Leukemia Research Reports 9 (2018) 9–13

Contents lists available at ScienceDirect

Leukemia Research Reports

journal homepage: www.elsevier.com/locate/lrr

Expression of a specific extracellular matrix signature is a favorable prognostic factor in acute myeloid leukemia

Valerio Izzi[a,*], Juho Lakkala[a], Raman Devarajan[a], Eeva-Riitta Savolainen[b], Pirjo Koistinen[c], Ritva Heljasvaara[a,d], Taina Pihlajaniemi[a]

[a] Centre of Excellence in Cell-Extracellular Matrix Research and Biocenter Oulu, Faculty of Biochemistry and Molecular Medicine, University of Oulu, Oulu, Finland
[b] Nordlab Oulu and Institute of Diagnostics, Department of Clinical Chemistry, Oulu University Hospital, University of Oulu, Oulu, Finland
[c] Medical Research Center Oulu, Institute of Clinical Medicine, Oulu University Hospital, University of Oulu, Oulu, Finland
[d] Centre for Cancer Biomarkers (CCBIO), Department of Biomedicine, University of Bergen, N-5009 Bergen, Norway

ARTICLE INFO

Keywords:
Extracellular matrix (ECM)
Acute myeloid leukemia (AML)
Hematopoiesis
Clinical study
PCR

ABSTRACT

Relapse of acute myeloid leukemia (AML) is still dramatically frequent, imposing the need for early markers to quantify such risk. Recent evidence point to a prominent role for extracellular matrix (ECM) in AML, but its prognostic value has not yet been investigated. Here we have investigated whether the expression of a 15-ECM gene signature could be applied to clinical AML research evaluating a retrospective cohort of 61 AML patients and 12 healthy donors. Results show that patients whose ECM signature expression is at least twice as that of healthy donors have considerably longer relapse-free survival, with further stage-specific therapy outcomes.

1. Introduction

Although up to 80% of acute myeloid leukemia (AML) patients can expect to enter a first complete remission period (CR1) after appropriate induction regimen, many of them will subsequently relapse and face a dismal prognosis [1]. This adverse outcome is at the root of AML's still dramatically high death rate (approximately 21380 new AML cases will be diagnosed in USA in 2017, 50% of whom estimated to die within the same year- the highest death rate among hematological malignancies) [2], and the identification of new prognostic factors predisposing to either a better or a worse outcome is imperative to increase patients' chance to survive AML. Gene expression signatures have long since proven their potential usefulness in AML [3–6], but their translation to the clinics has been largely unsuccessful mainly because of the sophisticated methods they are based on (such as microarrays or specific chip platforms and RNA-seq) which are not readily available in clinical labs [7].

The extracellular matrix (ECM), the non-cellular microenvironment in which cells are embedded, plays crucial roles in both tissue homeostasis and disease [8]. In particular, in the hematopoietic niches, the ECM has key roles in anchoring hematopoietic stem cells (HSC) to the endosteal or the vascular structures, instructing the balance between proliferative and anti-proliferative signals and ultimately allowing for fine-tuning of the hematopoietic process throughout the life of the organism [8,9]. On the other hand, the ability of leukemia stem cells (LSC) and AML cells to interact with the ECM is a detrimental feature which generally fosters resistance to therapy and survival of minimal leukemic clones, which relapse in time and re-install the disease [10]. It is the case, in example, of CD44, the prototypical hyaluronic acid receptor with the further ability to bind to other ECM components (such as osteopontin, fibronectin and selectins) [11]. It has been reported, in fact, that CD44 expression on LSC and AML cells associates with resistance to chemotherapy and increased aggressiveness of the disease [11]. Much alike, integrin-mediated sensing of fibronectin determines post-therapy survival of AML clones, thus ultimately facilitating its relapse [12].

While many evidence can be found in the literature about the ability of both normal and neoplastic hematopoietic clones to sense, and to bind to, ECM, there is conversely a dramatic scarcity of knowledge on the production of ECM by AML cells themselves, which also implies an almost complete lack of knowledge on what roles direct ECM regulation by AML cells play in the context of biological and clinical features of AML.

Recently, we and others have reported on common and widespread mechanisms controlling the expression of extracellular matrix genes in AML and leukemic precursors [5,13], and shown the prognostic value of what we called the "extracellular matrix signature of AML" [5]. Also, we showed that machine-learning algorithms such as support vector machine (SVM) can reduce the 80-genes signature to a more practicable 15-genes signature (which can be assessed by real-time quantitative

* Correspondence to: Center for Cell-Matrix Research, Biocenter Oulu and Faculty of Biochemistry and Molecular Medicine, P.O. Box 5000, FIN-90014 University of Oulu, Finland.
E-mail address: valerio.izzi@oulu.fi (V. Izzi).

https://doi.org/10.1016/j.lrr.2017.12.001
Received 9 November 2017; Received in revised form 27 November 2017; Accepted 10 December 2017
Available online 13 December 2017

PCR – RT-qPCR) without losing sensitivity [5], but did not test whether this reduced signature could be applied to define patients' prognosis.

Combining the need for a better understanding of ECM roles in AML with the necessity of having tests that can be performed in clinical laboratories without the need for sophisticated methods and high-end mathematics, we have here addressed the question whether the restricted set of ECM genes which we previously identified [5] could provide relevant clinical information on AML patients, and found that the expression of this ECM signature at levels twice as that of healthy donors marked patients with a better response to therapy, reduced minimal residual disease (MRD) and overall longer relapse-free survival. We also observe that these findings, obtained using the simplest techniques currently in use in hematological laboratories worldwide, can be largely recapitulated in previously-published AML cohorts investigated via microarrays, further suggesting the importance of this signature in the biology and clinical features of AML.

2. Material and methods

2.1. Analysis of the Oulu AML retrospective cohort

The Oulu AML retrospective cohort was assembled with approval of the Institutional Review Board and informed written consent of the patients, in accordance with the declaration of Helsinki. Details about the 73 patients studied (61 AML + 12 healthy controls), as well as about the composition of the reduced ECM signature and the primers used for RT-qPCR are reported in the Appendix. The expression values of the 15 genes constituting the ECM signature (normalized to *GAPDH*) were collapsed to a single value per AML patient or healthy donor by calculating their geometric mean, using the formula:

$$\left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}} = \sqrt[n]{a_1 a_2 \cdots a_n}.$$

in which the geometrical mean is defined as the n^{th} root of the product of n elements a (n being the number of elements, in this case the genes - a). The arithmetic mean of all geometric mean values from the healthy donors was then calculated and the standard deviation value multiplied by 2 and then added to the average to obtain the upper and lower cutoff thresholds. All AML patients whose gene expression (geometric mean) fell within the thresholds were allocated to the ECM^{norm} group, while those whose expression was higher than the upper 2-SD threshold were allocated into the ECM^{high} group. In the Oulu cohort there were also 3 AML patients whose expression was lower than the bottom 2-SD threshold. Upon analysis, we found that these patients had no difference with the ECM^{norm} group, while showed exactly the same differences that the ECM^{norm} exhibited in respect to the ECM^{high} group. Hence, these patients were allocated back into the ECM^{norm} group.

For the analysis of outcome (post therapy)-specific results, patients were assessed at the following time-points: end of the induction protocol (Ind1), end of the first consolidation protocol (Cons1), and last available follow-up (Last).

2.2. Analysis of ECM signature expression in hematopoietic precursors

Raw microarray data (Affymetrix Human Genome U133 Plus 2.0 Array) were downloaded for the samples reported by Gentles et al. (GSE24006) [3] and by Novershtern et al., (GSE24759) [14], imported into Chipster (http://chipster.csc.fi/), normalized using robust multi-array average (RMA) protocol and the expression of the ECM signature studied. To facilitate cross-comparison with GSE24006, data from the GSE24759 were subset (post-normalization) to remove more mature cells, finally including only hematopoietic precursors ($CD133^+$ and $CD34^+$ HSC), committed progenitors (CMP, GMP and MEP), single-colony forming unit (CFU) progenitors (monocytic, granulocytic and megakaryocytic), and naïve B and T lymphocytes

2.3. Statistics

Fisher's Exact test (2-sided), Mann-Whitney U test, Analysis of Variance (ANOVA) followed by Tukey's HSD or Dunnett's T3 post-hoc tests, Kaplan-Meier (Log-Rank method, KM) and Cox proportional hazards (Cox-PH) survival analyses were performed in IBM SPSS Statistics 21, and all tests were bootstrapped 1000 times unless otherwise specified. Gene network enrichment analysis was performed in *String-DB* (http://string-db.org/) and the results imported into Cytoscape for easier visualization. The Linear Support Vector Machine (LSVM) algorithm used to analyze the contribution of the ECM gene expression to prognosis was trained and tested as reported in the Appendix, using IBM SPSS Modeler 18. In all analyses, a value of $P < 0.05$ was considered significant.

3. Results

3.1. Features of the ECM signature

The ECM signature which we tested in this work was previously reported [5] and comprises the following genes: *ADAM17*, *COL24A1*, *EMILIN2*, *CHI3L1*, *COL17A1*, *COL18A1*, *CRISP3*, *CRISPLD2*, *DEFA1*, *ELANE*, *LGALS3*, *MMP8*, *MMP9*, *PRTN3* and *SLPI*. This specific ECM signature is significantly enriched for protein-protein interactions (PPI) and includes ECM regulators (proteinases, 45% of the total gene-set), collagens (27%), glycoproteins (18%) and ECM-affiliated proteins (9%) (Fig. 1A) and overlaps with human AML signatures and mouse models of immunological and hematological phenotypes, which is an indication of the specific involvement of its constituents in the development (either normal or neoplastic) and functions of white blood cells (Fig. 1B,C and Appendix Table 1). Further ontological analyses of the signature are reported in Appendix Table 1.

Notably, signature expression is overall low in early hematopoietic stem and progenitor cells ($CD133^+$ and $CD34^+$ hematopoietic stem cells -HSC- and multipotent precursors -MPP), while it significantly increased with differentiation along the erythro-myeloid branch (myelo-erythroid progenitors -MEP-, common myeloid progenitors -CMP-, and granulocyte-monocyte progenitors -GMP-) and reached its maximum at the monocytic stage (CFU-mono) (Appendix Fig. 1**A,B**). In a similar way, the expression of the ECM signature in neoplastic clones was at its lowest in leukemia stem cells (LSC), while it increased constantly with more-differentiated cell states (leukemia precursor cells -LPC- and AML blasts) (Appendix Fig. 1B). Altogether, these results indicate that acquisition of this signature is globally associated with a more mature phenotype and, accordingly, we observed a significant negative association between signature expression and mRNA levels for CD34, a typical HSC and LSC marker [15], and a positive association with CD14, the phenotyping marker of monocytes [16].

3.2. Clinical significance of the ECM signature

Since this signature includes genes both up-and down-regulated in respect to healthy donors (Appendix Fig. 2) [5], and since relative expression values could not be collapsed into a single "global" value without using complicate approaches (such as principal component analysis) [3,6] unsuitable for direct clinical use, we undertook a

different approach, which separated AML patients into those who expressed the signature more than 2-times standard deviation (2-SD) that of healthy donors' expression and those whose expression was less than 2-SD that of healthy donors (see Supplemental Material for further details). All AML patients within the 2-SD limit were considered as "normal-like ECM" (ECM^{norm}), while patients outside these borders were considered significant outliers. Interestingly, we could not detect AML patients below the lower 2-SD threshold, but we could identify patients above the highest 2-SD thresholds, which we termed ECM^{high}. We found that ECM^{high} patients (in total 24 out of the 61 patients) had