Gene 591 (2016) 490–498

Contents lists available at ScienceDirect

Gene

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

Letter to the Editor

Does AL amyloidosis have a unique genomic profile? Gene expression profiling meta-analysis and literature overview

Fedor Kryukov [a,b,*], Elena Kryukova [a,b], Lucie Brozova [c,d], Zuzana Kufova [a,b], Jana Filipova [a], Katerina Growkova [a], Tereza Sevcikova [a,b], Jiri Jarkovsky [c,d], Roman Hajek [a,b]

[a] Department of Haematooncology, Faculty of Medicine, University of Ostrava, Czech Republic
[b] Department of Haematooncology, University Hospital Ostrava, Czech Republic
[c] Babak Myeloma Group, Department of Pathological Physiology, Faculty of Medicine, Masaryk University, Brno, Czech Republic
[d] Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic

ARTICLE INFO

Article history:
Received 15 March 2016
Received in revised form 2 June 2016
Accepted 4 June 2016
Available online 7 June 2016

Keywords:
AL amyloidosis
Monoclonal gammapaties
Gene expression profile
Ribosome
Meta-analysis

ABSTRACT

Immunoglobulin light chain amyloidosis (ALA) is a plasma cell dyscrasia characterized by deposition of amyloid fibrils in various organs and tissues. The current paper is devoted to clarify if ALA has a unique gene expression profile and to its pathogenetic argumentation.
The meta-analysis of ALA patients vs. healthy donors, monoclonal gammopathy of undetermined significance, smoldering and multiple myeloma patients' cohorts have revealed molecular signature of ALA consists of 256 genes representing mostly ribosomal proteins and immunoglobulin regions. This signature appears pathogenetically supported and elucidates for the first time the role of ribosome dysfunction in ALA.
In summary of our findings with literature overview, we hypothesize that ALA development is associated not only with changes in genes, coding amyloidogenic protein itself, but with post-transcriptional disbalance as well. Based on our data analysis in ALA, ribosome machinery is impaired and the affected link mainly involves translational initiation, elongation and co-translational protein folding.

1. Background

Immunoglobulin light chain amyloidosis (AL amyloidosis, ALA) is a plasma cell dyscrasia characterized by deposition of amyloid fibrils in various organs and tissues, derived from monoclonal immunoglobulin light chains (LC), leading to organ dysfunction (Rosenzweig and Landau, 2011). While the clonal light chain repertoire and the pathogenetic role of chromosomal aberrations (CAs) have been intensively studied in ALA, until now there is no systematic study of genomic expression of patients with AL amyloidosis compared to other monoclonal gammopathies. Meanwhile, molecular testing has become a standard in diagnostic evaluation to identify patient subgroups regarding prognosis (Bochtler et al., 2008).

AL amyloidosis may coexist with any of the other plasma cell dyscrasias (Kyle and Gertz, 1995; Rajkumar et al., 2006). It has been reported that approximately 10% of patients with AL may have multiple myeloma (MM) at the time of diagnosis (Kyle and Gertz, 1995), while only a minority will develop delayed MM (Rajkumar et al., 1998). Conversely, it has been reported that up to 30% of MM patients may have subclinical amyloid deposits (Desikan et al., 1997; Bahlis and Lazarus, 2006; Vela-Ojeda et al., 2009).

It is reasonably that many studies have been devoted to investigation of cytogenetic basement of AL amyloidosis, which is supposed to be similar to MGUS and MM. Finally, it was confirmed that all major chromosomal aberrations identified in MGUS were also detected in ALA group (Bochtler et al., 2008). However, the frequencies of aberrations in AL differ compared to MM and MGUS.

Based on the observation that hyperdiploidy (HY) is rare in AL compared with MGUS and MM and at the same time there is markedly higher prevalence of HY in AL with concomitant MM compared with AL without MM, it was suggested that the hyperdiploid karyotypes are biologically distinct and are manifested with a higher plasma cell burden (Bochtler et al., 2011).

Conversely, based on a higher frequency of t(11;14) in MGUS than in MM, it was proposed that t(11;14) is negatively selected for progression from an early-stage plasma cell disorder to symptomatic MM (Bochtler et al., 2011). The high frequency of t(11;14) in AL suggests that this translocation – though less aggressive – nevertheless sustains the

Abbreviations: ALA, immunoglobulin light chain amyloidosis; AL, amyloidosis; MM, multiple myeloma; MGUS, monoclonal gammopathy of undetermined significance; PC, plasma cell; SAM, Significant Analysis of Microarrays; PAM, prediction analysis of microarrays; MDS, multidimensional scaling; HD, healthy donors; SMM, smoldering myeloma; IRES, internal ribosome entry site; LC, light chains; PIs, proteasome inhibitors.

* Corresponding author at: Department of Haematooncology, Faculty of Medicine, University of Ostrava, Czech Republic.
E-mail address: f.kryukov@gmail.com (F. Kryukov).

http://dx.doi.org/10.1016/j.gene.2016.06.017
0378-1119/

proliferation of the aberrant plasma cell clone so that amyloidosis-related symptoms prompt the detection of this "more benign" gammopathy that otherwise may not have become symptomatic so soon (Durie, 1986; Hayman et al., 2001; Harrison et al., 2002; Bryce et al., 2009).

The t(11;14) and gain 11q can result cyclin D1 overexpression in multiple myeloma (Lesage et al., 2005). The high frequency of t(11;14) in AL suggests that overexpression of *CCND1* in ALA is likely more common than in myeloma. Nevertheless, overexpression of *CCND1* in the clonal plasma cells of patients with AL amyloidosis has been shown to occur even in cases without t(11;14) (Abraham et al., 2005). Transcriptional profiling identified that *CCND1* overexpression in ALA clonal plasma cells associated with significantly higher levels of endoplasmic reticulum protein processing genes such as *SEL1L*, *Sec63*, and *PDIA6* (Zhou et al., 2012).

Weinhold et al. conclude that MM and AL amyloidosis showing immunoglobulin light-chain amyloidosis shares genetic susceptibility with multiple myeloma based on similar association of the CCND1 c.870G4A and risk of t(11;14) MM and MGUS as well as t(11;14) AL amyloidosis (Weinhold et al., 2014).

Exome sequencing of plasma cells in systemic AL amyloidosis revealed 21 mutated genes in common with MM including *DIS3* and *NRAS* (Brian A Walker et al., Abstract on ASH 2013). Authors conclude that genetic signature of ALA is similar to the other plasma cell dyscrasia: it involves copy number variants (CNV), translocation and number of nonsynonymous mutations similar to MGUS but fewer to MM. It is worth to mention, that only 18 ALA patients participated in this pilot study.

There is an opinion that AL amyloidosis, MGUS and MM are hypothesized to be the same disease entity at the cellular level, with AL amyloidosis as being just a clonal plasma cell disorder with an "unlucky protein" (Hayman et al., 2001).

The comparison of gene expression profiles of clonal plasma cells (PC) revealed that AL had an intermediate transcript level in between of MM and healthy individuals (Abraham et al., 2005). This analysis suggests several mechanisms could be relevant for understanding the differences in the pathogenesis of ALA vs. MM.

One of these, *TNFRSF7*, a member of the tumor necrosis factor (TNF) receptor superfamily, which codes for CD27, a marker expressed on memory B cells (Colonna-Romano et al., 2003) has a higher average expression in AL PCs (Abraham et al., 2005). Surface CD27 plays role in differentiation of B cell into plasma cells and loss of its expression in MGUS plasma cells has been linked to MM progression (Guikema et al., 2003).

Another gene was the chemokine stromal cell-derived factor 1 (*SDF-1*), which is highly expressed in ALA PCs compared to MM PCs. However, SDF-1 levels in normal PCs are higher than in ALA (Abraham et al., 2005). SDF-1, also known as pre-B cell growth-stimulating factor, is produced by bone marrow stromal cells and was reported to act together with interleukin-7 as co-mitogen for pre-B cells (D'Apuzzo et al., 1997).

Lastly, authors point out a number of deregulated genes and pathways in ALA PCs that related to protein processing and folding including *APP*, *A2M*, *PSEN1* and *2*, *UCHL1*, and *CASP3* (Abraham et al., 2005).

In conclusion, authors suggest that exactly *TNFRSF7* and *SDF-1* and their interactions with other regulatory genes may help understand some of the differences between ALA and MM (Abraham et al., 2005).

Nevertheless, it is still unclear what conditions predetermine either one way of monoclonal gammopathy (MG) development or another. It is still not elucidated if aberrant plasma cell (PC) clone with special genetic background defines such clinical manifestation or this predisposition can be mostly defined by target tissue dysfunction with its biophysical, biochemical, and cellular properties. It is still a question if studies should focus on amyloidogenic light chains to get the picture of the role of somatic mutations in the disease process, on genetically divergent malignant PC clone that determine predisposition to ALA, or on pathologic conditions responsible for amyloid generation and deposition. The aim of this study was to systemize previously published data to highlight specific transcriptomic changes with potential pathogenic impact in AL amyloidosis.

2. Methods

2.1. Data preprocessing

Gene expression profiling data from GEO database (http://www.ncbi.nlm.nih.gov/geo/) were used for the analysis. From the accession GSE6477 (Chng et al., 2007), data of healthy donors ($n = 15$), MGUS ($n = 21$), smoldering MM ($n = 24$) and newly diagnosed MM ($n = 69$) were merged with the data of ALA patients ($n = 16$) from the accession GSE24128 (Zhou et al., 2012). Appropriate ethical statements have been published in source articles (Chng et al., 2007; Zhou et al., 2012). It is worth to mention that accession GSE24128 was the only one available ALA dataset in repositories supporting MIAME-compliant data (Brazma et al., 2001). Data were merged together using spot ID; spots without any annotation were deleted. Finally, patients were described by 13,209 genes represented by the median value of spot intensities. To ensure the comparability between arrays, quantile normalization was performed. At first, two principal components' difference between arrays from the two datasets was not observed (data not shown) which proves their comparability.

2.2. Statistical analysis

Unsupervised hierarchical clustering (complete linkage with Euclidean distance metric) and multidimensional scaling (MDS) were done on the whole set of genes to find out which diagnoses are more or less similar based on the overall gene expression profile. Only those genes with the most variable expression across all samples (having standard deviation above the median value) entered supervised analysis. Supervised analysis was performed applying Significant Analysis of Microarrays (SAM) (Tusher et al., 2001). The genes that satisfied the condition FDR < 0.001 and fold change more than 1.5 or vice versa less than 0.5 were assessed as significant. To extract biological meaning from revealed gene list, DAVID Bioinformatics Resources 6.7 was used (Huang da et al., 2009a,b). Functional Annotation Tool with "GOTERM_BP_5" and "GOTERM_CC_5" was applied for both up and down regulated genes, respectively. Obtained results were analyzed with REViGO to find a subset of the GO terms with common biological or cellular function (Supek et al., 2011). Using TreeMaps, biological and cellular processes were visualized proportionally according to their significance. To determine the smallest subset of genes that best distinguish between ALA and other diagnoses, prediction analysis of microarrays (PAM) using nearest shrunken centroids and cross-validation assay were done (Tibshirani et al., 2002). Analyses were performed in the R software (http://www.r-project.org/).

3. Results

Firstly, unsupervised analysis was performed to find out which diagnoses are more or less similar based on the overall gene expression profile. The results of hierarchical clustering based on Euclidian metric and complete linkage are shown in Fig. 1A. Two distinct groups of MM patients which were observed, can probably be explained with ploid category (hyperdiploid and nonhyperdiploid). This anticipation cannot be checked because of absence of required data in the used GEO datasets. Moreover, we have observed similarities between SMM and MM expression profile. MGUS gene expression profile was similar to the HD profile or SMM/MM profile but did not form a separate cluster.

The results from cluster analysis were confirmed using MDS (multidimensional scaling) method in Fig. 1B. It was observed that the healthy donors together with four MGUS samples form the most distinct unit.