ORIGINAL ARTICLE

Identification of pathway-based prognostic gene signatures in patients with multiple myeloma

MOHAMAD ZAMANI-AHMADMAHMUDI, SHAHREYAR DABIRI, and NADIA NADIMI

KERMAN, IRAN

Molecular profiling is used to extract prognostic gene signatures in different cancers such as multiple myeloma (MM), which is the second most common hematological malignancy. In this study, we utilized gene expression profiles to find biological pathways that could efficiently predict survival time in patients with MM. Four data sets—namely GSE2658 (559 samples), GSE9782 (264 samples), GSE6477 (147 samples), and GSE57317 (55 samples)—were employed. GSE2658 was used as a training data set and the others as validation data sets. The genes significantly associated with survival were identified using the univariate Cox proportional hazards analysis, and their roles in the biological pathways were explored using the Gene-Set Enrichment Analysis (GSEA) in the training data set. Next, the significant genes and their corresponding pathways were used to reconstruct pathway-based prognostic signatures. Thereafter, the significant gene signatures were externally validated in 3 independent cohorts—namely GSE9782, GSE6477, and GSE57317. Our results revealed that 9 pathway-based prognostic signatures were able to efficiently predict survival time in the training data set (*Ps* < 0.01). The testing of these signatures in the validation data sets demonstrated that 3 signatures—namely MYC targets, spliceosome, and metabolism of RNA—were able to strongly predict the clinical outcome in the 3 cohorts at *P* values < 0.01. In addition, in the multivariate Cox analysis, the 3 gene signatures remained as independent prognostic factors compared with the routine prognostic variables in MM—namely serum albumin, serum β2-microglobulin, and age. These signatures were by far the most powerful independent prognostic factors (MYC targets: $P = 0.009$, spliceosome: $P = 0.024$, and metabolism of RNA: $P < 0.001$). (Translational Research 2017;■:1–11)

Abbreviations: AREs = AU-rich elements; β2M = β2-microglobulin; GSs = gene expression signatures; GSEA = Gene-Set Enrichment Analysis; ISS = International Staging System; MM = multiple myeloma; MSigDB = Molecular Signatures Database; NF-κB = nuclear factor-kappa B; PI3K/Akt = phosphatidyl/inositol3-kinase/Akt; WNT = wingless

From the Department of Clinical Science, Faculty of Biomedical and Veterinary Medicine, Shahid Bahonar University of Kerman, Kerman, Iran; Department of Pathology, Faculty of Medicine, Kerman University of Medical Science, Kerman, Iran.

All authors have read the journal's policy on conflicts of interest, and all authors declared no conflicts of interest.

Submitted for publication January 25, 2017; revision submitted April 25, 2017; accepted for publication May 1, 2017.

Reprint requests: Mohamad Zamani-Ahmadmahmudi, Shahid Bahonar University, Faculty of Biomedical and Veterinary Medicine, Kerman 76169133, Iran; e-mail: zamani_2012@alumni.ut.ac.ir.

1931-5244/$ - see front matter

http://dx.doi.org/10.1016/j.trsl.2017.05.001

AT A GLANCE COMMENTARY

Zamani-Ahmadmahmudi M, et al.

Background

■■■.

Translational Significance

■■■.

INTRODUCTION

Multiple myeloma (MM) is the second most common hematological malignancy, with its worldwide incidence rates varying from 0.4 to 5 per 100,000 person-years.[1,2]

The most routine prognostic system in patients with MM is the International Staging System (ISS), which uses serum β2-microglobulin (β2M) and albumin as prognostic elements.[3] It has been stated that the ISS does not have enough efficacy to detect highest risk patients.[3] In addition, cytogenetic aberrations such as translocations t(4; 14) and t(14; 16) and the deletion of 17p [del(17p)] have been proposed as prognostic factors capable of identifying highest risk patients.[4] Combination of these systems (ie, ISS and genetics aberrations) can define more robust prognostic subtypes.[4]

With the advent of high-throughput technologies such as microarray and next-generation sequencing, the molecular classification of various cancers—including MM—has been extensively developed with a view to providing more precise subtypes with distinct prognosis. Through gene expression profiling, some prognostic gene expression signatures (GSs) have been proposed for MM.[5-10] For example, a 70-gene signature, developed by researchers at the University of Arkansas for Medical Sciences, indicated that the signature was associated with an increased risk of relapse and poor overall survival.[6] As another model, a 15-gene signature was developed by Decaux et al (2008) which could accurately predict survival in patients with newly diagnosed MM treated with high-dose therapy. In general, in these studies, a top-down approach was employed to create prognostic signatures, where the constituting elements (genes) in each signature were merely detected using computational analyses and proposed as a prognostic signature without considering the potential relationships of the genes in each signature which manifest in biological pathways. Indeed, many biological pathways containing various genes are dysregulated (activated or suppressed) in cancers such as MM. For example, IL-6R/STAT3, Ras/MAPK, phosphatidyl/inositol3-kinase/Akt (PI3K/Akt), notch, wingless (WNT), and nuclear factor-kappa B are examples of the activated pathways in MM.[11] When a pathway is activated or suppressed in a cancer, relevant genes in that pathway will majorly show similar expression patterns (upregulated or downregulated).[12-14] Hence, these series of the genes that act in the same cancer-associated pathway can be used as a prognostic gene signature.[12-15] Ignoring biological relationships between the genes in a prognostic gene signature can negatively affect the interpretation of the findings insofar as these prognostic gene signatures lack biological meaning.[16,17] To overcome these challenges and to reduce the difficulty in biological interpretation, researchers widely draw upon some tests such as gene set analysis that combine gene expression analysis and biological knowledge.[15,18] In the present study, we sought to fill this gap and find biological pathways which could robustly predict survival time in patients with MM. Hence, we explored gene expression profiling in 559 patients with MM (training set) to find pathways that were significantly associated with the patients' survival. We found 9 pathways that efficiently predicted prognosis in patients with MM. Thereafter, we tested the significant pathway-based gene signatures in 466 patients, comprising 3 cohorts (validation data sets). Our analysis revealed that 3 pathway-based gene signatures—namely MYC targets, spliceosome, and metabolism of RNA—were able to strongly predict the patients' outcome in the validation patients at a P value < 0.01.

METHODS

A schematic diagram depicting the analysis pipeline in our study is presented in Fig 1.

Training and validation data sets. Gene expression profiling data sets—namely GSE2658, GSE9782, GSE6477, and GSE57317—were downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/). GSE2658 comprised 559 samples of patients with MM and was used as the training data set. GSE9782, GSE6477, and GSE57317—as validations data sets—contained 264, 147, and 55 samples, respectively. In all studies (data sets), samples were obtained from bone marrow and enriched for plasma cells (CD138 + cells) using immunomagnetic bead separation. Furthermore, in all studies, authors confirmed that all patients had provided written informed consent to sample procurement, in accordance with the Declaration of Helsinki. The data sets were downloaded in