Video Article

Using RNA-sequencing to Detect Novel Splice Variants Related to Drug Resistance in *In Vitro* Cancer Models

Rocco Sciarrillo[1,2,3], Anna Wojtuszkiewicz[1], Irsan E. Kooi[4], Valentina E. Gómez[3], Ugo Boggi[5], Gerrit Jansen[6], Gert-Jan Kaspers[1,7], Jacqueline Cloos*[1], Elisa Giovannetti*[3,8,9]

[1]Department of Pediatric Oncology/Hematology, VU University Medical Center

[2]Department of Hematology, VU University Medical Center

[3]Department of Medical Oncology, VU University Medical Center

[4]Department of Clinical Genetics, VU University Medical Center

[5]Division of General and Transplant Surgery, Azienda Ospedaliera Universitaria Pisana, Universita' di Pisa

[6]Amsterdam Immunology and Rheumatology Center, VU University Medical Center

[7]Princess Máxima Center for Pediatric Oncology

[8]Cancer Pharmacology Lab, AIRC Start-Up Unit, University of Pisa

[9]Institute of Nanoscience and Nanotechnology, CNR-Nano

*These authors contributed equally

Correspondence to: Elisa Giovannetti at e.giovannetti@vumc.nl

URL: http://www.jove.com/video/54714
DOI: doi:10.3791/54714

Keywords: Cancer Research, Issue 118, chemotherapy, resistance, cytotoxicity, alternative splicing, RNA-seq, transcriptomics

Date Published: 12/9/2016

Citation: Sciarrillo, R., Wojtuszkiewicz, A., Kooi, I.E., Gómez, V.E., Boggi, U., Jansen, G., Kaspers, G.J., Cloos, J., Giovannetti, E. Using RNA-sequencing to Detect Novel Splice Variants Related to Drug Resistance in *In Vitro* Cancer Models. *J. Vis. Exp.* (118), e54714, doi:10.3791/54714 (2016).

Abstract

Drug resistance remains a major problem in the treatment of cancer for both hematological malignancies and solid tumors. Intrinsic or acquired resistance can be caused by a range of mechanisms, including increased drug elimination, decreased drug uptake, drug inactivation and alterations of drug targets. Recent data showed that other than by well-known genetic (mutation, amplification) and epigenetic (DNA hypermethylation, histone post-translational modification) modifications, drug resistance mechanisms might also be regulated by splicing aberrations. This is a rapidly growing field of investigation that deserves future attention in order to plan more effective therapeutic approaches. The protocol described in this paper is aimed at investigating the impact of aberrant splicing on drug resistance in solid tumors and hematological malignancies. To this goal, we analyzed the transcriptomic profiles of several *in vitro* models through RNA-seq and established a qRT-PCR based method to validate candidate genes. In particular, we evaluated the differential splicing of DDX5 and PKM transcripts. The aberrant splicing detected by the computational tool MATS was validated in leukemic cells, showing that different DDX5 splice variants are expressed in the parental vs. resistant cells. In these cells, we also observed a higher PKM2/PKM1 ratio, which was not detected in the Panc-1 gemcitabine-resistant counterpart compared to parental Panc-1 cells, suggesting a different mechanism of drug-resistance induced by gemcitabine exposure.

Video Link

The video component of this article can be found at http://www.jove.com/video/54714/

Introduction

Despite considerable advances in cancer treatment, resistance of malignant cells to chemotherapy, either intrinsic or acquired upon prolonged drug exposure, is the major reason for treatment failure in a wide range of leukemia and solid tumors[1].

In order to delineate the mechanisms underlying drug resistance, *in vitro* cell line models are developed by stepwise selection of cancer cells resistant to chemotherapeutic agents. This procedure mimics the regimes used in the clinical settings and therefore allows in depth investigation of relevant resistance mechanisms. Resistant cells which survive the treatment are then distinguished from parental sensitive cells by using cell viability/cytotoxicity assays[2]. *In vitro* drug resistance profiles of primary cells have been shown to be significantly related to clinical response to chemotherapy[3].

High-throughput cytotoxicity assays constitute a convenient method to determine drug sensitivity *in vitro*. Herein, the viability of cells is assessed by for instance the 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide - MTT assay[4], which is based on metabolic conversion of certain substrates (*i.e.,* tetrazolium salts) into colored products, thereby reflecting the mitochondrial activity of cells. Alternatively, the cellular protein content can be quantified using the sulforhodamine B (SRB) assay[5]. Here, the number of viable cells is proportional to the optical density (OD) measured at an appropriate wavelength in a spectrophotometer, with no need of extensive and time-consuming cell counting procedures. The growth inhibition induced by a certain chemotherapeutic drug can be calculated based on the OD of the wells in which cells were treated with

a test agent and compared with the OD of untreated control cells. A dose-response curve is obtained by plotting drug concentrations versus percentages of viable cells relative to control cells. Finally, drug sensitivity can be reported as the concentration that results in 50% of cell growth inhibition as compared to untreated cells (IC_{50}).

The mechanisms underlying drug resistance include many different abnormalities, such as alterations affecting gene expression of determinants of drug activity and cellular metabolism. These molecular lesions, including mutations, aberrations at a transcriptional and post-transcriptional level as well as disturbed epigenetic regulation often affect genes involved either in drug metabolism or apoptosis[6].

Alternative pre-mRNA splicing and its intricate regulation have recently received considerable attention as a novel entity that may dictate drug resistance of cancer cells[7]. Up to 95% of human genes are alternatively spliced in normal cells by means of this tightly regulated process which produces many different protein isoforms from the same gene. Alternative splicing is often deregulated in cancer and several tumors are characterized by altered splicing of a growing number of genes involved in drug metabolism (*i.e.*, deoxycytidine kinase, folylpolyglutamate synthetase, or Multidrug resistance proteins)[6,8]. However, comprehensive analysis of splicing profiles of drug resistant cells is painfully lacking. Therefore, it is imperative to develop high throughput methods for alternative splicing analysis. This could help to develop more effective therapeutic approaches.

During the last decade, the rapid development of next generation sequencing (NGS) technologies has enriched biomedical research with new insights into the molecular mechanisms governing regulation of genome expression and their role in various biological processes[9]. RNA-sequencing (RNA-seq) is a powerful sub-application of NGS in the field of transcriptomics. It allows a genome-wide profiling (both qualitatively and quantitatively) of the expression patterns of thousands of genes simultaneously and is well suited for the characterization of novel coding mRNAs as well as long non-coding RNA, miRNA, siRNA, and other small RNA classes (*e.g.*, snRNA and piRNA)[10,11].

RNA-Seq has many advantages over previous technologies for transcriptome characterization (*e.g.,* Sanger sequencing and expression microarrays). It is not based on existing genome annotation, it has a single-nucleotide level of resolution and it has a broader dynamic range for expression level estimation. Briefly, the basic experimental workflow of RNA-seq experiments consists of polyadenylated transcript (mRNA) selection and fragmentation, followed by conversion into cDNA, library construction and finally, massively parallel deep sequencing[12,13]. Due to rapid drop of sequencing costs over the last few years, RNA-seq is gradually replacing other technologies and significant efforts are being made to improve the library preparation protocols. For instance, it is now possible to retain the strand information of mRNA transcripts by marking the second strand cDNA with deoxyuridine triphosphate (dUTP) and, prior to PCR amplification, digesting the marked strand with uracil-DNA-glycosilase (UDG). This process enhances the accuracy of gene annotation and estimation of the expression levels[14,15].

The analysis and interpretation of RNA-seq data require complex and powerful computational software packages and processing within bioinformatic pipelines[16,17]. First, the raw reads undergo quality control by removing technical and biological artefacts and discarding (trimming) the sequences which do not reach stringent quality requirements. Subsequently the reads for each sample are mapped to a reference genome and indexed into gene-level, exon-level, or transcript-level, in order to determine the abundance of each category. Depending on the application, refined data are then computed through statistical models for the identification of allele-specific expression, alternative splicing, gene fusions and single nucleotide polymorphisms (SNPs)[12]. Finally, differential analysis on selected level (*i.e.,* gene expression or alternative splicing) can be used to compare samples obtained under different conditions.

Differential splicing analysis describes the differences in splice site usage between two samples. An increasing number of software packages devoted to this purpose are available based on different statistical models, performances and user interface[18]. Among these, MATS (Multivariate Analysis of Transcript Splicing) emerges as a freely available and precise computational tool based on a Bayesian statistical framework and designed to detect differential splicing events from either single or paired end RNA-seq data. Starting from the aligned (.bam) files, MATS can detect all major types of alternative splicing events (exon skipping, alternative 3' splice site, alternative 5' splice site, mutually exclusive exons and intron retention - also see **Figure 1**).

First, the software identifies reads which support a certain splice event, for instance exon skipping, and classifies them into two types. "Inclusion reads" (for the canonical splice event) map within the investigated exon and span the junctions between that specific exon and the two upstream and downstream flanking exons. "Skipping reads" (for the alternative splice event) span the junction between the two flanking exons. Subsequently, MATS returns the normalized inclusion level for both the canonical and alternative events and compares values between samples or conditions. Ultimately, it calculates P-value and false discovery rate (FDR) assuming that the difference in the variant ratio of a gene between two conditions exceeds a given user-defined threshold for each splicing event[19,34].

Following differential splicing analysis in conjunction with RNA-seq, an extensive experimental validation is warranted in order to identify true-positive gene candidates[18]. Quantitative reverse transcribed-polymerase chain reaction (qRT-PCR) is the most commonly used and optimal method in validation of candidates obtained from RNA-Seq analysis[20]. The aim of this paper is to provide a robust methodology to investigate drug resistance-related splicing profiles in solid tumors and hematological malignancies. Our approach utilizes RNA-seq-based transcriptome profiling of selected cell line models of drug resistant cancers in combination with an established qRT-PCR method for the validation of candidate genes implicated in drug resistance.

The human leukemia cell line models used in this study included pediatric T-cell acute lymphoblastic leukemia (T-ALL) cell line CCRF-CEM (CEM-WT), its two glucocorticoid (GC)-resistant subclones CEM-C7H2-R5C3 (CEM-C3) and CEM-C7R5 (CEM-R5)[21,22] and the methotrexate (MTX)-resistant subline CEM/R30dm[23]. Although current therapies based on GCs and MTX establish clinical benefit in about 90% of cases, the emergence of GC-resistance still represents an unsolved problem with an unclear molecular mechanism. To isolate GC-resistant sub-clones, CEM-WT cells were cultured in 1 µM dexamethasone (Dex) for 2 to 3 weeks. MTX-resistant subline CEM/R30dm was developed through repeated short- term (24 hr) exposure of CEM-WT cells to 30 µM MTX as a mimic of clinical protocols. Interestingly, this cell line also displayed cross-resistance to Dex (unpublished results) for which the mechanism is not fully understood.

The solid tumor model investigated in the present study is pancreatic ductal adenocarcinoma, notorious for its extraordinary refractoriness to chemotherapy. To this end, we selected Panc-1 cell line and its gemcitabine-resistant sub-clone Panc-1R obtained by continuous incubation with 1 µM of the drug[24]. Here we describe an approach to discover novel mechanisms underlying *in-vitro* drug resistance by combining three

protocols: colorimetric cytotoxicity assays to assess drug sensitivity in leukemic cells and cancer cells from solid tumors, RNA-seq-based pipeline to identify novel splice variants related to drug sensitivity/resistance and RT-PCR and qRT-PCR analysis to validate potential candidates.

Protocol

1. Characterization of Drug Resistance Profiles through Cytotoxicity Assays

1. **Leukemic Cell Line Culture**
   1. Maintain the parental T-cell ALL cell line CCRF-CEM (CEM-WT) as well as its drug resistant sublines, including CEM/R30dm, CEM-R5 and CEM-C3, in 25 cm^2 in 10 ml RPMI-1640 medium containing 2.3 µM folic acid supplemented with 10% fetal calf serum and 100 units/ml penicillin G and 100 µg/ml streptomycin.
   2. Culture the cells in a humidified atmosphere at 37 °C in a 5% CO_2 incubator.
   3. Allow cell growth to concentrations between 2 - 3 x 10^6 cell/ml.
   4. Split the cells twice a week at an initial concentration of 0.3 x 10^6 cells/ml (*e.g.,* if the cell concentration is 3 x 10^6 cells/ml, transfer 1 ml cell suspension in a new flask containing 9 ml fresh medium). Discard the culture after 20 consecutive passages.

2. **Pancreatic Carcinoma Cell Line Culture**
   NOTE: Maintain the human pancreatic carcinoma cell line Panc-1 in 75 cm^2 culture flasks in 10 ml DMEM medium with high glucose and L-Glutamine supplemented with 10% fetal bovine serum and 100 units/ml penicillin G and 100 µg/ml streptomycin. The drug resistant variant, Panc-1R, is cultured in the same culture medium containing 1 µM gemcitabine dissolved in sterile water. Further details are provided in [24].
   1. Culture the cells at 37 °C in a 5% CO_2 incubator.
   2. Split the cells every 2 - 3 days at a ratio of 1:5 when cells reach confluence of about 90%.
   3. To split the cells, wash the twice with phosphate buffered saline (PBS), add 1 ml of Trypsin/EDTA per 75 cm^2 culture flask and incubate at 37 °C for 3 min.
   4. Add 9 ml medium and harvest the detached cells in a 15 ml tube. Seed 2 ml cell suspension in a new flask containing 8 ml medium. Discard the culture after 20 consecutive passages.

3. **MTT Assay for Leukemic Cells**
   1. Prepare in advance the MTT solution: dissolve 500 mg of MTT formazan in 10 ml PBS and stir (protected from light) with a magnetic stirrer for approximately 1 hr. Sterilize the solution with a 0.22 µm filter. NOTE: The solution can be stored in 10 ml aliquots at −20 °C. Protect from direct light after thawing.
   2. Prepare in advance the acidified isopropanol: add 50 ml of 2 M HCl to 2.5 L of isopropanol. NOTE: Store the solution for at least one month at room temperature before use. If the isopropanol is not acidified correctly, it might form precipitates with the medium and compromise the spectrophotometric readout.
   3. Prepare a separate 96-well flat bottom "Day 0" (control) plate to ensure more accurate estimation of growth inhibition: dedicate 3 to 6 wells per cell line, add 30 µl of growth medium and 120 µl of cell suspension (8,000 cells) per each well and 150 µl of growth medium to wells corresponding to blanks (no cells). Proceed from step 1.3.9 to step 1.3.13 of this section to measure the optical density (OD). NOTE: Further details are provided in [4].
   4. Prepare a 96-well flat bottom experimental plate: dedicate 30 wells to drug concentrations (10 concentrations, each in triplicate), 10 wells to control cells and 10 wells to control medium without cells (blanks) and prepare a drug dilution range of dexamethasone (Dex) using Dimethyl Sulfoxide as a solvent.
      NOTE: For CEM-WT cells the Dex dilution range is between 2 µM and 0.97 nM. For CEM/R30dm, CEM-R5 and CEM-C3 cells the Dex dilution range is between 640 µM and 0.33 nM.
   5. Add 30 µl from each Dex dilution into an appropriate well of the 96-well plate. Make sure to include each concentration in a triplicate.
   6. Add 30 µl of growth medium to wells corresponding to control cells and 150 µl of growth medium to wells corresponding to blanks.
   7. Harvest exponentially growing cells and resuspend at their optimal seeding concentration.
      NOTE: In order to determine optimal starting cell concentrations, it is recommended to assess the growth profile of each cell line cell line in a 96-well plate by seeding cells at several concentrations and measuring it daily for at least 4 days. Choose a seeding concentration which prevents overgrowth of cells after 72 hr, as this will influence the experiment by saturating the OD values. For CEM-WT, CEM-C5 and CEM-R5 the optimal seeding concentration is 8,000 cells/well, while for CEM/R30dm it is 5,000 cells/well.
   8. Add 120 µl of cell suspension to each well containing either the drug solution or the growth medium (wells corresponding to control cells). Fill the empty outer wells of the plate with 150 µl PBS to ensure good humidity in the plate and incubate the plates for 72 hr at 37 °C with 5% CO_2 in a cell culture incubator.
   9. Add 15 µl of the MTT solution to each well and shake the plate for 5 min with a plate-shaker up to a maximum of 900 shakes/min.
   10. Place the plates back at 37 °C with 5% CO_2 in a cell culture incubator and incubate for another 4 - 6 hr.
   11. Add 150 µl of the acidified isopropanol to each well and mix well with a multichannel pipette to thoroughly resuspend all the formazan crystals. Start with the blank wells and make sure to rinse the tips well before proceeding to another row of the plate.
   12. Incubate the plate at room temperature (RT) for 10 min.
   13. Using a microplate reader, determine the OD at 540 and 720 nm to ensure an accurate measurement by correcting for background OD. Then save data in a spreadsheet file and analyze it[4].

4. **SRB Assay for Pancreatic Carcinoma Cells**
   1. Dissolve SRB reagent in 1% acetic acid at final concentration of 0.4% (w/v).
   2. Dissolve trichloroacetic acid (TCA) in ultrapure water at final concentration of 50% (w/v).
   3. Dissolve Tris(hydroxymethyl)-aminomethane in ultrapure water at final concentration of 10 mM.
   4. Prepare a separate 96-well flat bottom "Day 0" control plate to ensure more accurate estimation of growth inhibition: seed 6 wells with cells growing in exponential phase in 100 µl medium at appropriate seeding concentration and add 100 µl medium only to wells corresponding to blanks. Incubate overnight at 37 °C with 5% CO_2 to ensure proper adhesion of the cells to the plate. Then add 100 µl medium to all the wells and proceed to steps 1.4.8 - 1.4.16 of this section.

5. Prepare a 96-well flat bottom experimental plate: seed cells growing in exponential phase in triplicate in 96 wells flat bottom plates at the appropriate density in 100 µl of medium by using a multichannel pipet.
   NOTE: In order to determine optimal starting cell concentrations, it is recommended to assess the growth profile of each cell line cell line in a 96-well plate by seeding cells at several concentrations and measuring it daily for at least 4 days. Choose a seeding concentration which prevents overgrowth of cells after 72 hr, as this will influence the experiment by saturating the OD values. For Panc-1 and Panc-1R cells, the optimal seeding concentration is 8000 cells/well.
6. Add 100 µl medium to medium-only wells and incubate overnight at 37 °C with 5% CO_2 to ensure proper adhesion of the cells to the plate.
7. Prepare a drug dilution range of gemcitabine between 1 µM and 10 nM for Panc-1 and 1 mM and 100 nM for Panc-1R cells. Add 100 µl from each dilution into an appropriate well of the 96-well plate by using a multichannel pipet. Make sure to have each concentration in triplicate. In addition, add 100 µl medium to the medium-only wells and the control cells. Incubate at 37 °C with 5% CO_2 for 72 hr.
8. Add 25 µl cold TCA solution to the wells by using a multichannel pipet and incubate the plates for at least 60 min at 4 °C in order to precipitate and fix the proteins at the bottom of the wells.
9. Empty the plate by removing the medium and dry briefly on a tissue.
10. Wash 5 times with tap water, then empty the plate and let dry at room temperature.
11. Add 50 µl SRB solution per well by using a repeat-pipet and stain for 15 min at room temperature.
12. Empty the plate by removing the SRB stain.
13. Wash 4 times with 1% acetic acid, then empty the plate and let it dry at room temperature.
14. Add 150 µl Tris solution per well by using a multichannel pipet and mix for 3 min on a plate-shaker up to a maximum of 900 shakes/min.
15. Read the optical density at 540 nm (or 492 nm if the OD values are too high).
16. Analyze the data.

5. **Data Analysis for MTT and SRB Assays**
   1. Calculate the OD values of cells at "Day 0" using the following formula: $OD_{Day\ 0} = OD_{control\ cells}$ - Average $OD_{blank\ wells}$
   2. Calculate the percentage of surviving cells for each drug concentration according to the following formula: % Treated cells = Average $[OD_{treated\ cells}$ - Average $OD_{blank\ wells}$ - $OD_{Day\ 0}]/[OD_{control\ cells}$ - Average $OD_{blank\ wells}$ - $OD_{Day0}]*100$
   3. Plot the dose-response curve (drug concentration vs. growth inhibition in %).
   4. Calculate the concentration of the drug which inhibits the growth of cells by 50% (IC_{50}) using the dose-response curve.

2. RNA Isolation and Library Preparation for RNA-sequencing

1. **Sample Collection and RNA Isolation**
   1. For CEM cells: harvest 10^6 cells directly from culture medium.
   2. For Panc-1 and Panc-1R: remove medium, wash the cells twice with PBS and detach by adding trypsin-EDTA and incubating at 37 °C for 3 min. Add culture medium and harvest 10^6 cells.
   3. Spin down samples at 300 x g for 3 min, remove supernatant and extract total mRNA using commercially available silica membrane spin columns, by following the manufacturer`s protocol.
   4. Determine the concentration and purity of total RNA by using a UV-Vis spectrophotometer.
      NOTE: RNA is considered of high purity if the 260 nm / 280 nm absorbance ratio is above 1.8. The samples can be stored at - 80 °C.
   5. Assess total RNA integrity by electrophoresis of 200 ng of sample on 1% agarose gel stained with ethidium bromide.
      NOTE: The detection of two intact bands corresponding to mammalian 28S and 18S rRNAs at approximately 2 kb and 1 kb in size is indicative of good total RNA integrity.

2. **Sequencing Library Preparation**
   1. Use 2 µg of total mRNA per each sample. Follow mRNA library preparation protocol according to manufacturer's instructions.
   2. Determine quality and concentration of each library by using a bioanalyzer system. Pool the libraries in a single sample up to a final concentration of 10 nmol/L and measure it with the bioanalyzer.
   3. Use high-throughput sequencing system with Single Read 100 bp mode.
      NOTE: Read length of > 80 bp is necessary for identifying transcriptional isoforms.

3. Detection of Differential Splicing from Sequencing Reads

1. **Alignment of Reads to Reference Genome and Quality Check**
   1. Compile a gene annotation (.gtf file) from the NCBI refGene table using UCSC table browser (25,963 genes).
   2. Perform annotation-aware gapped alignment of sequencing reads to the reference genome (GRCh37) using STAR.
   3. Sort and index the resulting alignment files with picard tools.
   4. Perform post-mapping quality control by using RSeQC and samtools.

2. **Differential Splicing Detection between Sample Groups**
   1. Install Python and corresponding versions of NumPy and SciPy. Download and install samtools. Download and install bowtie and tophat,
   2. Add the Python, bowtie, tophat and samtools directories to the $PATH environment variable. Download pre-built bowtie indexes (hg19). Download rMATS version 3.0.9.
      NOTE: Further details about rMATS are provided in [19] and [34].
   3. Detect alternative splicing events by comparing each Dex-resistant cell line to CEM-WT and Panc-1 to Panc-1R in separate MATS runs according to **Figure 5A**.

4. To detect differential splicing events from previously aligned sequencing reads (.bam files), run MATS using the commands from **Figure 5B** for CEM-WT vs. CEM- C3, CEM-WT vs. CEM-R5, CEM-WT vs. CEM/R30dm and Panc1 to Panc-1R comparisons.
   NOTE: MATS will create an output folder with two .txt files with results per type of splicing event analyzed (SE - exon skipping, A5SS - alternative 5' splice site, A3SS - alternative 3' splice site, MEX - mutually exclusive exons and RI - intron retention): one file containing results based on junction counts only and a second file with results based on junction counts and reads on target. Moreover, an additional .txt file with a result overview is generated in the same folder.
5. For SE, A5SS, A3SS and RI import to spreadsheets the .txt files based on junction counts and reads on target. For MEX import the .txt files based on junction counts only.
   NOTE: MATS output file is sorted by ascending P-values and contains several parameters: gene ID, gene symbol, chromosome and strand position, genomic coordinates of the alternatively spliced fragments, counts as well as the length of inclusion and skipping forms for both analyzed samples, the length of inclusion and skipping form used for normalization, p-value, false discovery rate (FDR), inclusion level for each sample based on normalized counts and the differential inclusion score (average(IncLevel1) - average(IncLevel2)).
6. Select statistically significant candidate genes with an FDR<10% for further validation (*e.g.*, DDX5 and PKM2).
7. Visualize alternative splicing events with Integrative Genomics Viewer (IGV, https://www.broadinstitute.org/igv/), as reported in **Figure 5**.

4. Validation of the Results by RT-PCR and qRT-PCR Assays

1. **Primer Design**
   NOTE: In order to provide a reliable validation of results obtained using the bioinformatic pipeline, the mRNA transcripts resulting from alternative splicing events are amplified by using RT-PCR and visualized by agarose gel electrophoresis of the PCR products. Cyanine green dye such as SYBR green qRT-PCR assay is used to quantify specific splice variants relative to a housekeeping gene. **Figure 7** depicts the strategy applied to visualize the exon 12 skipping event of the DDX5 gene and to quantify the mutually exclusive exon 9 and exon 10 of the PKM gene.
   1. For RT-PCR assay, design primer pairs which anneal to constitutive exons (exon 10 and exon 13) located upstream and downstream the alternative splicing sites (**Figure 7A**).
      NOTE: The amplicon size should cover between 100 and 800 bp in order to ensure a clear separation of the predicted PCR products on agarose gel. The annealing temperature of the primers should be between 55 and 65 °C and the GC content should not exceed 60%.
   2. Cyanine green assay (**Figure 7B**).
      1. Check the sequence homology of the two mutually exclusive exons and design two primer pairs each of which specifically and exclusively detects only one of the two splice variants.
         NOTE: For PKM, the reverse primer anneals to exon 11 common to both variants, while the forward primers are variant-specific and anneal to exon 9 (PKM1) or exon 10 (PKM2).
      2. In order to detect DDX5 full-length gene, anneal the reverse primer within the skipped exon (exon 12).
         NOTE: For the specific detection of the DDX5 ΔEx12 variant, the reverse primer spans the exon11/exon13 boundary. Use the same forward primer which anneals to constitutive exon 10 for both reactions.
         NOTE: The amplicon size should be between 80 and 200 bp.

2. **First strand cDNA Synthesis**
   1. Set up reverse transcription of 1 µg of the isolated RNA to cDNA by using 200 U/µl of Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase in its reaction buffer diluted 1:5 with sterile water. Add DTT 1 µM, 0.05 µg of random hexamers, deoxynucleotide mix (dNTP) 1 mM, and 40 U/µl of ribonuclease inhibitor.
   2. Vortex briefly and incubate the reaction mix at 37 °C for 2 hr.
   3. Incubate the reaction mix at 70 °C for 5 min to inactivate the reverse transcriptase, transfer the mix on ice and spin down by using a microcentrifuge. Samples can be used immediately or stored at - 20 °C.

3. **RT-PCR Reaction and Agarose Gel**
   1. Set up the PCR reaction in a PCR tube per each sample by mixing 12.5 µl of 2x concentrated PCR mastermix, 1.25 µl of 10 µM forward primer and the same amount for the reverse primer up to a final volume of 25 µl with sterile water.
   2. Add 1 µl of cDNA to the mix and place the tubes in the thermocycler. Run the program as follows: initial denaturation: 95 °C for 2 min. Reiterate the following steps for 35 cycles: denaturation at 95 °C for 25 sec; annealing at 52 °C for 35 sec; extension at 72 °C for 1 min. Set final extension at 72 °C for 5 min.
   3. Prepare 1% agarose gel by dissolving 1 g of agarose in 100 ml 1x TBE buffer. Add 2.5 µl of ethidium bromide to the solution.
      CAUTION: ethidium bromide is carcinogenic, handle with care by using a chemical fume hood.
   4. Load the samples and run the gel in 1x TBE buffer at 100 V for approximately 30 min in an electrophoresis system.
   5. Reveal the gel with a digital UV camera and save the image.

4. **qRT-PCR and Data Analysis**
   1. Set up the Cyanine green PCR reaction per each sample by mixing 12.5 µl of 2x green PCR Mastermix, 2 µl of 5 µM forward primer and the same amount for the reverse primer up to a final volume of 15 µl with sterile water in a PCR tube. Prepare a mix for each specific splice variant to be detected (PKM1, PKM2, DDX5 full length, DDX5 ΔEx12) and for the housekeeping gene (GUS).
   2. Load the mastermix on a white 96-well plate and prepare duplicates per each sample.
   3. Dilute the cDNA 10x in water, add 5 µl to each mix and place the plate in the thermocycler. Run the program as follows: initial denaturation: 95 °C for 5 min. Reiterate the following steps for 45 cycles: denaturation at 95 °C for 10 sec; annealing at 58 °C (for DDX5 and DDX5 ΔEx12 mixes) or 60 °C (for PKM1 and PKM2 mixes) for 20 sec. Set final extension at 72 °C for 20 sec. Set melting curve by applying a gradient from 65 to 97 °C.

4. In order to assess the specificity of the primer sets to their target, check the melting curves and verify that a single peak is formed per each primer set.
5. Calculate the second derivative values of the amplification curves and export the cycle threshold values (Ct).
6. Calculate the relative expression levels (REL) of the mRNA splice variants compared to the GUS housekeeping (reference) gene in each sample by using the "delta Ct (ΔCt)" method. The formula is: $REL = 2^{-\Delta Ct}$, where $\Delta Ct = Ct_{\text{target splice variant}} - Ct_{\text{reference gene}}$
7. In order to quantify the relative abundance of splice variants, calculate the REL ratio by using the formula: Ratio REL ratio = $REL_{\text{splice variant 1}} / REL_{\text{splice variant 2}}$.

Representative Results

The cytotoxicity assays described in the protocol provide a reliable and robust method to assess the resistance of cancer cells to chemotherapeutic agents *in vitro*. By means of the MTT assay, sensitivity to Dex was determined in four T-ALL cell lines, including Dex-sensitive parental CEM-WT cells, and three Dex-resistant sublines: CEM/R30dm, CEM-R5 and CEM-C3. Two different concentration ranges had to be used due to the large difference in sensitivity between CEM-WT (2 μM - 0.97 nM) and the Dex resistant cell lines (640 μM - 0.33 nM). The MTT assay clearly showed Dex resistance in CEM/R30dm (IC_{50} = 456 ± 49 μM), CEM-R5 (IC_{50} > 640 μM) and CEM-C3 (IC_{50} = 386 ± 98 μM) as compared to the CEM-WT cells (IC_{50} = 0.028 ± 0.003 μM). Similarly the SRB assay demonstrated high gemcitabine resistance in Panc-1R cell line (IC_{50} = 3.16 ± 0.01 μM) compared to the parental Panc-1 cells (IC_{50} = 0.077 ± 0.03 μM) as showed in **Figure 2.**

Following confirmation of drug resistance in all considered cell lines, we next proceeded to mRNA isolation, library preparation and RNA-sequencing. Extraction of total mRNA by using silica membrane spin columns is the preferred choice over other methods since it avoids phenol and protein contamination and provides high purity of the sample with 260 nm / 280 nm absorbance ratio well above 1.8. This is a crucial requirement for complex downstream applications such as deep sequencing. The method used to check the integrity of RNA samples by agarose 1% is particularly suitable for freshly isolated cells (**Figure 3**). Since the purpose of this protocol is to detect alternative splice variants aberrantly expressed in drug resistant cell lines, we choose positive selection of polyadenylated mRNA for the sequencing library preparation, by using stranded mRNA kit. Electropherograms of single and pooled libraries are depicted in **Figure 4**, showing an average fragment size of about 300 bp, consistent with sequencing system requirements. The prepared libraries were then sequenced using a chip with Single Read 100 bp mode. The choice of sequencing reads of 100 bp is necessary to detect alternative splicing through downstream bioinformatic pipelines.

After initial processing steps and quality control, the clean reads aligned to human genome (hg19) were subjected to differential splicing analysis using MATS. In this analysis we made comparisons between the drug sensitive parental cell line and each of its drug resistant sublines separately (*i.e.,* CEM WT *vs.* CEM/R30dm, CEM WT *vs.* CEM-C3, *etc.*). MATS relies on a flexible and precise statistical model used to detect differential splicing between samples. By using default analysis options and a FDR value < 10% as a cut-off (depicted in **Figure 5B**), we were able to identify 38 ± 12 significant differentially spliced gene candidates per comparison ordered by type of alternative splicing event, with most hits classified as exon skipping. **Figure 6** illustrates a typical analysis output for the comparison Panc-1 vs. Panc-1R.

We further focused our study on two most common types of alternative splicing events: exon skipping and mutually exclusive exon events with one representative candidate per category described below. DDX5 (DEAD-box helicase 5) has been detected by MATS analysis as statistically significant in the comparison CEM-WT *vs.* CEM-C3 and CEM-R5, but not significant in the comparison CEM-WT *vs.* CEM/R30dm. Given its putative role in leukemia[25,26], we choose this candidate for further validation. It is highly recommended to visualize RNA-seq data in a genome browser-like tool. IGV provides a versatile and user-friendly interface for this purpose, as shown in **Figure 7** for the gene candidate DDX5. PKM (pyruvate kinase muscle isozyme) is a statistically significant mutually exclusive exon event in the comparison CEM-WT *vs.* all Dex resistant sublines, but not in the comparison Panc-1 *vs.* Panc-1R. Given the relevance of this enzyme in solid tumor metabolism[27,28] and the emerging role of cell metabolism in glucocorticoid resistance in T-ALL[29], we choose this candidate for further validation using RT-PCR.

Primer design must be conducted with extreme care in order to amplify the correct amplicon, especially when primers anneal to exon-exon boundaries (in case of the reverse primer detecting DDX5 ΔEx12 variant) or when they anneal to mutually exclusive exons with high sequence homology (exon 9 and exon 10 of PKM). **Figure 8** shows the primer design strategy, while **Figure 9** shows the results of an RT-PCR for DDX5 gene candidate. DDX5 ΔEx12 is detected in the sample CEM-WT and CEM/R30dm but not in CEM-C3 and CEM-R5, thus confirming MATS data in a qualitative fashion. Cyanine green qRT-PCR assay accurately quantifies the mRNA expression levels of the DDX5 and PKM splice variants, as shown in **Figure 10** and **Figure 11**, respectively.

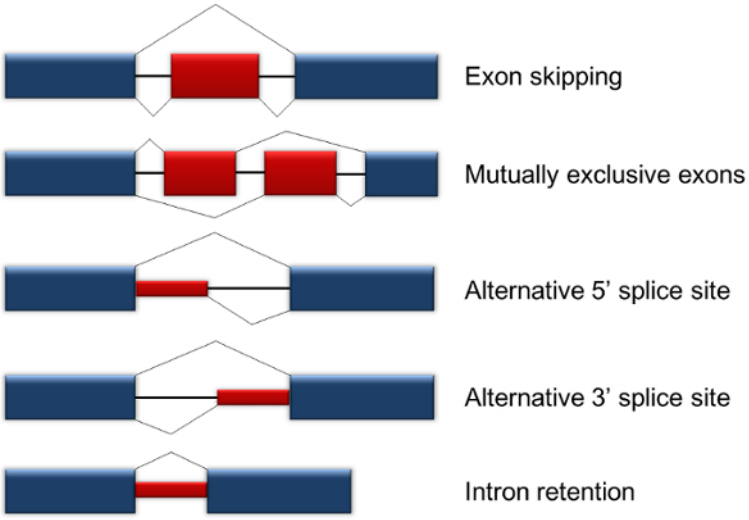


Figure 1: Schematic Representation of Alternative Splicing Events. Schematic representation of the possible patterns of alternative splicing of a gene. Boxes are discrete exons that can be independently included or excluded from the mRNA transcript. Please click here to view a larger version of this figure.

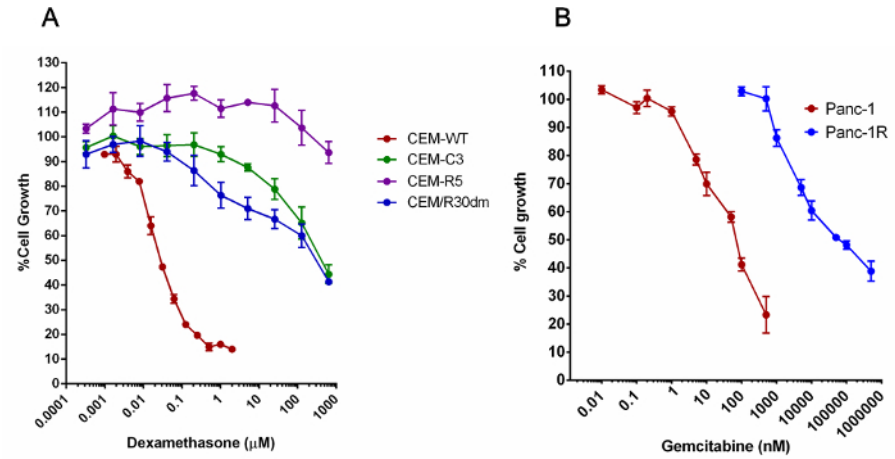


Figure 2: Results of Cytotoxicity Assays. (A) MTT assay for leukemic cell lines shows high levels of dexamethasone resistance in CEM-C3 (IC_{50} = 386 ± 98), CEM-R5 (IC_{50} > 640 μM) and CEM/R30dm (IC_{50} = 456 ± 49 μM) as compared to the parental CEM-WT (IC_{50} = 0.028 ± 0.003 μM). **(B)** SRB assays for pancreatic carcinoma cell lines reveal high levels of gemcitabine resistance in Panc-1R (IC_{50} = 3.16 ± 0.01 μM) as compared to the parental Panc-1 (IC_{50} = 77.22 ± 2.76 nM). The graphs report mean cell growth % ± SEM of three independent experiments. Please click here to view a larger version of this figure.

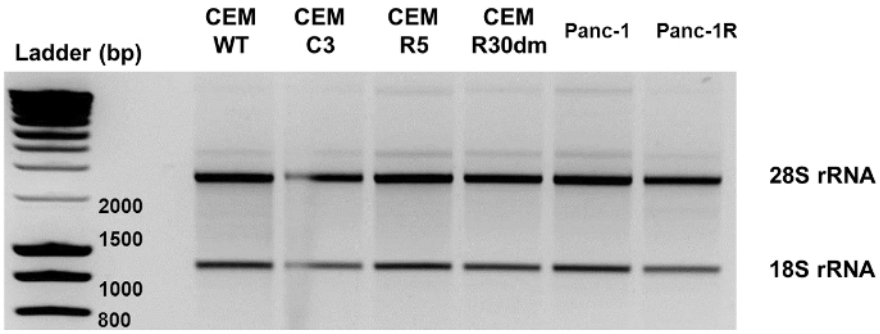


Figure 3: RNA Quality Assessment using an Agarose Gel. Two hundred ng of total mRNA were run on 1% agarose gel stained with ethidium bromide. The presence of intact bands corresponding to ribosomal RNA (rRNA) species 18S and 28S and the absence of smears at lower molecular weights are indicative of good quality RNA. Please click here to view a larger version of this figure.

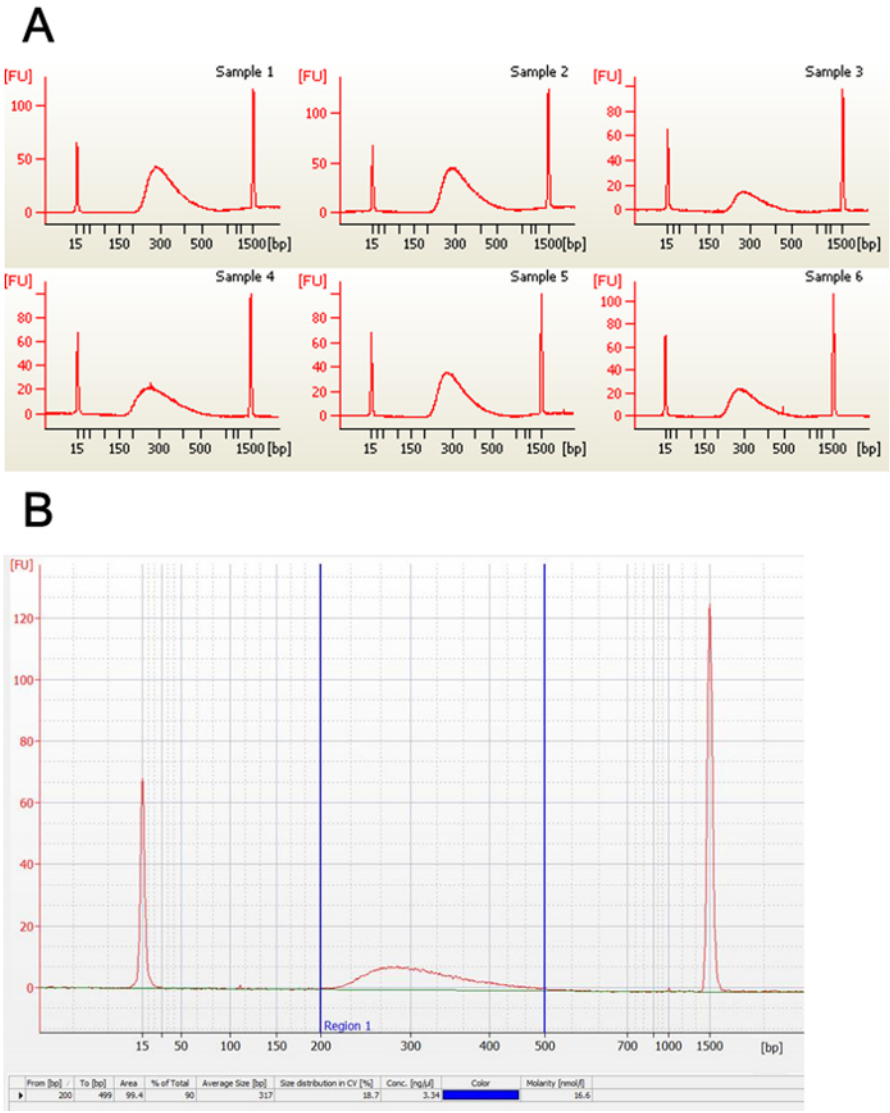


Figure 4: Bioanalyzer Traces of Sequencing Libraries. (A) Electropherograms of sequencing libraries show peaks at approximately 300 bp, which is indicative of good quality. Sample 1 to 6 = CEM-WT, CEM-C3, CEM-R5, CEM/R30dm, Panc-1 and Panc-1R. **(B)** Electropherogram of the pooled samples (FU, Fluorescence Units).

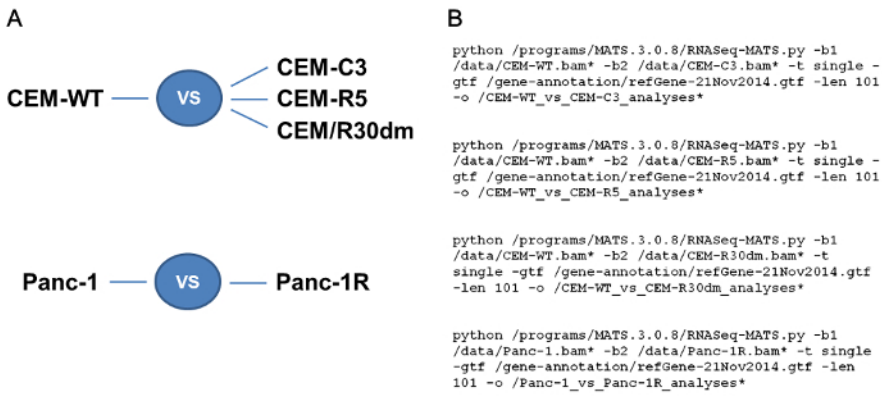
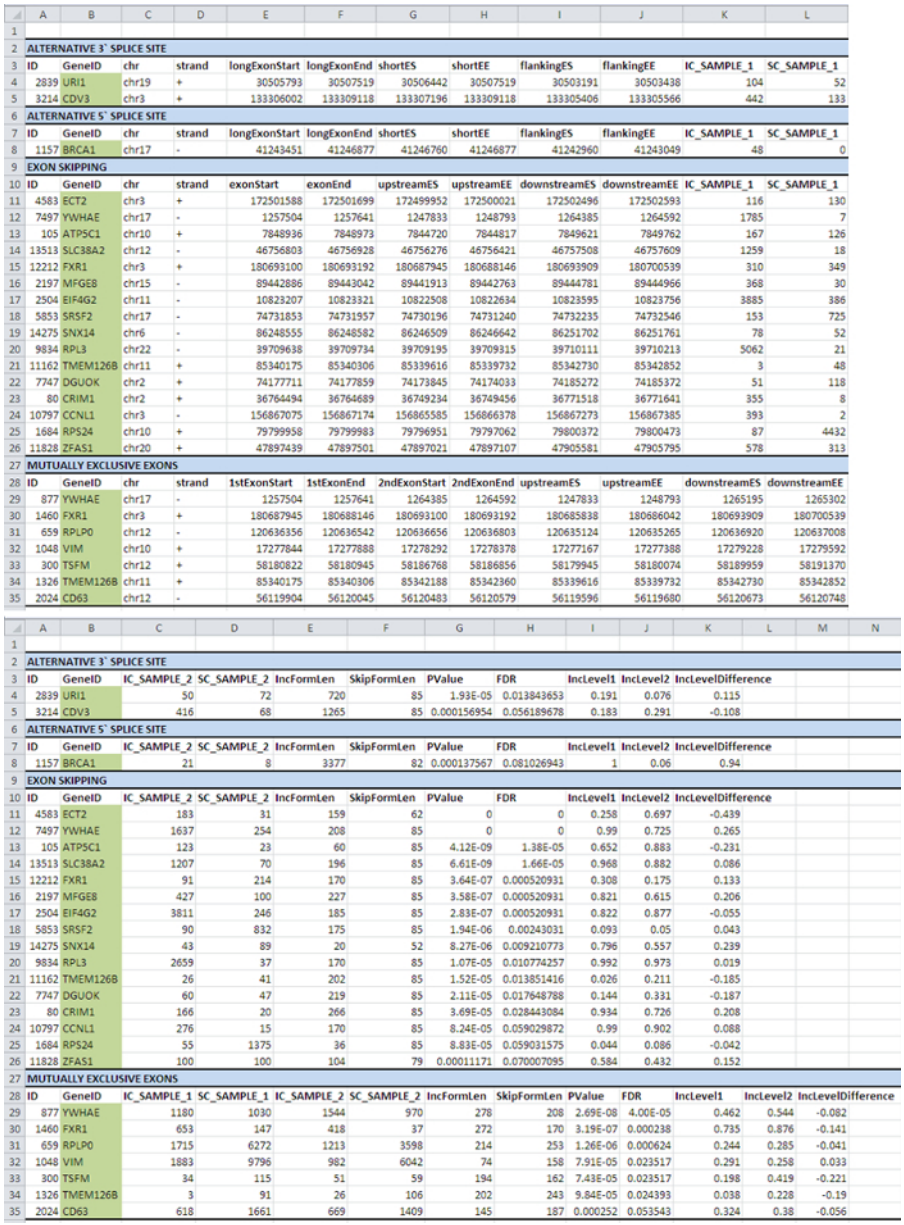


Figure 5: Detection of Differential Splicing with MATS. (A) Group comparisons and **(B)** scripts used to run MATS. Please click here to view a larger version of this figure.



	A	B	C	D	E	F	G	H	I	J	K	L
1												
2	ALTERNATIVE 3` SPLICE SITE											
3	ID	GeneID	chr	strand	longExonStart	longExonEnd	shortES	shortEE	flankingES	flankingEE	IC_SAMPLE_1	SC_SAMPLE_1
4	2839	URI1	chr19	+	30505793	30507519	30506442	30507519	30503191	30503438	104	52
5	3214	CDV3	chr3	+	133306002	133309118	133307196	133309118	133305406	133305566	442	133
6	ALTERNATIVE 5` SPLICE SITE											
7	ID	GeneID	chr	strand	longExonStart	longExonEnd	shortES	shortEE	flankingES	flankingEE	IC_SAMPLE_1	SC_SAMPLE_1
8	1157	BRCA1	chr17	-	41243451	41246877	41246760	41246877	41242960	41243049	48	0
9	EXON SKIPPING											
10	ID	GeneID	chr	strand	exonStart	exonEnd	upstreamES	upstreamEE	downstreamES	downstreamEE	IC_SAMPLE_1	SC_SAMPLE_1
11	4583	ECT2	chr3	+	172501588	172501699	172499952	172500021	172502496	172502593	116	130
12	7497	YWHAE	chr17	-	1257504	1257641	1247833	1248793	1264385	1264592	1785	7
13	105	ATP5C1	chr10	+	7848936	7848973	7844720	7844817	7849621	7849762	167	126
14	13513	SLC38A2	chr12	-	46756803	46756928	46756276	46756421	46757508	46757609	1259	18
15	12212	FXR1	chr3	+	180693100	180693192	180687945	180688146	180693909	180700539	310	349
16	2197	MFGE8	chr15	-	89442886	89443042	89441913	89442763	89444781	89444966	368	30
17	2504	EIF4G2	chr11	-	10823207	10823321	10822508	10822634	10823595	10823756	3885	386
18	5853	SRSF2	chr17	-	74731853	74731957	74730196	74731240	74732235	74732546	153	725
19	14275	SNX14	chr6	-	86248555	86248582	86246509	86246642	86251702	86251761	78	52
20	9834	RPL3	chr22	-	39709638	39709734	39709195	39709315	39710111	39710213	5062	21
21	11162	TMEM126B	chr11	+	85340175	85340306	85339616	85339732	85342730	85342852	3	48
22	7747	DGUOK	chr2	+	74177711	74177859	74173845	74174033	74185272	74185372	51	118
23	80	CRIM1	chr2	+	36764494	36764689	36749234	36749456	36771518	36771641	355	8
24	10797	CCNL1	chr3	-	156867075	156867174	156865585	156866378	156867273	156867385	393	2
25	1684	RPS24	chr10	+	79799958	79799983	79796951	79797062	79800372	79800473	87	4432
26	11828	ZFAS1	chr20	+	47897439	47897501	47897021	47897107	47905581	47905795	578	313
27	MUTUALLY EXCLUSIVE EXONS											
28	ID	GeneID	chr	strand	1stExonStart	1stExonEnd	2ndExonStart	2ndExonEnd	upstreamES	upstreamEE	downstreamES	downstreamEE
29	877	YWHAE	chr17	-	1257504	1257641	1264385	1264592	1247833	1248793	1265195	1265302
30	1460	FXR1	chr3	+	180687945	180688146	180693100	180693192	180685838	180686042	180693909	180700539
31	659	RPLP0	chr12	-	120636356	120636542	120636656	120636803	120635124	120635265	120636920	120637008
32	1048	VIM	chr10	+	17277844	17277888	17278292	17278378	17277167	17277388	17279228	17279592
33	300	TSFM	chr12	+	58180822	58180945	58186768	58186856	58179945	58180074	58189959	58191370
34	1326	TMEM126B	chr11	+	85340175	85340306	85342188	85342360	85339616	85339732	85342730	85342852
35	2024	CD63	chr12	-	56119904	56120045	56120483	56120579	56119596	56119680	56120673	56120748

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2	ALTERNATIVE 3` SPLICE SITE													
3	ID	GeneID	IC_SAMPLE_2	SC_SAMPLE_2	IncFormLen	SkipFormLen	PValue	FDR	IncLevel1	IncLevel2	IncLevelDifference			
4	2839	URI1	50	72	720	85	1.93E-05	0.013843653	0.191	0.076	0.115			
5	3214	CDV3	416	68	1265	85	0.000156954	0.056189678	0.183	0.291	-0.108			
6	ALTERNATIVE 5` SPLICE SITE													
7	ID	GeneID	IC_SAMPLE_2	SC_SAMPLE_2	IncFormLen	SkipFormLen	PValue	FDR	IncLevel1	IncLevel2	IncLevelDifference			
8	1157	BRCA1	21	8	3377	82	0.000137567	0.081026943	1	0.06	0.94			
9	EXON SKIPPING													
10	ID	GeneID	IC_SAMPLE_2	SC_SAMPLE_2	IncFormLen	SkipFormLen	PValue	FDR	IncLevel1	IncLevel2	IncLevelDifference			
11	4583	ECT2	183	31	159	62	0	0	0.258	0.697	-0.439			
12	7497	YWHAE	1637	254	208	85	0	0	0.99	0.725	0.265			
13	105	ATP5C1	123	23	60	85	4.12E-09	1.38E-05	0.652	0.883	-0.231			
14	13513	SLC38A2	1207	70	196	85	6.61E-09	1.66E-05	0.968	0.882	0.086			
15	12212	FXR1	91	214	170	85	3.64E-07	0.000520931	0.308	0.175	0.133			
16	2197	MFGE8	427	100	227	85	3.58E-07	0.000520931	0.821	0.615	0.206			
17	2504	EIF4G2	3811	246	185	85	2.83E-07	0.000520931	0.822	0.877	-0.055			
18	5853	SRSF2	90	832	175	85	1.94E-06	0.00243031	0.093	0.05	0.043			
19	14275	SNX14	43	89	20	52	8.27E-06	0.009210773	0.796	0.557	0.239			
20	9834	RPL3	2659	37	170	85	1.07E-05	0.010774257	0.992	0.973	0.019			
21	11162	TMEM126B	26	41	202	85	1.52E-05	0.013851416	0.026	0.211	-0.185			
22	7747	DGUOK	60	47	219	85	2.11E-05	0.017648788	0.144	0.331	-0.187			
23	80	CRIM1	166	20	266	85	3.69E-05	0.028443084	0.934	0.726	0.208			
24	10797	CCNL1	276	15	170	85	8.24E-05	0.059029872	0.99	0.902	0.088			
25	1684	RPS24	55	1375	36	85	8.83E-05	0.059031575	0.044	0.086	-0.042			
26	11828	ZFAS1	100	100	104	79	0.00011171	0.070007095	0.584	0.432	0.152			
27	MUTUALLY EXCLUSIVE EXONS													
28	ID	GeneID	IC_SAMPLE_1	SC_SAMPLE_1	IC_SAMPLE_2	SC_SAMPLE_2	IncFormLen	SkipFormLen	PValue	FDR	IncLevel1	IncLevel2	IncLevelDifference	
29	877	YWHAE	1180	1030	1544	970	278	208	2.69E-08	4.00E-05	0.462	0.544	-0.082	
30	1460	FXR1	653	147	418	37	272	170	3.19E-07	0.000238	0.735	0.876	-0.141	
31	659	RPLP0	1715	6272	1213	3598	214	253	1.26E-06	0.000624	0.244	0.285	-0.041	
32	1048	VIM	1883	9796	982	6042	74	158	7.91E-05	0.023517	0.291	0.258	0.033	
33	300	TSFM	34	115	51	59	194	162	7.43E-05	0.023517	0.198	0.419	-0.221	
34	1326	TMEM126B	3	91	26	106	202	243	9.84E-05	0.024393	0.038	0.228	-0.19	
35	2024	CD63	618	1661	669	1409	145	187	0.000252	0.053543	0.324	0.38	-0.056	

Figure 6: MATS Output List. The figure depicts a typical output of MATS analysis in a software with spreadsheets: here are reported the differentially spliced candidates in the comparison Panc-1 *vs.* Panc-1R for exon skipping events (FDR < 10%). Please click here to view a larger version of this figure.

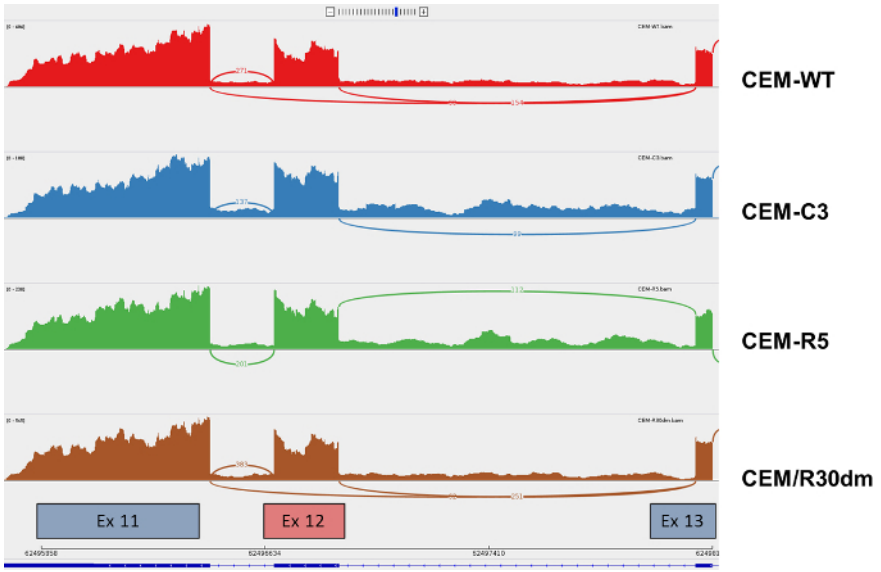


Figure 7: Visualization of Differential Splicing of DDX5 Candidate Gene through IGV Genome Browser. Files with aligned reads (.bam) corresponding to leukemic cells have been uploaded on IGV genome browser and visualized by using Sashimi plots (minimum junction counts value = 10 to visualize significant splicing events). Splice junction counts are represented by connecting lines and a number corresponding to the RNA-seq reads spanning the exons. CEM-WT and CEM/R30dm show skipping counts spanning exon 11 to exon 13 compared to CEM-C3 and CEM-R5 which do not show any exon 12 skipping. Please click here to view a larger version of this figure.

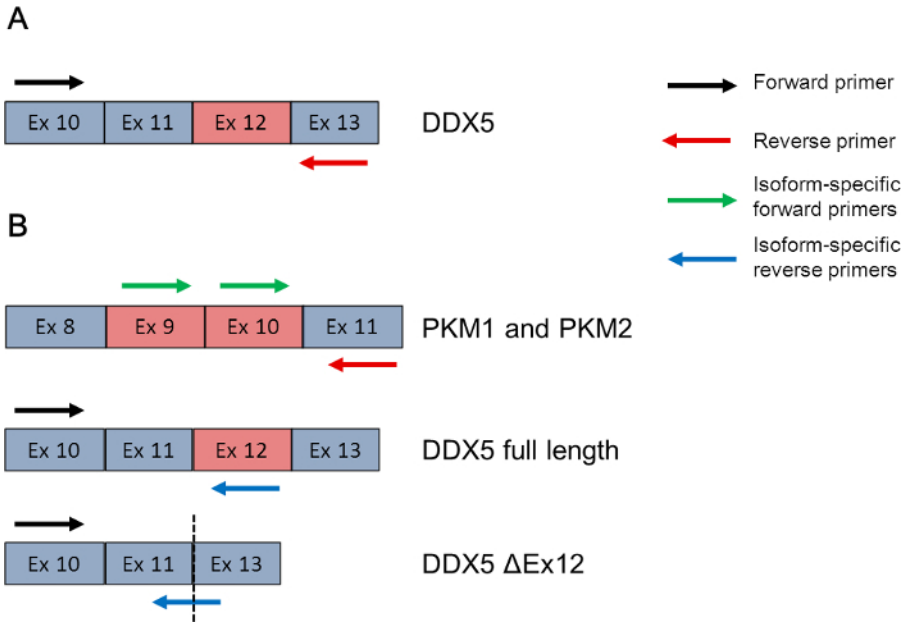


Figure 8: Primer Design for RT-PCR and Cyanine Green qRT-PCR. (A) RT-PCR: primer pairs detecting differential splicing of DDX5 anneal to constitutive exons (exon 10 and exon 13) located upstream and downstream from the alternatively spliced exons. **(B)** qRT-PCR assay: for the relative quantification of transcripts resulting from exon skipping events compared to the canonical transcripts, the reverse primer anneals either within the skipped exon (alternative variant) or to the exon11/exon13 boundary (canonical variant) of DDX5 gene. For the quantification of mutually exclusive exons, the reverse primer anneals to exon 11 common to both isoforms, while the forward primer anneal either to exon 9 (PKM1) or exon 10 (PKM2). Please click here to view a larger version of this figure.

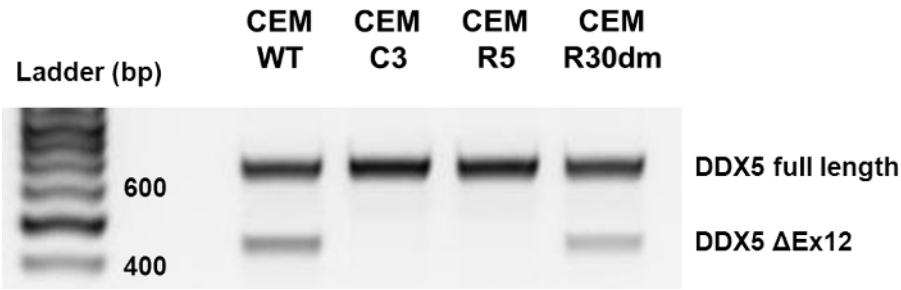


Figure 9: RT-PCR Validation of DDX5 Differential Splicing. The 1% agarose gel shows differential splicing of the DDX5 gene in leukemic cells. The fragment corresponding to DDX5 full length amplicon (650 bp) is amplified in all samples while DDX5 ΔEx12 (430 bp) variant is amplified in CEM-WT and CEM/R30dm samples and the size corresponds to Exon 12 skipping. This is not detected in CEM-C3 and CEM-R5 cells, as determined by MATS analysis. Please click here to view a larger version of this figure.

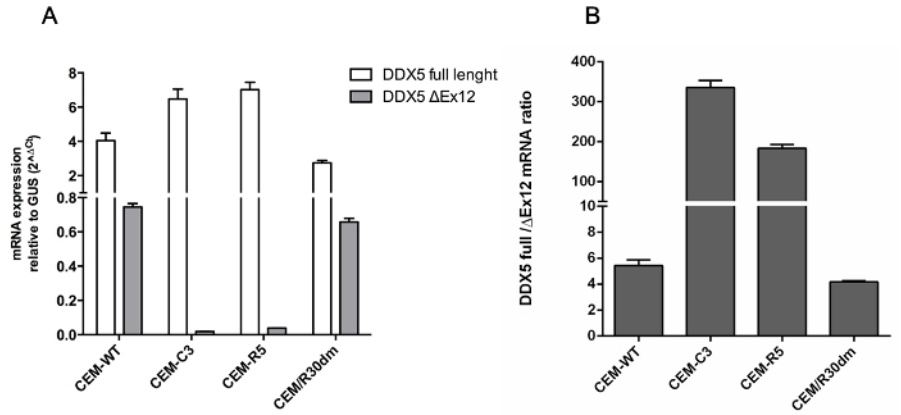


Figure 10: mRNA Expression Levels of DDX5 Splice Variants in CEM Cells. (A) qRT-PCR assay. Mean relative expression levels and standard error of the mean (REL ± SEM) of two independent experiments. **(B)** Ratio of REL (± SEM) of the splice variants. Please click here to view a larger version of this figure.

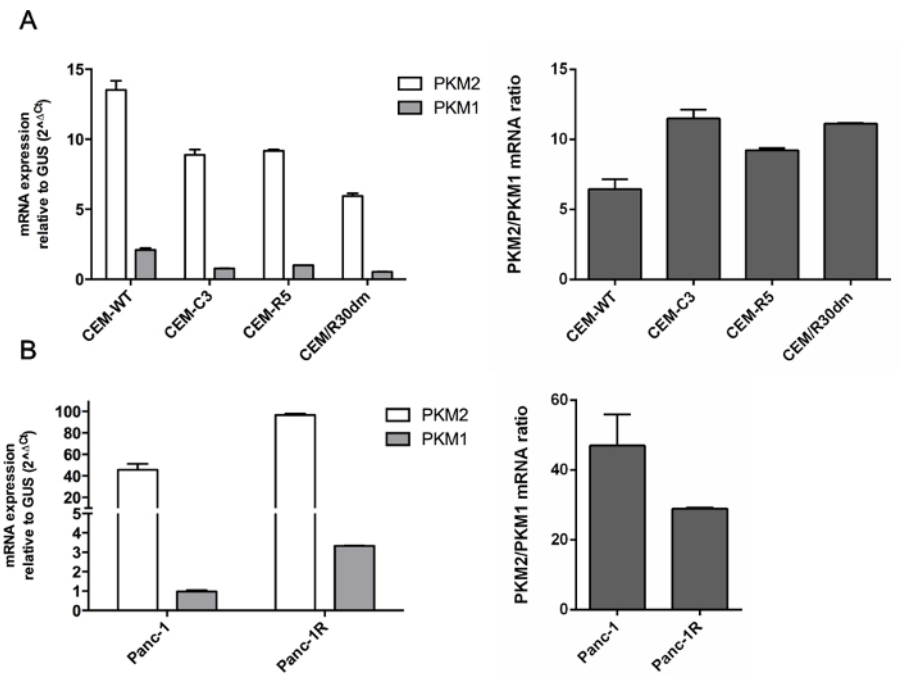


Figure 11: mRNA Expression Levels of PKM Splice Variants in CEM and Panc-1 Cells. (A) qRT-PCR assay. Mean relative expression levels and standard error of the mean (REL ± SEM) of two independent experiments and ratio of REL of the splice variants for CEM cells. **(B)** qRT-PCR assay. Mean REL ± SEM of two independent experiments and ratio of REL (± SEM) of the splice variants for Panc-1 cells. Please click here to view a larger version of this figure.

Discussion

Here we describe a novel approach that combines well-established cytotoxicity screening techniques and powerful NGS-based transcriptomic analyses to identify differential splicing events in relation to drug resistance. Spectrophotometric assays are convenient and robust high-throughput methods to assess drug sensitivity in *in vitro* cancer models and represent the first choice for many laboratories performing cytotoxicity screenings. Troubleshooting as well as possible variations for this method were extensively described elsewhere[4,5].

High-throughput genomic analyses currently used to explore drug resistance mechanisms rely mainly on SNPs detection and differential expression estimation of genes associated with a certain drug-resistant phenotype. In this study, we describe the use of RNA-sequencing methods, together with robust bioinformatics pipelines for precise annotation of mRNA transcripts and detection of differential splicing. A particularly important feature of the described protocol is the ability to identify novel splice variants between two sample groups with distinct drug sensitivity profiles. One of the crucial steps for an accurate and unbiased analysis is the isolation of RNA, which must be of high purity and integrity.

MATS is the software we choose among a series of similar bioinformatics tools available (*e.g.,* Cuffdiff 2, DEXseq, DiffSplice and Splicing Compass) for the detection of alternative splicing[18]. The main key features that make it the preferred option are its superior precision and accuracy as well as the possibility to identify novel events. MATS generates two types of output containing differential splicing analysis: the first is based only on exon junction counts and the second is based on junction counts as well as reads on target. While the latter is preferred for detecting exon skipping events, it is recommended to use the first option for analysis of mutually-exclusive exons as this approach reduces the number of false positive candidates for this particular type of splicing alteration.

Moreover, for analyses focusing on intron retention, alternative 3' and 5' splice site events, a .gtf file with annotated introns should be used. Ultimately, in order to reduce biological and technical variability within sample groups and ensure high true positive rates, it is strongly recommended to sequence at least three replicates[34]. The selection of differentially spliced gene candidates based on MATS output was combined with a validation step using RT-PCR. This is critically important for the selection of true positive variants among a large list of statistically significant candidates. The key factor for an accurate validation is the design of oligonucleotides and the optimization of the PCR reactions according to molecular biology standards. Particular care should be taken when designing primers spanning exon-exon junctions and additional validation steps, such as sequencing of the amplicons by Sanger method, are warranted in order to confirm their specificity.

The differential splicing of DDX5 and PKM transcripts detected by MATS represent two examples of an aberrant splicing related to drug resistance. DDX5 ΔEx12 was not expressed in the GC-resistant cell lines (CEM-C3 and R5), which have been selected after prolonged Dex exposure. DDX5 ΔEx12 was expressed in the parental cell line but also in the subclone CEM/R30dm, which was selected for resistance to the chemotherapeutic agent MTX rather than Dex. In cancer cells, PKM2 was highly expressed compared to its splice variant PKM1, but the ratio PKM2/PKM1 was higher in Dex-resistant cells, as suggested by NGS results. This was not observed for Panc-1 sample compared to its gemcitabine-resistant counterpart and, indeed, this candidate gene was not among the statistically significant events in the MATS analysis. This might reflect the different cell type and mechanism of drug-resistance induced by gemcitabine exposure.

In conclusion, this protocol constitutes a suitable approach for the discovery of splice variants which may underlie drug resistance and can be applied to either leukemic cells[30] or solid tumor cells[31]. A clear limitation is that tumor cell lines capture only a small part of cancer heterogeneity. Moreover, most cell lines have been maintained for many years in monolayers in growth-promoting media. These conditions affect the cellular characteristics, resulting in the selection of subpopulations that differ dramatically from the cells of the primary tumors from which they originate. However, many of the genes that are involved in drug resistance are also involved in other pivotal cell functions such as cell growth and apoptosis that might be affected by long-term culturing in plastic. Therefore, in order to improve the study of drug resistance, more effort should be directed toward the development of novel preclinical models, such as primary cultures and xenografts, that more closely mimic the *in vivo* cancer microenvironment so as to avoid relevant changes in cellular characteristics caused by extended periods of cell culture and culture conditions.[32] Remarkably, our protocol could be applied also to primary cells, using cytotoxicity assays in order to determine *ex vivo* IC_{50} values. Another limitation is that since many mechanisms of resistance exist for each anticancer drug, similar or different mechanisms of resistance could develop in cells exposed to identical but independent treatments. Therefore, a comparative selection strategy should involve parallel selections and analyses, including genetic analyses on splicing variants, of the same parental cells treated with the same chemotherapy agent.

Additional approaches for functional validation should be aimed at overexpressing the splice variant of interest in cell lines or specifically downregulate their expression by using RNA interference or splice-switching oligonucleotides[33].

Disclosures

The authors have nothing to disclose.

Acknowledgements

The authors would like to acknowledge prof. J.J. McGuire, prof. R. Kofler and dr. K. Quint for providing the resistant cell lines used in this work. The study has been founded by grants from Cancer Center Amsterdam (CCA) Foundation (to JC, EG and RS), the KiKa (Children Cancer-free grant for AW) foundation, the Law Offices of Peter G. Angelos Grant from the Mesothelioma Applied Research Foundation (to VEG and EG), Associazione Italiana per la Ricerca sul Cancro (AIRC), Istituto Toscano Tumori (ITT), and Regione Toscana Bando FAS Salute (to EG).

References

1. Gottesman, M.M. Mechanisms of cancer drug resistance. *Annu Rev Med.* **53:** 615-27. (2002).
2. McDermott, M. *et al.* In vitro Development of Chemotherapy and Targeted Therapy Drug-Resistant Cancer Cell Lines: A Practical Guide with Case Studies. *Front Oncol.* **4**,40. (2014).
3. Kaspers, G. J. *et al.* Prednisolone resistance in childhood acute lymphoblastic leukemia: vitro-vivo correlations and cross-resistance to other drugs. *Blood.* **92**(1),259-66, (1998).
4. van Meerloo, J. , Kaspers, G.J., Cloos, J. Cell sensitivity assays: the MTT assay. *Methods Mol Biol.* **731,** 237-45. (2011).
5. Vichai, V., Kirtikara, K. Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nat Protoc.* **1**(3), 1112-6. (2006).
6. Wojtuszkiewicz, A., Assaraf, Y.G., Maas, M.J., Kaspers, G.J., Jansen, G., Cloos, J. Pre-mRNA splicing in cancer: the relevance in oncogenesis, treatment and drug resistance. *Expert Opin Drug Metab Toxicol.* **5**, 673-89. (2015).
7. Liu, S., Cheng, C. Alternative RNA splicing and cancer. *Wiley Interdiscip Rev RNA.* **4**(5), 547-66. (2013).
8. Wojtuszkiewicz, A. *et al.* Folylpolyglutamate synthetase splicing alterations in acute lymphoblastic leukemia are provoked by methotrexate and other chemotherapeutics and mediate chemoresistance. *Int J Cancer.* **138**(7), 1645-56. (2016).
9. Shendure, J., Ji, H. Next-generation DNA sequencing. *Nat Biotechnol.* **10**, 1135-45. (2008).
10. Wang, Z., Gerstein, M, Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* **1**, 57-63. (2009).
11. Ozsolak, F., Milos, P.M. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet.* **2**, 87-98. (2011).
12. Han, Y., Gao, S., Muegge, K., Zhang, W., Zhou, B. Advanced Applications of RNA Sequencing and Challenges. *Bioinform Biol Insights.* **9** (Suppl 1), 29-46. (2015).
13. Khatoon, Z., Figler, B., Zhang, H., Cheng, F. Introduction to RNA-Seq and its applications to drug discovery and development. *Drug Dev Res.* **5**, 324-30. (2014).
14. Zhao, S. *et al.* Comparison of stranded and non-stranded RNAseq transcriptome profiling and investigation of gene overlap. *BMC Genomics.* **16**, 675. (2015).
15. Carrara, M. *et al.* Alternative splicing detection workflow needs a careful combination of sample prep and bioinformatics analysis. *BMC Bioinformatics.* **16** (Suppl 9). (2015).
16. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* **29**, 15-21. (2013).
17. Wang, L., Wang, S., Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics.* **28**, 2184-5. (2012).
18. Hooper, J.E. A survey of software for genome-wide discovery of differential splicing in RNA-Seq data. *Hum Genomics.* **8**, 3. (2014).
19. Shen, S. *et al.* MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. *Nucleic Acids Res.* **8**, e61. (2012).
20. Vargas, I.M., Vivas-Mejìa, P.E. Assessment of mRNA splice variants by qRT-PCR. *Methods Mol Biol.* **1049**, 171-86. (2013).
21. Hala, M., Hartmann, B.L., Böck, G., Geley, S., Kofler, R. Glucocorticoid-receptor-gene defects and resistance to glucocorticoid-induced apoptosis in human leukemic cell lines. *Int J Cancer.* **68**(5), 663-8 (1996).
22. Schmidt, S., *et al.* Glucocorticoid resistance in two key models of acute lymphoblastic leukemia occurs at the level of theglucocorticoid receptor. *FASEB J.* **20**(14), 2600-2 (2006).
23. McCloskey, D. E., McGuire, J.J., Russell, C.A., Rowan, B.G., Bertino, J.R., Giuseppe Pizzorno, G. *et al.* Decreased folylpolyglutamate synthetase activity as a mechanism of methotrexate resistance in CCRF-CEM human leukemia sublines. *The Journal of Biological Chemistry.* **266** (10), 6181-6187, (1991).
24. Quint, K. *et al.* Pancreatic cancer cells surviving gemcitabine treatment express markers of stem cell differentiation and epithelial-mesenchymal transition. *Int J Oncol.* **41**(6), 2093-102. (2012).
25. Lin, S. *et al.* DDX5 is a positive regulator of oncogenic NOTCH1 signaling in T cell acute lymphoblastic leukemia. *Oncogene.* **32**(40), 4845-53. (2013).
26. Mazurek, A. *et al.* Acquired dependence of acute myeloid leukemia on the DEAD-box RNA helicase DDX5. *Cell Rep.* **7**(6), 1887-99. (2014).
27. Calabretta, S. *et al.* Modulation of PKM alternative splicing by PTBP1 promotes gemcitabine resistance in pancreatic cancer cells. *Oncogene.* (2016).
28. Azoitei, N. *et al.* PKM2 promotes tumor angiogenesis by regulating HIF-1α through NF-κB activation. *Mol Cancer.* **15**(1), 3. (2016).
29. Samuels, A.L., Heng, J.Y., Beesley, A.H., Kees, U.R. Bioenergetic modulation overcomes glucocorticoid resistance in T-lineage acute lymphoblastic leukaemia. *Br J Haematol.* **165**(1), 57-66. (2014).
30. Rots, M.G., Pieters R., Kaspers G.J., Veerman A.J., Peters G.J., Jansen G. Classification of ex vivo methotrexate resistance in acute lymphoblastic and myeloid leukaemia. *Br J Haematol.* **110**(4), 791-800. (2000).
31. Funel, N. *et al.* Laser microdissection and primary cell cultures improve pharmacogenetic analysis in pancreatic adenocarcinoma. *Lab Invest.* **88**(7), 773-84. (2008).
32. Gillet, J.P., *et al.* Redefining the relevance of established cancer cell lines to the study of mechanisms of clinical anti-cancer drug resistance. *Proc Natl Acad Sci U S A.* **108**(46), 18708-13. (2011).
33. Kole, R., Krainer A.R., Altman S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat Rev Drug Discov.* **11**(2), 125-40. (2012).
34. Shen, S. *et al.* rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A.* **111**(51), E5593-601. (2014).

