[Elucidation of the UPR signalling pathway mediated by IRE1 protein in cancer using bioinformatics]

Abstract

Proteostasis imbalance is emerging as a major hallmark of cancer, driving tumor aggressiveness. Evidence suggests that the endoplasmic reticulum (ER), a major site for protein folding and quality control, plays a critical role in cancer development. This concept is valid in glioblastoma multiform (GBM), the most lethal primary brain cancer with no effective treatment. Previous studies demonstrated that the ER stress sensor IRE1α (also referred to as IRE1) contributes to GBM progression, through XBP1 mRNA splicing and regulated IRE1-dependent decay (RIDD) of RNA. Here, we first demonstrated IRE1 signaling significance to human GBM and defined specific IRE1-dependent gene expression signature of two main components with 19 genes each, reflecting the XBP1s and RIDD signaling activity, respectively, that were confronted to human GBM transcriptomes. The activity assessment of both signaling axes and their correlation with biochemical, clinical, immunological and phenotype characteristics allowed us to demonstrate the antagonistic roles of XBP1 mRNA splicing and RIDD on tumor outcomes, mainly through selective remodeling of the tumor stroma. Moreover, we mapped the miRNA network involved in the regulation of the two signaling axes for the first time and identified 11 candidate miRNAs that control IRE1 signaling activity in both GBM and Triple-Negative Breast Cancer (TNBC) by controlling biological processes involved in the infiltration/invasion, migration and metastasis of tumor cells and the cancer stem cell reprogramming as well. Next, we deciphered the nucleotide sequence and RNA secondary structure motifs of RIDD targets and we constructed *in silico* two scaffold RNA sequences (probes) folding in stem-loop and internal loop, respectively, after the encapsulation of identified IRE1 cleavage site motifs in the loops' region. Based on the RNA probes constructed *in silico*, RNA oligos were synthesized biochemically, and were validated as RIDD targets by an IRE1-mediated cleavage assay. In parallel, we developed a bioinformatics methodology for the identification of cleavage site motifs in the sequences of hairpin and internal loops in the precursor miRNAs of human miRBase, highlighting candidate miRIDD targets as well. The findings of RNA motif sequence and structure analysis helped us to develop a new methodology for the identification and quantification of cleavage site motifs in RNAseq data. Moreover, the computational methodologies that were utilized throughout the GBM for the evaluation of IRE1 activity was applied to other 11 solid tumors of Cancer Genome Atlas (TCGA) and their patients were grouped into high and low IRE1/XBP1s/RIDD molecular subtypes with a direct impact on patient survival rates, while significant correlation

was demonstrated with specific molecular, clinical and immunological markers. This study provides the first demonstration of a dual role of IRE1 downstream signaling in cancer and opens a new therapeutic window to abrogate tumor progression. The IRE1α-centric landscape of tumor microenvironment is summarized in the "Synopsis" figure which highlights the pivotal role of IRE1 protein in cancer.

Finally, for the sake of translational research and precision medicine we developed a GUI (graphical user interface) R Shiny platform that allows users to i) retrieve molecular and clinical TCGA data, ii) process and filter molecular data by increasing the ratio between signal to noise, iii) analyze and integrate NGS data from different platforms by using cross-platform harmonization, iv) derive gene signatures with potential biomarkers that characterize the phenotype under study, v) perform differential expression analysis and functional enrichment analysis, vi) stratify cancer patients based on the gene signature expression profile and vii) visualize their results with a series of sophisticated ways including hierarchical clustering heatmaps, PCA plots, survival plots, oncoplots, Volcano plots and gene network representations.

Keywords: *Endoplasmic Reticulum (E.R.), UPR, Cancer, IRE1, XBP1, RIDD, GBM, TCGA, motifs, miRNAs, RNA secondary structure, IRE1sign38, R Shiny.*