biomedical Research (iCBR) – Group of Environment, Genetics and Oncobiology (CIMAGO) - Faculty of Medicine of University of Coimbra (FMUC), Coimbra, Portugal
4. Center for Innovative Biomedicine and Biotechnology (CIBB), Coimbra, Portugal
5. Clinical Hematology Department/Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal

Duarte Gil Pacheco Silva
Azinhaga de Santa Comba, Celas
duarte.silva29@gmail.com

ABSTRACT

Aldehyde dehydrogenase (ALDH) is highly expressed in hematopoietic stem cells (HSC) and inhibition of ALDH promotes HSC self-renewal via reduction of retinoic acid activity. Since myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML) pathogenesis is owed to genetic and epigenetic changes in HSC, dysregulation of processes related with differentiation and cellular proliferation of these stem cells may be associated with the development of these hematological neoplasms. Thus, this study aims to analyze the role of ALDHs genetic variability in myeloid neoplasms (MN), namely in MDS and AML development and progression, particularly *ALDH1A2*, *ALDH3A1*, and *ALDH16A1* genes, in order to identify new potential risk factors and/or prognostic markers.

The selected ALDH single nucleotide polymorphisms (SNP) (*ALDH1A2*: rs4646626, *ALDH3A1*: rs2228100, *ALDH3A1*: rs887241 and *ALDH16A1*: rs1320303) were genotyped using tetra-primer-ARMS-PCR, in 99 MN patients (49 MDS patients and 50 AML patients), and in 118 control individuals. The role of SNPs in MDS and AML susceptibility and their association with clinical and laboratory characteristics was assessed by logistic regression analysis and/or by Fisher's exact test. All statistical analysis was two sided and a $p<0.05$ was considered statistically significant.

Our results show that *ALDH3A1* (rs2228100) CG genotype was associated with risk of MN and MDS development (MN: codominant model: OR=1.779, 95%CI 1.006-3.145, p=0.048; overdominant model: OR=2.072, 95%CI 1.197-3.585, p=0.009; MDS: overdominant model: OR=2.204, 95%CI 1.119-4.344, p=0.022). However, *ALDH3A1* (rs2228100) GG genotype conferred a protective effect against MN (recessive model: OR=0.296, 95%CI 0.105-0.828, p=0.02), having GG carriers 3.4x lower risk of MN development. Additionally, *ALDH3A1* CG and GT haplotypes were associated with MDS risk (CG: OR=1.901, 95%CI 1.181-3.062, p=0.0104; GT: OR=2.855, 95%CI 1.253-6.504, p=0.0200). However, CT haplotype conferred protection against MDS development (OR=0.444, 95%CI 0.243-0.812, p=0.0092), having CT haplotype carriers 2.25x lower risk of MDS development. *ALDH1A2* (rs4646626) GG genotype was associated with higher risk of MDS progression. In AML group, *ALDH1A2* (rs4646626) heterozygous carriers (AG) showed higher overall survival than homozygous carriers (AA + GG) (HR=0.513, 95%CI 0.273-0.966, p=0.035).

Altogether, these findings reinforce ALDH association with MN susceptibility. Because of its role in oxidative stress and mutagenesis, ALDH may play an important role in MDS development. Furthermore, these enzymes may also be important in AML prognosis and overall survival due to its importance in chemotherapy resistance and its association with stemness properties.

KEYWORDS

Aldehyde dehydrogenase; myelodysplastic syndrome; acute myeloid leukemia; polymorphism; genetic variability; risk factors; prognostic factors

RESUMO

A aldeído desidrogenase (ALDH) é uma enzima com expressão elevada nas células estaminais hematopoiéticas (HSC) e a sua inibição promove a autorrenovação das HSC através da redução da atividade da via do ácido retinóico. Uma vez que a patogénese da síndrome mielodisplásica (SMD) e da leucemia mieloide aguda (LMA) envolve alterações genéticas e epigenéticas nas HSC, a desregulação dos processos relacionados com a diferenciação e proliferação celular dessas células estaminais pode estar associada desenvolvimento destas neoplasias hematológicas. Neste contexto, o presente estudo teve como objetivo analisar melhor o papel da variabilidade genética das ALDHs, particularmente dos genes *ALDH1A2*, *ALDH3A1* e *ALDH16A1*, nas neoplasias mieloides (NM), particularmente no desenvolvimento da SMD e da LMA, de modo a identificar novos potenciais fatores de risco e/ou marcadores prognósticos.

Os polimorfismos de nucleotídeo único (SNP) selecionados (*ALDH1A2*: rs4646626, *ALDH3A1*: rs2228100, *ALDH3A1*: rs887241 e *ALDH16A1*: rs1320303) foram genotipados por tetra-primer-ARMS-PCR, em 99 doentes com NM (49 doentes com SMD e 50 doentes com LMA), assim como 118 indivíduos controlo. O papel dos SNPs na suscetibilidade da SMD e da LMA e a associação com características clínicas e laboratoriais foi avaliado por regressão logística e/ou pelo teste exato de Fisher. A análise estatística foi considerada estatisticamente significativo quando $p<0,05$.

Os nossos resultados mostram uma associação do genótipo CG da *ALDH3A1* (rs2228100) a maior risco de desenvolvimento NM e SMD (MN: modelo codominante: OR=1,777, IC 95% 1,006-3,145 , $p=0,048$, modelo sobredominante: OR=2,072, IC 95% 1,197-3,585, $p=0,009$; SMD: modelo sobredominante: OR=2,204, IC 95% 1,119-4,344, $p=0,022$). No entanto, o genótipo GG da *ALDH3A1* (rs2228100) apresenta efeito protetor (modelo recessivo: OR=0,296, IC 95% 0,105-0,828, $p=0,02$), tendo estes indivíduos 3,4x menos probabilidade de desenvolver de NM. Adicionalmente, os haplótipos CG e GT da *ALDH3A1* foram associados a risco de desenvolvimento de SMD (GC: OR=1,901, IC 95% 1,181-3,062, $p=0,0104$; GT: OR=2,855, IC 95% 1,253-6,504, $p=0,0200$). No entanto, o haplótipo CT confere proteção contra o desenvolvimento de SMD (OR=0,444, IC 95% 0,243-0,812, $p=0,0092$), tendo os portadores deste haplótipo 2,3x menos risco de desenvolvimento de SMD. Por outro lado, o genótipo GG da *ALDH1A2* (rs4646626) parece estar associado a maior risco de evolução de SMD para leucemia aguda. Os doentes com LMA, com o genótipo AG da *ALDH1A2* (rs4646626) apresentaram maior sobrevivência que os doentes com os genótipos AA e GG (HR=0,513, IC 95% 0,273-0,966, $p=0,035$).

Em suma, estes resultados reforçam a associação da ALDH com a suscetibilidade para NM. Devido ao seu papel no stresse oxidativo e na mutagénese, a ALDH parece desempenhar um papel importante no desenvolvimento da SMD. Além disso, as ALDHs podem ser também importantes no prognóstico e na sobrevivência da LMA devido ao seu papel na resistência à quimioterapia e sua associação com as propriedades estaminais.

PALAVRAS-CHAVE

Aldeído desidrogenase; síndrome mielodisplásica; leucemia mieloide aguda; polimorfismos; variabilidade genética; fatores de risco; fatores de prognóstico

ABBREVIATIONS

ALDH – Aldehyde dehydrogenase

AML – Acute myeloid leukemia

ARMS-PCR – Amplification-refractory mutation system-polymerase chain reaction

bp – Base pairs

CHUC, EPE – Centro Hospitalar e Universitário de Coimbra, EPE

CI – Confidence interval

CSC – Cancer stem cells

dbSNP – Dingle nucleotide polymorphisms database

DNA – Deoxyribonucleic acid

FI – Forward inner primer

FO – Forward outer primer

FMUC – Faculty of Medicine of University of Coimbra

HDFF, EPE – Hospital Distrital da Figueira da Foz, EPE

HR – Hazard ratio

HSC – Hematopoietic stem cells

HWE – Hardy-Weinberg equilibrium

IPSS – International Prognostic Scoring System

IPSS-R – Revised International Prognostic Scoring System

MAF – Minor allele frequency

MDS – Myelodysplastic syndromes

MDS-5q – Myelodysplastic syndrome with del(5q)

MDS-EB1 – Myelodysplastic syndrome with excess blasts type 1

MDS-EB2 – Myelodysplastic syndrome with excess blasts type 2

MDS-MLD – Myelodysplastic syndrome with multilineage dysplasia

MDS-RS – Myelodysplastic syndrome with ring sideroblasts

MDS-SLD – Myelodysplastic syndrome with single lineage dysplasia

MDS-U – Unclassified myelodysplastic syndrome

MN – Myeloid neoplasms

OS – Overall survival

OR – Odds ratio

RI – Reverse inner primer

RO – Reverse outer primer

ROS – Reactive oxygen species

SNP – Single nucleotide polymorphisms

WHO – World Health Organization

LIST OF ILLUSTRATIONS AND TABLES

Figure 1 – Regulation and function of ALDH in cancer stem cells (CSC) 7

Figure 2 – Representative examples of SNPs analysis by tetra-primer-ARMS-PCR 11

Figure 3 – Allelic distribution of ALDH1A2 (rs4646626), ALDH3A1 (rs2228100 and rs887241) and ALDH16A1 (rs1320303) polymorphisms in MDS, AML and controls 12

Figure 4 – Time to progression curves according to *ALDH3A1* (rs2228100) genotypes in MDS patients 15

Figure 5 – Overall survival curves in AML patients, according to *ALDH1A2* (rs4646626) genotypes 16

Table 1 – Basic demographic and clinical characteristics of myeloid neoplasms patients and controls 8

Table 2 – Selected SNPs details 9

Table 3 – PCR characteristics 10

Table 4 – Genotype distribution of selected SNPs in MN patients and controls, and their associated risk 13

Table 5 – Genotype distribution of selected SNPs in MDS and AML patients, and their associated risk 14

Table 6 – Haplotype distribution of *ALDH3A1* in MN and controls, and its association with risk of MDS and AML 15

BACKGROUND

Myelodysplastic syndromes (MDS) are a heterogeneous spectrum of clonal disorders that affect both hematopoietic stem cells and progenitor cells within the erythroid, megakaryocytic and/or granulocytic lineages. These diseases are characterized by ineffective hematopoiesis associated with peripheral blood cytopenias, myelodysplasia of one to all three myeloid cell lineages, genetic instability, and by its predisposition toward transformation into acute leukemia, essentially of the myeloid lineage.[1,2] Although molecular basis of MDS pathogenesis and progression remain unclear, a model of MDS molecular pathogenesis has been proposed whereby a normal hematopoietic stem cell acquires successive genetic abnormalities that ultimately lead to malignant transformation and clonal expansion.[2,3] In 1997, International Prognostic Scoring System (IPSS) was created, being an important standard for assessing prognosis of primary untreated adult MDS patients.[4] In order to become a more precise predict of patient outcomes, Revised International Prognostic Scoring System (IPSS-R) was created, later in 2012. Patient age, performance status, serum ferritin, and lactate dehydrogenase were significant additive features for survival but not for acute leukemia transformation.[5]

AML is a group of diseases characterized by infiltration of bone marrow, blood and other tissues by proliferative, clonal, abnormally differentiated hematopoietic cells, which leads to a series of fatal clinical problems.[6,7] Smith and collaborators referred AML pathogenesis as a complex process, pointing numerous mechanisms such as specific gene translocations or duplications, point mutations or larger deletions, as well as a variety of epigenetic changes leading to under- and overexpression of many genes involved in hematopoietic cell growth, differentiation, and self-renewal.[7] Analysis of NPM1, CEBPA, and, more importantly, FLT3 mutations has entered in clinical practice and affects diagnosis, risk assessment, and also guidance of therapy decision in AML. The evaluation of molecular genetic lesions as prognostic and predictive markers is an active research area.[3,8]

The aldehyde dehydrogenase (ALDH) superfamily is a group of NADP(+)-dependent enzymes, that participate in aldehyde metabolism, catalyzing exogenous and endogenous aldehydes oxidation.[9] In the human genome, 19 ALDH functional genes have been identified, with distinct chromosomal locations, and links between mutations in these genes and the molecular basis of several diseases.[3,10] ALDHs are involved in a broad spectrum of biosynthetic and metabolic processes and have a wide variety of functions, for instance catalytic (ester hydrolysis), binding and antioxidant (production of NADPH).[10] ALDH activity has been vastly studied in oncology and evidence proposes it as a universal cancer stem cell (CSC) marker.[11] ALDH stands out for its association with different types of solid tumors and its biological functions, including antioxidant protection and oxidation of retinal, resulting in an important role in developing tissues and in the self-renewal, differentiation, and self-protection of stem cells (Fig.1). Moreover, high levels of ALDH activity was associated with bad prognosis in several solid tumors, such as breast and prostate cancer, being not only a marker for aggressive stem-like and metastatic tumors, but also mechanistically involved in these behaviors.[12] However, validation of ALDH as a prognostic biomarker and/or therapeutic target in the clinical setting remain unclear. Hence, the study of ALDH as

a biomarker and functional mediator of metastasis *in vivo* is a promissory field for discovering targets that might interfere with solid tumor progression. Therefore, potential prognostic application involving the use of CSC prevalence in tumor tissue may require identification and quantification of specific ALDH isoforms.[12–14]

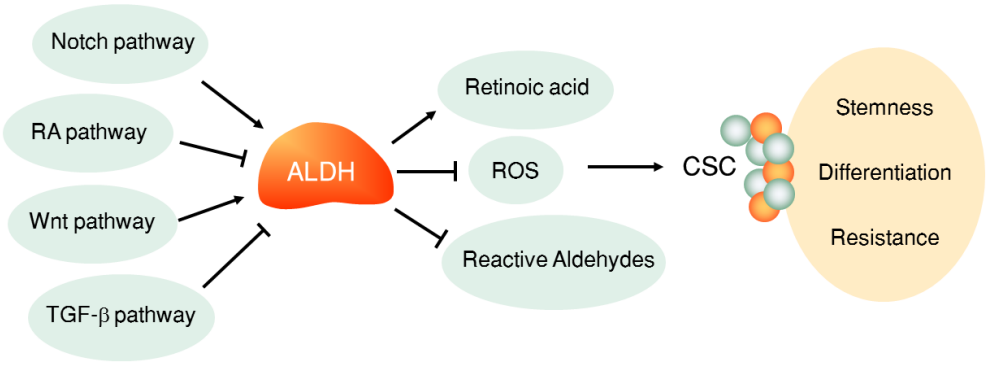

Figure 1 – Regulation and function of ALDH in cancer stem cells (CSC). ALDH metabolizes retinoic acid and reduces reactive oxygen species (ROS) and reactive Aldehydes, regulating CSC self-renewal, differentiation and resistance. CSC, cancer stem cells; RA, retinoic acid; ROS, reactive oxygen species; TGF-β; tumor growth factor β. (Adapted from X. Xu *et al*. 2015)

Genetic variations or polymorphisms existing in the human genome can confer genetic susceptibility to cancer, namely polymorphisms in genes involved in cell cycle control, carcinogen metabolism, DNA repair apoptosis, inflammation and epigenetic regulation.[15] Single nucleotide polymorphisms (SNP) are polymorphisms (which occur with a frequency of >1% in population) characterized by a change of a single nucleotide, that seem to be responsible for most of human interindividual genetic variability.[16] ALDHs interindividual genetic susceptibility, conferred primarily by SNP, may be particularly relevant in AML and MDS susceptibility, as well as influence their survival and risk of progression.

ALDH is also highly expressed in hematopoietic stem cells (HSC) and inhibition of ALDH activity promotes HSC self-renewal via reduction of retinoic acid activity.[17] As pathogenesis of MDS and AML is owed to several mutations in hematopoietic stem cells, dysregulation of processes related with differentiation and cellular proliferation of these cells may be associated with their development.[2,18] In addition to a possible role for ALDHs and their reactive aldehyde and reactive oxygen species (ROS) substrates in leukemic transformation, some of these dysregulated biological processes in these disorders may be dependent of ALDH activity.[19] In this context, it is reasonable to connect these enzymes with development these disorders

Thus, this study aims to analyze the involvement of ALDHs genetic variability in MDS and AML development, particularly *ALDH1A2*, *ALDH3A1*, and *ALDH16A1* genes, in order to identify new potential risk factors and/or prognostic markers.

MATERIAL AND METHODS

Ethical Statement

The present study was performed in accordance with the Helsinki declaration. Faculty of Medicine of University of Coimbra's (FMUC) ethics committee (Coimbra, Portugal) approved the study protocol and all patients and controls signed an informed consent.

Study Population

We conducted a hospital-based case-control study that included 99 myeloid neoplasia (MN) patients (49 MDS patients and 50 AML patients), as well as 118 control individuals (Table 1). Patients were diagnosed according to 2016 World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia.

Table 1 – Basic demographic and clinical characteristics of myeloid neoplasms patients and controls.

	Cases (*n* = 99)				Controls (*n* = 118)	
	MDS (*n* = 49)		AML (*n* = 50)			
Characteristics	*n*	%	*n*	%	*n*	%
Demographic data						
Gender						
Male	26	53.1	26	52.0	61	51.7
Female	23	46.9	24	48.0	57	48.3
Age (years)						
Median age	73		57.5		68	
Range	22-89		12-80		20-92	
Clinical data						
Myelodysplastic syndrome (*n*=47)						
MDS-RS	7	14.3	-	-		
MDS-SLD	7	14.3	-	-		
MDS-MLD	16	32.7	-	-		
MDS-5q	5	10.2	-	-		
MDS-EB1	3	6.1	-	-		
MDS-EB2	4	8.2	-	-		
MDS-U	7	14.3	-	-		
Acute myeloid leukemia (*n*=52)						
AML with recurrent genetic abnormalities	-	-	7	14		
AML with myelodysplasia-related changes	-	-	8	16		
Therapy-related myeloid neoplasms	-	-	3	6		
AML, NOS	-	-	32	64		

AML, acute myeloid leukemia; MDS, myelodysplastic syndrome; MDS-5q, myelodysplastic syndrome with del(5q); MDS-EB1, myelodysplastic syndrome with excess blasts type 1; MDS-EB2, myelodysplastic syndrome with excess blasts type 2; MDS-MLD, myelodysplastic syndrome with multilineage dysplasia; MDS-RS, myelodysplastic syndrome with ring sideroblasts; MDS-SLD, myelodysplastic syndrome with single lineage dysplasia; MDS-U, unclassified myelodysplastic syndrome; NOS, not otherwise specified.

All cases and controls were enrolled from two hospitals of Portugal's central region – Centro Hospitalar e Universitário de Coimbra, EPE (CHUC, EPE) and Hospital Distrital da Figueira da Foz, EPE (HDFF, EPE) – from January 2003 to September 2015. The follow-up information ended January 2019 and was updated via medical appointment or telephone contact. Controls were selected from healthy individuals, without evidence of neoplasms, age and gender matched, during the same period of time.

Genes and SNPs Selection

For this study realization, four ALDH genes were selected due to their role in stem cell signaling and their stablished relation with numerous neoplasms.[3,10]

The SNPs to be studied were selected based on minor allele frequency (MAF> 20%) in the Iberian population and its functional consequence (missense substitution). Thus, the following SNPs were selected: rs4646626 (*ALDH1A2*), rs2228100 (*ALDH3A1*), rs887241 (*ALDH3A1*), and rs1320303 (*ALDH16A1*) (Table 2).

Table 2 – Selected SNPs details.

Gene	dbSNP ID	Chr. position[a]	Variants	Consequence	MAF[b]
ALDH1A2	rs4646626	15:57963929	A>G	Missense	0.4439 (G)
ALDH3A1	rs2228100	17:19739639	C>G	Missense	0.2243 (G)
ALDH3A1	rs887241	17:19742625	G>T	Missense	0.3692 (T)
ALDH16A1	rs1320303	19:49461720	G>C	Missense	0.2290 (C)

ALDH, aldehyde dehydrogenase; Chr., chromosome; dbSNP ID, single nucleotide polymorphism database identification; MAF, minor allele frequency.

[a]According to GRCh38.p12

[b]MAF source: 1000 Genomes (Iberian population)

DNA Extraction and SNP Genotyping

DNA from cases and controls was extracted from blood samples through salting out method, as previously described by Miller, Dykes *et al*.[20] DNA was then quantified using NanoDrop ND-100 (NanoDrop Technologies, Wilmington, DE) and 100 ng of DNA was used in each genotyping assay. SNPs were genotyped using tetra-primer-ARMS-PCR with primers designed with BatchPrimer3 1.0 software (https://probes.pw.usda.gov/batchprimer3/). PCR conditions are described in Table 3. For each assay, three previously genotyped samples covering three possible genotypes were used as positive controls, along with water as a no template control (NTC). Electrophoresis was then performed on a 4% agarose gel and stained with Green Safe dye (Xpert Green DNA Stain, GRiSP, Lda., PT) (Fig.2).

Table 3 – PCR characteristics and conditions.

Gene	dbSNP ID	Primers (5' → 3')	PCR program	Buffer	Product
ALDH1A2	rs4646626	FI: AAGCGTGGAGCGGGCCAAGAGGCTCA RI: TGGTGGGGTCAAAGGGACTCCCCACGAC FO: CTATGCTGTGGAGCAGGCCCACCAGGGTG RO: TCCCTCTCTGGCTCCATTTCCAGCCACGAA	30'' 95ºC 30'' 63ºC x35 30'' 72ºC	MgCl2 2nM FI, FO, RI, RO 0.4μM	**G allele**: 164bp **A allele**: 229bp Outers: 339bp
ALDH3A1	rs2228100	FI: ATCCTCACGGACGTGGACCCCCAGTGCC RI: GGCCCGAAGATCTCCTCTTGCATCACGGC FO: TCCAAGCCTGGGCAGGTTTGCCTGGAAA RO: AACCCAGGTCTGTGGGCCCCAGGACCCT	30'' 95ºC 30'' 63ºC x35 30'' 72ºC	MgCl2 1.5nM FI, FO, RI, RO 0.2μM	**G allele**: 180bp **C allele**: 218bp Outers: 341bp
ALDH3A1	rs887241	FI: CTGGCTCGGTGTGCTCTGCAGGGAGCT RI: AGCTCCGAGGGCTTGAGGACCACCGC FO: CTCAGAGGGCCAGGTGGTGCAGCTCTGG RO: CACCCCCCAAAGACCTGGCCCTAGCTCA	30'' 95ºC 30'' 63ºC x35 30'' 72ºC	MgCl2 2nM FI, FO, RI, RO 0.2μM	**T allele**: 179bp **G allele**: 265bp Outers: 391bp
ALDH16A1	rs1320303	FI: GGCCCCTTCCCGGGAATCCTGAATGACG RI: GGCACCAGGGACGCAGGGCCACTCAG FO: CTGGGCCTTTGAGCTGCCCCACTTCCCC RO: CACGCCCTTCCTAGGGACCCCCCCCAGT	30'' 95ºC 30'' 60ºC x35 30'' 72ºC	MgCl2 2.5nM FI, RI 0.5μM FO, RO 0.3μM 10% DMSO	**C allele**: 157bp **G allele**: 202bp Outers: 305bp

ALDH, Aldehyde dehydrogenase; dbSNP ID, single nucleotide polymorphism database identification; bp, base pair; FI, forward inner primer; FO, forward outer primer; RI, reverse inner primer; RO, reverse outer primer; PCR, polymerase chain reaction

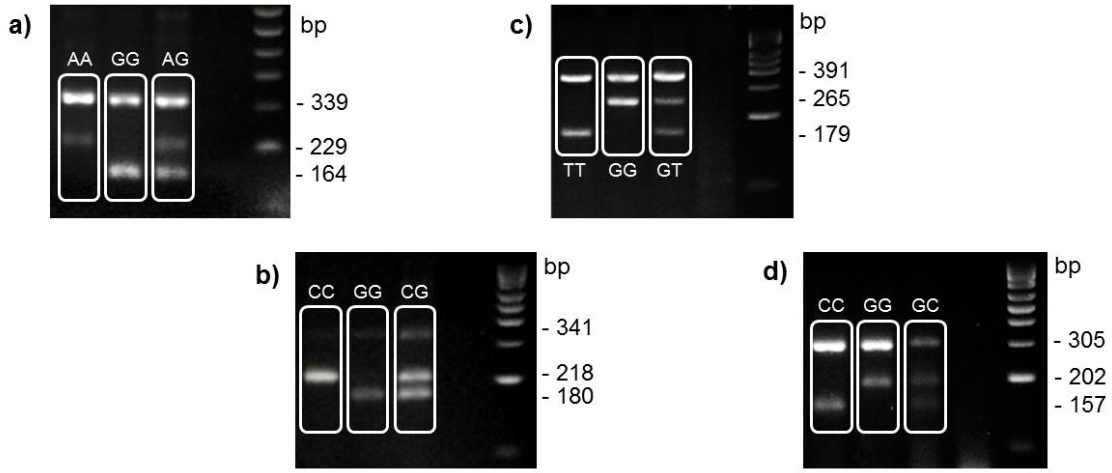

Figure 2 – Representative examples of SNPs analysis by tetra-primer-ARMS-PCR. *ALDH1A2* (rs4646626) (a), *ALDH3A1* (rs2228100) (b), *ALDH3A1*(rs887241) (c), and *ALDH16A1* (rs1320303) (d) genotypes and respective length.

Statistical Analysis

All statistical analysis was two-sided, using IBM SPSS Statistics 25 and Graphpad Prism 7.04. Age and gender differences between patients and controls were determined by Mann Whitney U test and Fisher's exact test, respectively. The Hardy-Weinberg (HWE) equilibrium of study population was determined by Arlequin software. Allelic and genotypic frequencies were determined by direct count. The analysis of genotype association was performed according to four genetic models by logistic regression: codominant model, dominant model, recessive model and overdominant model. Haplotypes and profiles were identified by Arlequin software, and risk analysis was performed by Fisher's exact test, calculating odds ratio (OR) and its 95% confidence interval (CI).[21] Survival of MDS and AML patients, as well as the analysis of MDS progression to acute leukemia according to genotype was determined using Kaplan-Meier's method and the hazard ratio (HR) calculated, with 95% CI. In all analysis, a value of $p<0.05$ was considered statistically significant.

RESULTS

Characteristics of study groups

This study enrolled a MN group (n=99) and a healthy control group (n=118). The MN group had a median age of 68 years (range 12-89), composed of 52 males (52.5%) and 47 females (47.5%). It included a MDS subgroup (n=49), with a median age of 73 years (range 22-89), of which 26 (53.1%) were males and 23 (46.9%) were females, and a AML subgroup (n=50), with a median age of 57.5 years (range 12-80), of which 26 (52%) were males and 24 (48%) were females. The healthy control group had a median age of 68 years (range 20-92) with of 61 males (51.7%) and 57 females (48.3%). In order to confirm adequate age and gender matching and avoid confounding bias, we assessed age and gender differences between groups. There were no significant differences between MN and controls groups in terms of age (p=0.41) or gender (p=1.00).

The MDS group included patients with the following subtypes, according to 2016 WHO classification: MDS with ring sideroblasts (MDS-RS: *n*=7, 14.3%), MDS with single lineage dysplasia (MDS-SLD: *n*=7, 14.3%), MDS with multilineage dysplasia (MDS-MLD: *n*=16, 32.7%), MDS with isolated del(5q) (MDS-del5q: *n*=5, 10.2%), MDS with excess blasts type 1 (MDS-EB1: *n*=3, 6.1%), MDS with excess blasts type 2 (MDS-EB2: *n*=4, 8.2%) and unclassifiable MDS (MDS-U: *n*=7, 14.3%). The IPSS was applied in 39 MDS patients. According to this score system, MDS group had 15 low risk patients (38.5%), 18 were intermediate-1 risk (46.2%), 5 intermediate-2 risk (12.8%) and 1 high risk patient (2.6%). IPSS-R was applied in 35 MDS patients, being 8 very low risk patients (22.9%), 14 low risk (40%), 8 intermediate risk (22.9%), 5 high risk (14.3%) and no very high risk patients. The AML subgroup included the following subgroups, according to 2016 WHO classification: 7 patients with AML with recurrent genetic abnormalities (14%) (AML with t(8;21)(q22;q22.1): *n*=1; AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22): *n*=2; APL with PML-RARA: *n*=4), 8 patients with AML with myelodysplasia-related changes (16.0%), 3 patients with therapy-related myeloid neoplasms (6.0%), and 32 patients with AML not otherwise specified (64.0%).

Allele and genotype distribution

The allele frequencies were assessed and then we studied the association with the selected SNPs and the risk of development of these MN. There were no differences in allele frequencies between MN patients and controls nor between MDS and AML patients and controls (Fig.3).

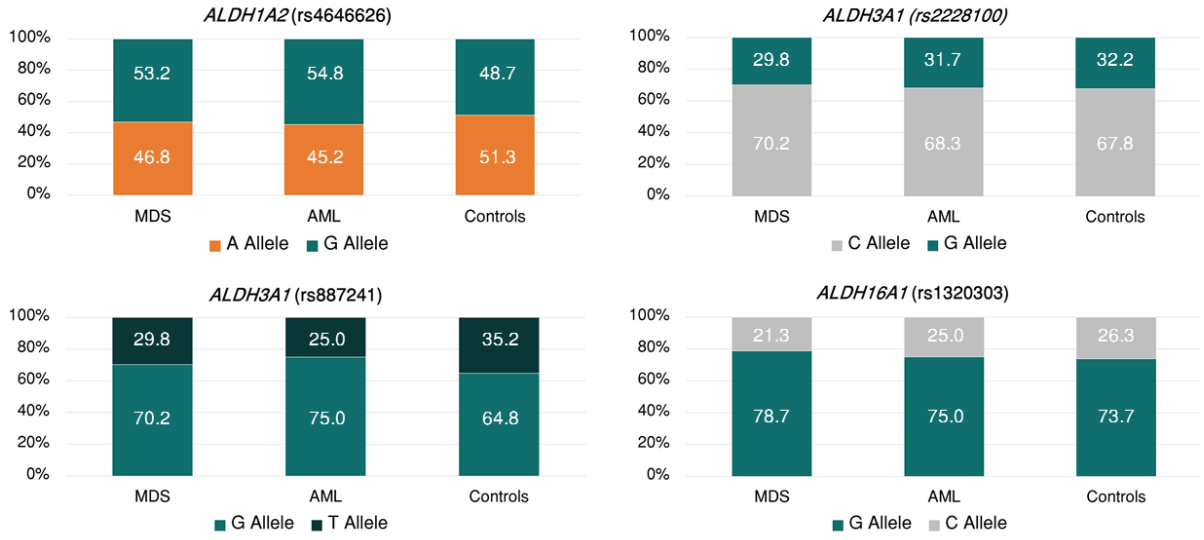

Figure 3 – Allelic distribution of *ALDH1A2* (rs4646626), *ALDH3A1* (rs2228100 and rs887241) and *ALDH16A1* (rs1320303) polymorphisms in MDS, AML and controls. AML, acute myeloid leukemia; MDS, myelodysplastic syndrome.

The genotype frequencies of selected SNPs were also compared and associations are shown in Table 4 and 5. All genotypes in study group were in HWE.

Table 4 – Genotype distribution of selected SNPs in MN patients and controls, and their associated risk.

	MN (*n*=99)				Controls (*n*=118)	
Gene: dbSNP ID	*n*	%	*p*-value	OR (95% CI)		
***ALDH1A2*: rs4646626**						
AA[a]	20	20.2	Ref.		28	23.7
AG[a]	51	51.5	0.787	1.098 (0.556-2.170)	65	55.1
GG[a]	28	28.3	0.263	1.568 (0.713-3.446)	25	21.2
AA[b]			0.533	0.841 (0.426-1.556)		
GG[c]			0.227	1.467 (0.788-2.731)		
AG[d]			0.600	0.866 (0.507-1.480)		
***ALDH3A1*: rs2228100**						
CC[a]	43	43.4	Ref.		60	50.8
CG[a]	51	51.5	**0.048**	**1.779 (1.006-3.145)**	40	33.9
GG[a]	5	5.1	0.081	0.388 (0.134-1.125)	18	15.3
CC[b]			0.277	0.742 (0.434-1.270)		
GG[c]			**0.020**	**0.296 (0.105-0.828)**		
CG[d]			**0.009**	**2.072 (1.197-3.145)**		
***ALDH3A1*: rs887241**						
GG[a]	53	53.5	Ref.		50	42.4
GT[a]	38	38.4	0.177	0.676 (0.383-1.194)	53	44.9
TT[a]	6	8.1	0.153	0.503 (0.196-1.289)	15	12.7
GG[b]			0.102	1.567 (0.915-2.683)		
TT[c]			0.273	0.604 (0.245-1.490		
GT[d]			0.332	0.764 (0.444-1.316)		
***ALDH16A1*: rs1320303**						
GG[a]	60	60.6	Ref.		61	51.7
GC[a]	32	32.3	0.105	0.626 (0.355-1.102)	52	44.1
CC[a]	7	7.1	0.565	1.423 (0.428-4.734)	5	4.2
GG[b]			0.189	1.438 (0.837-2.470)		
CC[c]			0.368	1.720 (0.528-5.597)		
GC[d]			0.078	0.606 (0.348-1.057)		

ALDH, aldehyde dehydrogenase; CI, confidence interval; dbSNP ID, single nucleotide polymorphism database identification; MN, myeloid neoplasms; OR, odds ratio; Ref., reference.
[a]Codominant model;
[b]Dominant model;
[c]Recessive model;
[d]Overdominant model.

The GG genotype of *ALDH3A1* (rs2228100) conferred a protective effect against MN (recessive model: OR=0.296, 95%CI 0.105-0.828, *p*=0.02), having the GG carriers 3.4x lower risk of MN development. However, the CG genotype of *ALDH3A1* (rs2228100) was associated with higher risk of MN predisposition (codominant model: OR=1.779, 95%CI 1.006-3.145, *p*=0.048; overdominant model: OR=2.072, 95%CI 1.197-3.585, *p*=0.009) (Table 4), in particular with a risk of MDS about 2-fold greater than the control population (overdominant model: OR=2.204, 95%CI 1.119-4.344, *p*=0.022) (Table 5). There were no significant differences between AML group and controls.

Table 5 – Genotype distribution of selected SNPs in MDS and AML patients, and their associated risk.

Gene: dbSNP ID	MDS (*n*=49)				AML (*n*=50)			
	n	%	*p*-value	OR (95% CI)	*n*	%	*p*-value	OR (95% CI)
***ALDH1A2*: rs4646626**								
AA[a]	11	22.4	Ref.		9	18.0	Ref.	
AG[a]	24	49.0	0.885	0.940 (0.406-2.177)	27	54.0	0.566	1.292 (0.539-3.100)
GG[a]	14	28.6	0.467	1.425 (0.548-3.709)	14	28.0	0.275	1.742 (0.643-4.717)
AA[b]			0.859	0.930 (0.421-2.058)			0.414	0.706 (0.306-1.629)
GG[c]			0.306	1.488 (0.695-3.185)			0.340	1.447 (0.677-3.090)
AG[d]			0.472	0.783 (0.402-1.526)			0.897	0.957 (0.493-1.859)
***ALDH3A1*: rs2228100**								
CC[a]	21	42.9	Ref.		22	44.0	Ref.	
CG[a]	26	53.1	0.083	1.857 (0.922-3.742)	25	50.0	0.135	1.705 (0.847-3.428)
GG[a]	2	4.1	0.145	0.317 (0.068-1.485)	3	6.0	0.240	0.455 (0.122-1.695)
CC[b]			0.348	0.725 (0.371-1.418)			0.417	0.760 (0.391-1.477)
GG[c]			0.060	0.236 (0.053-1.061)			0.110	0.355 (0.100-1.263)
CG[d]			**0.022**	**2.204 (1.119-4.344)**			0.052	1.950 (0.995-3.821)
***ALDH3A1*: rs887241**								
GG[a]	25	51.0	Ref.		28	56.0	Ref.	
GT[a]	19	38.8	0.359	0.717 (0.352-1.459)	19	38.0	0.211	0.640 (0.318-1.288)
TT[a]	5	10.2	0.478	0.667 (0.217-2.044)	3	6.0	0.127	0.357 (0.095-1.341)
GG[b]			0.307	1.417 (0.726-2.765)			0.107	1.731 (0.888-3.373)
TT[c]			0.650	0.780 (0.267-2.279)			0.209	0.438 (0.121-1.587)
GT[d]			0.466	0.777 (0.394-1.533)			0.434	0.767 (0.394-1.492)
***ALDH16A1*: rs1320303**								
GG[a]	32	65.3	Ref.		28	56	Ref.	
GC[a]	14	28.6	0.073	0.513 (0.248-1.064)	18	36	0.428	0.754 (0.375-1.516)
CC[a]	3	6.1	0.860	1.144 (0.257-5.095)	4	8	0.433	1.743 (0.435-6.989)
GG[b]			0.109	1.759 (0.882-3.508)			0.609	1.189 (0.612-2.312)
CC[c]			0.605	1.474 (0.338-6.422)			0.330	1.965 (0.505-7.647)
GC[d]			0.064	0.508 (0.247-1.041)			0.333	0.714 (0.361-1.412)

ALDH, aldehyde dehydrogenase; AML, acute myeloid leukemia; CI, confidence interval; dbSNP ID, single nucleotide polymorphism database identification; MDS, myelodysplastic syndrome; OR, odds ratio; Ref., reference.
[a]Codominant model;
[b]Dominant model;
[c]Recessive model;
[d]Overdominant model.

Regarding IPSS, in order to avoid sample dilution, we grouped patients in two groups: one group includes low risk and intermediate-1 risk patients (*n*=33, 84.6%) and other group the intermediate-2 risk with high risk patients (*n*=6, 15.4%). These risk groups were used in risk assessment but no significant differences in genotype distribution between these two risk groups were found.

Genotypic profile and haplotype analysis

Given the structural, functional and regulatory interaction of the majority of proteins, we identified the combination of the selected SNPs variants (genotypic profiles, GP) and assessed its associated risk with these MN. There were no significant associations between GPs and disease risk. However, the two *ALDH3A1* SNPs, rs2228100 and rs887241, are likely to be transmitted as a unit (haplotype). Therefore,

we assessed the association of *ALDH3A1* haplotype with disease development risk. Four possible haplotypes were identified: CG (haplotype 1), CT (haplotype 2), GG (haplotype 3) and GT (haplotype 4) (Table 6). Haplotypes 1 and 4 were associated with MDS risk (haplotype 1: OR=1.901, 95%CI 1.181-3.062, *p*=0.0104; haplotype 4: OR=2.855, 95%CI 1.253-6.504, *p*=0.0200). However, Haplotype 2 conferred a protective effect against MDS (OR=0.444, 95%CI 0.243-0.812, *p*=0.0092), and carriers of haplotype 2 have 2.25x less risk of MDS development. When we analyzed the correlation between genetic variations and the MDS risk score, IPSS, there were no significant differences in genotypic profiles nor haplotype distribution between the two risk groups mentioned previously.

Table 6 – Haplotype distribution of *ALDH3A1* in MN and controls, and its association with risk of MDS and AML.

	MDS			AML			Controls
Haplotype	%	p-value	OR (95% CI)	%	p-value	OR (95% CI)	%
***ALDH3A1*: rs2228100 + rs887241**							
H1: CG	53.1	**0.0104**	**1.90 (1.18-3.06)**	44.0	0.2725	1.32 (0.82-2.13)	37.3
H2: CT	16.3	**0.0092**	**0.44 (0.24-0.81)**	25.0	0.3573	0.76 (0.45-1.29)	30.5
H3: GG	17.3	0.0683	0.56 (0.31-1.02)	31.0	0.5084	1.21 (0.72-2.01)	27.1
H4: GT	13.3	**0.0200**	**2.86 (1.25-6.50)**	0.0	0.0210	0.09 (0.01-1.52)	5.1

ALDH, Aldehyde Dehydrogenase; AML, acute myeloid leukemia; CI, confidence interval; H1, haplotype 1; H2, haplotype 2; H3, haplotype 3; H4, haplotype 4; MDS, myelodysplastic syndrome; OR, odds ratio.

Prognostic impact

In order to assess the prognostic impact of selected SNPs, we estimated overall survival (OS) in MDS and AML and rate of MDS transformation into acute leukemia in patients whose follow-up information was available (MDS: *n*=40 and AML: *n*=48). We assessed MDS transformation into acute leukemia by evaluating, in MDS group, progression of disease and time to progression. The median time of progression was 43.7 months (range 1.7-195.2), being that 6 patients (15%) progressed. As observed in figure 4, MDS patients with *ALDH1A2* (rs4646626) GG genotype have higher risk of MDS progression.

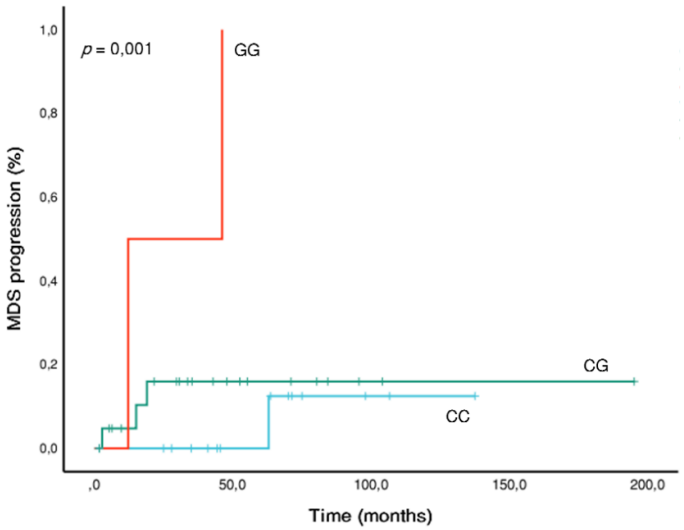

Figure 4 – Time to progression curves according to *ALDH3A1* (rs2228100) genotypes in MDS patients. Time to progression analysis was performed by Kaplan-Meier's method and differences were tested with log rank test.

Overall survival was evaluated in both MDS and AML using Kaplan-Meier's method. In MDS group, the median survival was 42.0 months (range 1.7-195.2) and, in AML group, the median survival was 9.2 months (range 0-124.8). There were no significant differences between selected SNPs genotypes and MDS overall survival. However, in AML group, *ALDH1A2* (rs4646626) heterozygous carriers (AG) showed higher overall survival that homozygous carriers (AA + GG) (HR=0.513, 95%CI 0.273-0.966, *p*=0.035), as shown in Figure 5.

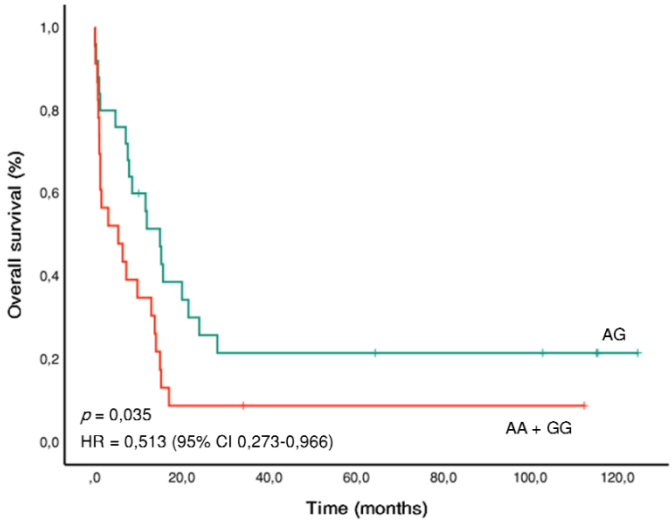

Figure 5 – Overall survival curves in AML patients, according to *ALDH1A2* (rs4646626) genotypes. Overall survival analysis was performed by Kaplan-Meier's method, differences were tested with log rank test, and hazard ratio (HR) with 95% confidence interval (CI) were calculated using Cox proportional hazard model.

DISCUSSION AND CONCLUSION

The role of ALDH has been studied in both normal and cancer stem cells and, to date, some studies have assessed its role in hematological malignancies. In this study, we were able to identify an association between *ALDH3A1* polymorphism and MN development, in particular MDS development. Additionally, we identified an association between *ALDH1A2* polymorphism and AML overall survival. We observed that *ALDH3A1* (rs2228100) GG genotype conferred protection against MN, but the CG genotype was associated with a higher risk of MDS development. We also observed *ALDH3A1* CG and GT haplotypes were associated with higher risk for developing MDS, however, CT haplotype was associated with less risk of MDS development. Many studies associated ALDH3A1 with breast and lung cancer, and others associated a decrease in ALDH3A1 and ALDH1A1 activity with a wide range of hematopoietic defects.[7,19,22–24] It has also been described that a deletion of both ALDH1A1 and ALDH3A1 could severely block B-cell development and reduce number of HSCs.[25] Furthermore, some studies described an association of ALDH1A1 and ALDH3A1, with drug resistance, namely to chemotherapeutic agents.[14,23] ALDH3A1 selective inhibition could increase chemosensitivity towards cyclophosphamide in ALDH3A1-expressing tumors, which could mean ALDH3A1 could have significance on AML overall survival.[7,23] However, we did not find significant results in the two studied *ALDH3A1* SNPs in overall survival. In addition, we found that *ALDH1A2* (rs4646626) AG genotype carriers showed a higher overall survival than homozygous carriers. It was described that, in human K562 leukemia cells, lentiviral gene transfer of *ALDH1A2* increased drug resistance to cyclophosphamide as well as to doxorubicin.[7] High expression of ALDH1A2 has also been associated with resistance to Ara-C.[23] The drug resistance associated with ALDH1A2 could explain its association with cancer overall survival, as it depends on the treatment used. However, our study did not assess the functional consequence of these SNPs, we cannot correlate our results with previous findings. Although we did not find significant results in ALDH16A1 SNP, this enzyme has been associated with higher risk for hyperuricemia and gout.[26]

There are few studies regarding ALDH specific isoforms association with cancer, however, many studies assessed its association with cancer pathogenesis and drug resistance. It has been described that ALDH1A1, ALDH1A2, and ALDH3A1 may be the main retinal dehydrogenases isoforms.[27] ALDH is the key enzyme in retinal oxidation to retinoic acid (RA), hence its important role in developing tissues and in self-renewal, differentiation, and self-protection of stem cells.[9] RA activates retinoic acid receptor (RAR) and retinoid X receptor (RXR), inducing transcription of target genes which, depending on cellular context, can lead to differentiation, apoptosis and cell arrest.[9,14,27]

Reactive aldehydes have also been studied in AML, described as being normal modifiers of signaling pathways and other biological processes in low concentrations, but at high concentrations it can damage DNA and proteins via adduct formation.[7] Reactive aldehydes are described by being more stable than ROS, diffusing to distant sites from where they were produced, propagating oxidative injury.[24] Moreover, it seems that a loss of ALDH activity can lead to reactive aldehydes-mediated increases in DNA adduct formation[7,24], and consequently may lead to neoplasm development.

Furthermore, besides being used as a HSC marker, it has been described that ALDH activity may also identify leukemic stem cells (LSC). In addition, $ALDH^+$ AML cells were described as having high levels of engraftment, quiescence, and increased long-term culture-initiating cell activity.[28] These cells were also associated with adverse cytogenetic abnormalities.[28] Thus, many studies suggest $ALDH^+$ AML patients as having a poorer outcomes due to either higher levels of LSCs or drug resistance.[22,24,25,28]

ALDH also participate in alcohol metabolism, and it is known that alcohol can be a cause of cytopenia, as well as its metabolite, acetaldehyde, can suppress neutrophil and macrophage function and increase intracellular ROS.[3,19] Thus, it is suggested that insufficient ALDH or DNA repair activity in HSCs may make HSCs more susceptible to accumulating DNA damage from ethanol-derived acetaldehyde, which may lead to mutagenesis, and higher susceptibility to MDS development.[19] In this context, life style choices should be accounted when analyzing the role of ALDH genetic variability in the development of cancer.

It has also been associated several SNPs in genes involved in oxidative stress with susceptibility of MDS and AML, MDS progression, as well as in survival of patients with MDS and AML.[29] This highlights the role that ALDH SNPs may have regarding oxidative stress and mutagenesis, as well as in MN pathogenesis and prognosis.

This work has some limitations, due to sample size and study design (hospital-based case-control). Besides sample size, it is important to note that MDS group was predominantly constituted by patients with low to intermediate risk, disenabling the correlation of ALDH genetic variability with the prognostic risk. Although sample size was reduced, allele frequencies observed in the study group were in HWE, suggesting that sampling was sufficiently random. Our results should be confirmed in other populations to exclude the probability of finding associations by chance.

It is vital to note that many ALDH related findings show different associations between various ALDH isoforms and several cancers, meaning that future studies regarding its mechanism in cancer pathogenesis may come down to specific isoform. Also, these various findings from different studies could be due to the different origins of ALDH. Thus, the study of ALDH role in hematological neoplasms seems important, not only for understanding MN pathogenesis, but also for a novel approach to treat a variety of cancers and for risk stratification. Introduction of SNPs in myeloid neoplasm prognostic systems may improve them, resulting in a better patient follow-up and earlier intervention. In this context, several studies enrolling a significant number of MDS and AML patients will be needed to confirm our results. There is also a need of functional studies, assessing these SNPs consequence, as well as correlating these genetic variants and respective ALDH expression and therapy response.

Altogether, these findings reinforce ALDH association with these MN susceptibility. Because of its role in oxidative stress and mutagenesis, ALDH may play an important role in MDS development. Furthermore, these enzymes may also be important in AML prognosis and overall survival due to its importance in chemotherapy resistance and its association with stemness properties.

ACKNOWLEDGMENTS

First, I thank Professor Ana Bela Sarmento-Ribeiro and Professor Ana Cristina Gonçalves, for all they taught me and for encouraging theorical and clinical discussion, teaching all that I needed to know.

To Joana Jorge and Raquel Alves, I thank all the support, for attending and answer all my newbie questions, for introducing me to all the techniques without which this work would not exist, and helping me until the final step of this work.

I thank my parents and brother for the unconditional support, despite the distance, not only throughout this work, but also throughout all past 6 years, unchangeable since day one. Also thanking my family, whose role was as equally vital to all my academic course.

Last, but not least, I thank all my friends, from Braga or Coimbra, besides listening to me rant over the technical and theorical aspects of this work, also turned this past 6 years easier to endure.

REFERENCES

1. Prebet T, Zeidan A. Trends in Clinical Investigation for Myelodysplastic Syndromes. Clin Lymphoma, Myeloma Leuk. 2016;16(No.S1):S57–63.
2. Look AT. Molecular Pathogenesis of MDS. 2005;156–60.
3. Macedo BBP. Aldehyde Dehydrogenases as potential biomarkers in Myeloid Neoplasias. Universidade de Aveiro. 2017;
4. Greenberg P, Cox C, LeBeau MM, Fenaux P, Morel P, Sanz G, Sanz M, Vallespi T, Hamblin T, Oscier D, Ohyashiki K, Toyama K, Aul C, Mufti G BJ, Greenberg P, Cox C, LeBeau MM, Fenaux P, Morel P, *et al*. International scoring system for evaluating prognosis in myelodysplastic syndromes. Blood. 1997;89(6):2079–88.
5. Yang Z, Zhang P, Tao K. Revised International Prognostic Scoring System for Myelodysplastic Syndromes. Blood. 2012;120(12):2454–65.
6. Döhner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. Longo DL, editor. N Engl J Med. 2015;373(12):1136–52.
7. Smith C, Gasparetto M, Humphries K, Pollyea DA, Vasiliou V, Jordan CT. Aldehyde dehydrogenases in acute myeloid leukemia. Ann N Y Acad Sci. 2014;1310(1):58–68.
8. Marcucci G, Haferlach T, Do H. Molecular Genetics of Adult Acute Myeloid Leukemia : Prognostic and Therapeutic Implications. J Clin Oncol. 2011;29(5):475–86.
9. Xu X, Chai S, Wang P, Zhang C, Yang Y, Yang Y, *et al*. Aldehyde dehydrogenases and cancer stem cells. Cancer Lett. 2015;369(1):50–7.
10. Marchitti SA, Brocker C, Stagos D, Vasiliou V. Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily. Expert Opin Drug Metab Toxicol. 2008;4(6):697–720.
11. Moreb J. Aldehyde Dehydrogenase as a Marker for Stem Cells. Curr Stem Cell Res Ther. 2008;3(4):237–46.
12. Rodriguez-Torres M, Allan AL. Aldehyde dehydrogenase as a marker and functional mediator of metastasis in solid tumors. Clin Exp Metastasis. 2016;33(1):97–113.
13. Clark DW, Palle K. Aldehyde dehydrogenases in cancer stem cells: potential as therapeutic targets. Ann Transl Med. 2016;4(24):518–25.
14. Marcato P, Dean CA, Giacomantonio CA, Lee PWK. Aldehyde dehydrogenase: Its role as a cancer stem cell marker comes down to the specific isoform. Cell Cycle. 2011;10(9):1378–84.
15. Tan IB, Ngeow J, Tan P. Role of Polymorphisms in Cancer Susceptibility. Encycl Life Sci. 2010;
16. Regateiro FJ. Manual de genética médica. 1ª ed. Imprensa da Universidade de Coimbra; 2003. 41-66 p.
17. Chute JP, Muramoto GG, Whitesides J, Colvin M, Safi R, Chao NJ, *et al*. Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. Proc Natl Acad Sci. 2006;103(31):11707–12.
18. Dohner H, Estey EH, Amadori S, Appelbaum FR, Buchner T, Burnett AK, *et al*. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert

panel, on behalf of the European LeukemiaNet. Blood. 2010;115(3):453–74.

19. Smith C, Gasparetto M, Jordan C, Pollyea DA, Vasiliou V. The Effects of Alcohol and Aldehyde Dehydrogenases on Disorders of Hematopoiesis. In: Advances in experimental medicine and biology. 2015. p. 349–59.
20. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215–1215.
21. Lewis CM. Genetic association studies: design, analysis and interpretation. Brief Bioinform. 2002;3(2):146–53.
22. Ran D, Schubert M, Pietsch L, Taubert I, Wuchter P, Eckstein V, *et al*. Aldehyde dehydrogenase activity among primary leukemia cells is associated with stem cell features and correlates with adverse clinical outcomes. Exp Hematol. 2009;37(12):1423–34.
23. Pors K, Moreb JS. Aldehyde dehydrogenases in cancer: An opportunity for biomarker and drug development? Drug Discov Today. 2014;19(12):1953–63.
24. Gasparetto M, Smith CA. ALDHs in normal and malignant hematopoietic cells: Potential new avenues for treatment of AML and other blood cancers. Chem Biol Interact. 2017;276:46–51.
25. Yang X, Yao R, Wang H. Update of ALDH as a Potential Biomarker and Therapeutic Target for AML. Biomed Res Int. 2018;2018:1–5.
26. Sulem P, Gudbjartsson DF, Walters GB, Helgadottir HT, Helgason A, Gudjonsson SA, *et al*. Identification of low-frequency variants associated with gout and serum uric acid levels. Nat Genet. 2011;43(11):1127–30.
27. Moreb JS, Ucar-Bilyeu DA, Khan A. Use of retinoic acid/aldehyde dehydrogenase pathway as potential targeted therapy against cancer stem cells. Cancer Chemother Pharmacol. 2017;79(2):295–301.
28. Blume R, Rempel E, Manta L, Saeed BR, Wang W, Raffel S, *et al*. The molecular signature of AML with increased ALDH activity suggests a stem cell origin. Leuk Lymphoma. 2018;59(9):2201–10.
29. Gonçalves AC, Alves R, Baldeiras I, Cortesão E, Carda JP, Branco CC, *et al*. Genetic variants involved in oxidative stress, base excision repair, DNA methylation, and folate metabolism pathways influence myeloid neoplasias susceptibility and prognosis. Mol Carcinog. 2017;56(1):130–48.