Cancer testis antigens in myelodysplastic syndromes revisited: a targeted RNA-seq approach

Ana María Hurtado López, Tzu Hua Chen-Liang, María Zurdo, Salvador Carrillo-Torneil, Joaquín Panadero, Eduardo José Salido, Victor Beltrán, Begoña Muñá, MariLuz Amigo, Noelia Navarro-Villamor, Rosa Cifuentes, Inés Calabria, Ana Isabel Antón, Raúl Teruel, Manuel Muro, Vicente Vicente, and Andrés Jerez

Hematology and Medical Oncology Department, Hospital Universitario Morales Meseguer, IMIB, Murcia, Spain; Genomics Unit, Health Research Institute La Fe, Valencia, Spain; Department of Hematology, Virgen De La Arrixaca University Hospital, Murcia, Spain; Hematology Unit, H Virgen Del Castillo, Yecla, Spain; Hematology Unit, Hospital Rafael Méndez, Lorca, Spain; Genomics Unit, IMIB-Arrixaca, Murcia, Spain; Immunology Department, Hospital Clínico Universitario Virgen De La Arrixaca, Murcia, Spain; CB15/00055-CIBERER, Murcia, Spain

ABSTRACT
Cancer-testis antigens (CTA) are named after the tissues where they are mainly expressed: in germinal and in cancer cells, a process that mimics many gametogenesis features. Mapping accurately the CTA gene expression signature in myelodysplastic syndromes (MDS) and chronic myelomonocytic leukemia (CMML) is a prerequisite for downstream immune target-discovery projects. In this study, we take advantage of the use of azacitidine to treat high-risk MDS and CMML to draw the CTAs landscape, before and after treatment, using an ad hoc targeted RNA sequencing (RNA-seq) design for this group of low transcript genes. In 19 patients, 196 CTAs were detected at baseline. Azacitidine did not change the number of CTAs expressed, but it significantly increased or decreased expression in nine and five CTAs, respectively. TFDP3 and DDX53, emerged as the main candidates for immunotherapeutic targeting, as they showed three main features: i) a significant derepression on day +28 of cycle one in those patients who achieved complete remission with hypomethylating treatment (FC = 6, p = .008; FC = 2.1, p = .008, respectively), ii) similar dynamics at the protein level to what was observed at the RNA layer, and iii) to elicit significant specific cytotoxic immune responses detected by TFDP3 and DDX53 HLA-A*0201 tetramers. Our study addresses the unmet landscape of CTAs expression in MDS and CMML and revealed a previously unrecognized TFDP3 and DDX53 reactivation, detectable in plasma and able to elicit a specific immune response after one cycle of azacitidine.

Background

Myelodysplastic syndromes (MDS) and chronic myelomonocytic leukemia (CMML) are clonal hematopoietic disorders that compromise the function and number of the blood cells, showing a high rate of progression to acute myeloid leukemia (AML). Recent therapeutic and survival advances in lymphoid malignancies have not been paralleled in the myeloid setting, where there is an urgent need for novel and personalized targets.

Cancer testis antigen (CTA) genes, under normal conditions, are expressed mainly in germ cells, in particular in the testis, and their expression is highly silenced in normal tissue. Because the testis is an immunologically privileged site, tolerance is not established against CTAs. CTAs have been described reactivated in a multitude of cancers and, in addition, their expression appears associated with more aggressive forms of the disease. Until now, the CTAs expression landscape in MDS and CMML has been drawn in a fragmented fashion, with qPCR experiment measuring a discreet set of candidates chosen, mainly, due to findings in solid cancer. Mapping accurately the CTA gene expression signature in MDS and CMML is a prerequisite for downstream target-discovery projects. Global attempts, by means of tag-based sequencing approaches or whole RNA-seq, show either high background level or scarce sensitivity for low transcripts. Using high throughput technology to sequence selected transcriptome areas has allowed researchers to identify and characterize transcripts whose rare or transient expression is not properly annotated with previous approaches.

Fenaux et al., in 2009, first depicted how a drug improved survival in patients with high-risk MDS and CMML. The drug was azacitidine, a chemical analog of cytidine, which exerts its effect through redressing of a dominant pathogenic mechanism in these disorders, hypermethylation. Several studies have shown CTAs derepression with the use of demethylating agents; thus, the current use of azacitidine in MDS and CMML represents an unique opportunity to analyze in vivo this phenomenon for potential therapeutic combinatorial strategies.

Motivated by these observations, the main goals of this study were: i) to screen by means of an ad hoc targeted RNA-seq assay for the expression of a comprehensive list of 210 CTAs, before and after treatment with azacitidine, in a cohort of treatment-naïve MDS and CMML patients; ii) to identify those CTAs whose derepression after azacitidine correlates with a better response; iii) to confirm that increase at the
proteomic level with an easy to use immunoassay; and iv) to assess a specific cytotoxic immune response against those CTAs with a relevant derepression.

**Patients and methods**

**Patients and samples**

Peripheral blood (PB) samples were obtained from 19 patients at diagnosis and on day +28 after the first hypomethylating cycle with azacitidine, between 2015 and 2018. Diagnoses were made according to 2008 and 2017 World Health Organization classification. In each case, informed consent was obtained in accordance with protocols approved by University Hospital Morales Meseguer Institutional Review Board (C.P. – C.I. EST: 47/15) and with the Declaration of Helsinki. The patients were treated according to the MDS and CMMl guidelines of the Spanish group of MDS (GESMD). The response criteria were those established by the International Working Group in 2006.

Azacitidine was administered at a dose of 75 mg/m² subcutaneously day 1 to 5 and days 8 and 9 (5 + 2 schedule). The cycles were planned each 28 days. For patients without reduced baseline counts prior to first cycle therapy, the next cycle was delayed until counts recovery for a maximum of 14 days or a 50% of the previous dose was scheduled otherwise.

**Targeted RNA sequencing (T-RNA-seq)**

It was carried out on the Ion Proton™ platform from Life Technologies on RNA extracted from whole nucleated cells from PB samples on day 0 and +28 of cycle one of azacitidine. In nine patients it was possible to study an additional sample on day +28 of cycle two of azacitidine. The analysis of the quality of the sequences, as well as the mapping and alignment of the readings, was carried out using the Ion Torrent Browser Suite software package. Homo sapiens genome assembly GRCh37 (hg19) was used as a reference genome. On average, we achieved 3.888.308 mapped reads per sample. Libraries were generated using the Ion AmpliSeq™ RNA Library Kit. Our design included 210 CTAs (supplemental Table 1). Granulomonocytic, lymphoid, and CD34+ myeloid compartments were quantified on the PB samples used for day 0 and +28 targeted RNA-seq, using CD11b-FITC, CD13-PE, CD45-PerCP, and CD34-APC performed on a FACScanto II (Becton-Dickinson).

**DNA targeted sequencing**

We designed a TruSeq Custom Amplicon panel (Illumina, Inc. San Diego, CA, USA) including 18 genes recurrently mutated in MDS and CMML. For some genes, known mutation hotspots were targeted; for those with widely dispersed mutations the entire coding sequence was analyzed (supplemental Table 2). The average amplicon size was 151 base pairs and ~99% of the regions were covered on both strands. Paired-end sequencing (2 x 150 bp) was performed with MiSeq v2.2 chemistry, and a mean depth of 997 reads/base within the regions of interest was obtained. The resulting reads were imported into the CLC Genomics Workbench v10 (CLCbio, Arhus, Denmark). The targeted amplicon sequencing workflow was followed for alignment, variant calling, and annotation of filtered mutations.

**Immunoblotting assay**

Plasma samples from patients were collected by PB centrifugation on day 0 and +28 of cycle one of azacitidine. Dilutions (1/20) were separated using 8–12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The proteins were transferred to a PVDF membrane using a semi-dry transfer apparatus. Finally, Western blot analysis was conducted by incubating the membrane with the primary antibodies listed in supplemental Table 3 overnight at 4°C. Densitometry analysis was carried out with Image J (https://imagej.nih.gov/ij/). For normalization, Ponceau staining was used.

**Enzyme-linked immunosorbent assay (ELISA)**

Quantitation of human TFD3 and DDX3 in patients’ plasma at diagnosis and through treatment was performed using Human transcription factor Dp family, member 3 ELISA Kit and Human Probable ATP-dependent RNA helicase DDX3 ELISA Kit, respectively (My Biosource, San Diego, USA) following the manufacturer’s protocols.

**Synthesis of TFD3/DDX3 HLA-A*0201 tetramers**

Regarding TFD3 and DDX3 proteins, the highest MHC-I-binding affinity peptides were chosen using the IEDB consensus methods. HLA-A*-02:01, TFD3-LMWMGMTFGL (LMW295-304), and DDX3- YLMPGFIHL (YLM276-284) peptides were combined to make monomers. HLA-A*0201 containing a biotinylation site and β2-microglobulin were overexpressed in Escherichia coli and purified by fast protein liquid chromatography as previously reported. Purified monomers were biotinylated by BirA ligase and complexed with phycoerythrin (PE), yielding TFD3 and DDX3/HLA-A*0201 tetramers.

**Detection and quantification of TFD3 and DDX3 specific CD8 + T cells**

Freshly isolated PB nucleated cells from two HLA-A*0201 positive and two HLA-A*0201 negative cases (HLA-A*2426 and HLA-A*0111) were analyzed longitudinally by a 6 colors flow cytometry design. Cells were labeled with anti-CD8-PERCP55, anti-CD4-APC, anti-, anti-CD45-PacificBlue, CD57-FITC, CD3-AmCyan (BD Bioscience), and PE-conjugated tetramers (LMW295-304 or YLM276-284). Cells were analyzed using a FACS Canto II flow cytometer with Paint-A-Gate™ (BD Biosciences, Immunocytometry Systems, San Jose, Ca).
Whole RNA-sequencing (W-RNA-seq)

As a secondary aim of this study, we performed whole RNA-seq on a HiSeq4000 sequencer (Illumina, San Diego, CA, USA) in RNA extracted from PB samples on day 0 and +28 of cycle one of azacitidine from 11 out of 19 patients included in the T-RNA-seq experiments. Trimmed reads were mapped to reference genome hg19 with TopHat, and Cufflinks was used for transcript assembly. Expression profile was calculated for each sample and transcript/gene as fragment per kilobase of transcript per million (FPKM) mapped reads achieving a mean of 106 million reads per sample. Libraries were generated using TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, CA, USA).

Statistical analysis

Differential expression between groups was carried out using limma package in R platform. In order to minimize false positives, p values were adjusted by Benjamini–Hochberg multiple comparisons procedure. Additional analysis were performed with the Statistical Package for Social Science (SPSS version 21.0, USA) using Student t-test, Mann-Whitney U test, and chi-square test as appropriate. Significance was determined at a two-sided alpha level of 0.05, except for p values in multiple comparisons, for which Bonferroni correction was applied.

Results

Baseline characteristics of the patients

We performed T-RNA-seq on PB samples from 14 MDS and 5 CMML patients on day 0 and on day +28 of cycle one of azacitidine. A detailed description of the clinical and laboratory characteristics is found in supplemental Table 4. The median age at diagnosis was 70 years old in MDS and 74 in CMML. Patients were classified, according to WHO 2017, as CMML-1 (n = 2), CMML-2 (n = 3), MDS with multilineage dysplasia (MLD) (n = 6), MDS with excess of blasts (EB) type 1 (n = 4) and MDS-EB type 2 (n = 4). We used the IPSS-R and CPSS predictive models to assess MDS (Intermediate, n = 4; High, n = 8; and Very High, n = 2) and CMML (Intermediate-1, n = 2 and Intermediate-2, n = 3) patients’ risk, respectively. MDS patients received, on average, 12 cycles of azacitidine, with a dose reduction in three cases and a median follow-up of 16 months. In the CMML group, a mean of 19 azacitidine cycles were administered with a dose reduction in two cases, and a median follow-up of 27 months. The bone marrow and cytogenetic reevaluation of the disease was carried out after the sixth cycle, resulting in five MDS patient with a complete response (CR), four MDS, and five CMML with a partial response (PR), and 2 MDS with a stable disease (SD). In three MDS patients, a disease progression was suspected and confirmed before the scheduled reassessment. We determined, by flow cytometry, the myeloid, monocytic, CD34+ myeloid and lymphoid subsets on the PB samples used for targeted RNA-seq on days 0 and +28 of the first azacitidine cycle. The median percentage represented by the granulomonocytic compartment (81%) was composed of a median of 69% granulocytes, 11% monocytes, and 0.9% myeloid CD34+ cells.

CTAs expression profile at diagnosis in MDS and CMML patients

Although CTAs expression was initially described as restricted to testis tissue, consecutive studies showed that these genes exhibit heterogeneous gene expression profiles, allowing their classification into testis-restricted, testis/brain-restricted, testis-selective that show additional expression in somatic tissues, and not found in testis but expressed in somatic tissues. In addition, it was also revealed that they were often expressed in different types of cancers.

Studying our T-RNA-seq data in the MDS and CMML cohort, patients were defined by a higher number of CTAs expressed at diagnosis than age-matched controls. A total of 195 CTAs out of the 210 included in T-RNA-seq design, presented expression detected at baseline with, at least, a median of 1 count per million reads (CPM) >1 (Supplemental table 5). In controls, our high throughput low transcript RNA-sequencing detected the expression of 164 CTAs. Supplemental Table 6 shows those CTAs detected in patients but not in controls.

We next classified those CTAs genes detected at diagnosis in T-RNA-seq into Hoffman’s four groups. As only 97 of our 195 detected CTAs were included in their work, those were the CTAs assigned: belonging 22 CTAs to the testis restricted group (21 in X and 1 non-X chromosome), 6 CTAs were classified as testis/brain-restricted (4-X and 2 non-X), 63 CTAs in the testis-selective group (20-X and 43 non-X) and 6 CTAs were classified as not found in testis group (supplemental Figure 1).

When assessing differences between MDS and CMML patients, both entities showed a similar pattern expressing 186 and 187 CTAs at diagnosis, respectively. Nineteen CTAs were differentially expressed reaching statistical significance. Out of the 19, 14 were overexpressed in CMML cases. Strikingly, 8 of these 14 belonged to the MAGE family CTAs (supplemental Table 7).

Azacitidine induces changes at mRNA level in CTAs in MDS and CMML patients

On day+28 of azacitidine, T-RNA-seq was able to detect a similar number of CTAs: 196. One hundred and ninety-one were the same CTAs detected at baseline with five CTAs re-expressed de novo (NOL4, PRM2, RBM46, SPANXN4, and XAGE1C) and four CTAs not detected at this second time point (SPO11, SSX1, THEG, and XAGE3). One hundred and five CTAs raised their expression and eighty-six decreased it.

With the aim to determine clinically significant changes in the CTAs expression pattern in our cohort on day +28 after azacitidine, we clustered the patients according to their response to treatment: i) complete responders (CR), partial responders (PR), any type of response (CR+PR) and non-responders (NR), following the criteria established by IW-MD. Again, the number of CTAs detected was high pre- and post-azacitidine, both in the responders (171 and 173,
respectively) and within the non-responder cohort (175 and 174, respectively). These results promoted us to focus our search for significant differences on the change of expression rather than the raw number of CTAs detected. Figure 1A shows those CTAs which met both a $p < .05$ and a fold change (FC) $> 1$ when contrasting day +28 and day 0 expression of the first azacitidine cycle. Non-responder patients are not represented in the figure as there was no CTA variation which met the established criteria for significance.

In a quantitative and qualitative approach, both the number of CTAs with a statistically significant increment and the magnitude of that increase were higher on day +28 in the 5 MDS patients who achieved a CR in their reevaluation after the sixth cycle. Specifically, the CTAs that showed the greatest upward modification in these CR patients were: TFDP3, DDX53, and MAGEB4 (Figure 1A). TFDP3 and DDX53 behavior in every CR patient was defined by a low-level expression at diagnosis, followed by a significant increase 28 days after azacitidine (Figure 2A, B). Other potential candidates with significantly re-expression in responders vs. nonresponder were: PRAME, TMPRSS12, and DNAJB8. VENTPX1 was discarded as a suitable candidate for proteomic validation as it is described as a pseudogene.

ADAM29 was found to be significantly re-expressed de novo in CR cases, in addition to an increase (FC = 1.28; $p = .03$) on day +28 in PR patients. Similarly, SLCO6A1 (FC = 1.6; $p = .002$) was found to be significantly re-expressed in CR patients, in addition to an increase (FC = 1.28; $p = .03$) on day +28 in patients with any response vs. no responders. Azacitidine CR patients group showed the higher number of CTAs reexpressed on day +28 (Figure 1B).

Figure 1. CTAs with a statically significant change in expression or reexpressed after one cycle of azacitidine. A) CTAs expressed at baseline that show a significant change after the first cycle of azacitidine, grouped by patients response. B) CTAs that do not detected at diagnosis (CPM $< 1$) but which experience the novo expression in patients who reached CR. The inclusion criteria are those described in the text. i) all patients, ii) in the group with any type of response, iii) in the group of patients with partial response and iv) the group of patients who achieved a complete response.

Figure 2. TFDP3 and DDX53T-RNA derepression in CR patients after one cycle of azacitidine and proteomic validation. Radar charts displaying TFDP3 (a) and DDX53 (b) CPM for each of the T-RNA-seq patients grouped by response. The blue line depicts the expression before treatment and the red line on day +28 of the first azacitidine cycle. (c,d) Upper panel showing western Blot experiment on plasma in 4 out of 5 CR showing dynamics in TFDP3 and DDX53 proteins levels through the course of treatment with azacitidine. For UPN15 and UPN5 plasma was available from pretreatment, cycle 1 + 28 and cycle2 + 28 timepoints; for UPN17 plasma was available from pretreatment and cycle 1 + 28 timepoints. Middle panel: ponceau staining was used for normalization. Bottom panel: pretreatment and cycle 1 + 28 linear representation of protein and gene expression level for UPN15, UPN17, UPN16 and UPN5. (e,f) Violin plots illustrating TFDP3 and DDX53 fold change, by means of ELISA, between cycle 1 + 28 and pretreatment sample in 20 patients including 14 T-RNA-seq cases.
TFDP3 and DDX53 plasmatic protein validation and dynamics

We next sought to confirm that the raise in expression of our chosen candidates after azacitidine correlated with an increment at the proteomic level. First, Western blot analysis revealed a correlation between TFDP3 and DDX53 gene expression and the plasmatic protein level change in the CR patients (Figure 2C, D). Conversely, no correlation was found in the Western blot analysis in the case of ADAM29, SLC06A1, MAGEB4, PRAME, TMPRSS12, and DNAJB8 (data not shown).

Second, as we had been able to detect the proteins in patients’ plasma, we decided to explore whether those findings could be assessed with an easy to use assay, susceptible to be adopted in the clinical routine. Thus, a TFDP3 and DDX53 ELISA allowed us to study a cohort of 20 patients. Fourteen T-RNA-seq patients plus six new patients (1 CMML-CR, 2 MDS-NR, and 3 MDS-PR) were included (supplemental Table 8). PB plasma samples on day 0 (sample 1) and on day +28 (sample 2) of the first cycle were measured. We grouped the patients according to their response to azacitidine: i) CR vs rest, ii) CR vs NR, iii) CR vs PR, iv) CR+PR vs NR. Regarding TFDP3, an increase in its plasmatic levels on day +28 vs. pre-treatment was statistically higher in CR patients in comparisons i) and ii) (1.4 FC vs. 1.1 FC, p = .035 1.4 FC vs. 0.8 FC, p = .03, respectively). Though DDX53 FCs were higher after azacitidine in all groups, in particular in CR patients (1.3 FC), no statistical significance was reached (Figure 2F and supplemental Table 9).

Ideally, the correct timing for using an anti-CTA vaccine in combination with an hypomethylating agent in MDS would be when azacitidine achieves the target derepression peak at the protein level, with the potential for eliciting the strongest immune response. Thus, we sought to assess the longitudinal behavior of TFDP3 and DDX53 after cycle 1 and 2 of azacitidine. In CR cases, both TFDP3 and DDX53 showed the most significant increase in gene expression and protein levels after the first cycle (Figure 3A,B,E,F); in contrast, PR and NR patients lacked that peak on day +28 (Figure 3C,D,G,H).

Azacitidine elicits cytotoxic responses against TFDP3 and DDX53 in MDS patients treated with azacitidine

In the light of our previous results, we decided to confirm that an ex vivo TFDP3-LMWMGMTFGL (LMW295-304) and DDX53- YLMGPFIHL(YLM276-283) CD8 + T cells specific response was detectable in MDS patients. In addition, we quantified that specific response in a longitudinal way with specific tetramers in two HLA-A*0201 patients and two non- HLA-A*0201 patients as controls (Figure 4 and supplemental Table 10).

TFDP3 and DDX53-specific CD8 + T cells were detected in our two HLA-A*0201 high-risk MDS patients at diagnosis: UPN53 showed a 5.83% and a 2.93% of total CD3+CD8 + cells, and UPN56 a 20.74% and a 24.89%, respectively. Conversely, that population was not detected or extremely reduced in the two non- HLA-A*0201 high-risk MDS patients, used as controls: UPN55 showed a 1.3% and a 0.6% of total CD3 +/CD8+ cells, and UPN57 a < 0.01 % and a 0.2%, respectively.

Serial analysis during hypomethylating treatment of the two HLA-A*0201 MDS cases showed differences in the quantitative and qualitative dynamics of the CD8 + T cells antigen-specific response. In UPN53, a CR patient after six cycles of azacitidine, both specific responses increased from baseline, reaching a peak for both CTAs on C3 + 28, with a proportional reduction on day +15 (nadir) of each cycle. UPN56, a PR patient after six cycles of azacitidine, showed the highest CD8 + T cells TFDP3 and DDX53-specific response at baseline.

To examine the memory phenotype of the TFDP3 and DDX53-specific CD8 + T-cell responses, we costained with CD57. This analysis showed an increased in the CD57+ subpopulation within the CD8+ tetramers-specific response with repeated cycles, showing an accumulation of memory lymphocytes which exceeded the CD57-counterpart on C2 + 28 for

| TFDP3 | DDX53 |
|-------|-------|
| Gene expression | Protein level | Gene expression | Protein level |
| CR | A | B | E |
| PR/NR | C | D | G |
| F | H |

Figure 3. Gene expression and protein level dynamics through the first and second azacitidine cycles. Gene expression as CPMs from T-RNA-seq assay and protein levels as ng/ml from plasma ELISA. a,b) TFDP3 in two CR patients. e,f) DDX53 in two CR patients. c,d) TFDP3 in two PR patient. g,h) in three PR patients (blue line), one nonresponder patient (orange line) and an stable disease patient (purple line).
both TFDP3 and DDX53 in UPN53. In UPN56, the percentage of CD57+ within the CD8+ tetramers-specific response reached its peak earlier, on day C1 + 15 for DDX53 and on day C1 + 28 for TFDP3 (supplemental Figure 2 and supplemental Table 10).

Cohort somatic recurrent mutations

We performed targeted amplicon sequencing using a panel of 18 genes known to be recurrently mutated in CMML in 16 samples with available DNA (11 MDS and 5 CMML). A total of 30 nonsynonymous somatic variants were identified including typical CMML and MDS associated mutations and frequencies: TET2 (total number of mutations = 8, 22.2% cases mutated), ASXL1 (n = 5, 18.5%) SRSF2 (n = 3, 11.1%), RUNXI (n = 3, 11.1%), TP53 (n = 3, 11.1%), NRAS, and IDH1 (n = 2, 7.4%) y SETBP1, JAK2, EZH2, and ZRSR2 (n = 1, 3.7%). No significant differences in clinical responses to azacitidine or in the expression level of TFDP3 and DDX53 on day 0 and on day +28 were found between TET2 mutated and no mutated cases (p > 0.05). No differences were found when comparing TP53 mutated and wild cases. Interestingly, a statistical trend for a higher increase of TFDP3 and DDX53 expression after the first cycle was noted when comparing abnormal versus normal karyotype cases (supplemental Table 11). The limited number of other genes’ mutated cases precluded us for further comparisons.

T-RNA-seq vs W-RNA-seq

In 11 patients, T-RNA and W-RNA-seq were performed in order to compare both approaches. Samples at diagnosis and on day +28 were sequenced. A total of 203 CTA genes could commonly be annotated in both experiments. T-RNA-seq technique was able to detect expression in 181 from 203 CTAs, either on day 0 and/or day +28. In contrast, the global RNA-seq experiment was only able to detect expression levels in 53 CTAs in the same samples. In 16 CTAs, expression was not detected by any technique.

Due to amplicon design issues, we could not include NY-ESO1 in the targeted panel. Data extracted from de W-RNA-seq assay did not show any increase of expression detected on day +28 of azacitidine (0.03 vs.0.03 CPM for CR; 0.1 vs. 0.02 for PR and 0.02 vs. 0.03 for NR patients). However, both at baseline and on day +28, the number of transcripts detected was extremely low.

Discussion

In this study, we take advantage of the use of hypomethylators to treat high-risk MDS and CMML to draw the CTAs landscape, before and after azacitidine, using targeted RNA sequen- cing directed to this group of low expression transcripts. We have defined the expression pattern of CTAs in MDS and CMML at diagnosis and highlighted the significant increase of expression of two of them, TFDP3 and DDX53 in those patients with a better response to azacitidine. In addition, we confirmed that increase at protein level, the presence of a specific T cytotoxic response against them, and described it through the course of the treatment.

Cancer-Testis antigens are named after the tissues where they are mainly expressed: in germinal and in cancer cells, a process that mimics many gametogenesis features.22 Thus, any gene that exhibits an mRNA expression restricted to the testis and neoplastic cells was termed as a CTA gene. With the advent of more sensitive gene expression techniques, a certain degree of expression was noted in healthy adult tissues; thus, a new and less stringent definition was coined for CTAs (genes with a dominant testicular expression) and a subclassification developed: testis restricted, testis/brain-restricted, testis-selective, and not found in testis.23 With over 200 CTAs annotated and many more candidates identified, it has become increasingly challenging to identify immunogenic or functionally relevant CTAs suitable for cancer vaccine development and/or biomarkers. In our experiment, considering a threshold for a detected expression as an average of CPM ≥ 1, we found that, out of 210 CTAs included in the design, 195 and 196 were expressed at baseline in MDS and CMML, respectively. Twenty-eight days after administering a first dose of azacitidine, we found a statistically significant increase or decrease in nine and five genes, respectively. It has already
been suggested previously that, for some CTAs, the degree of methylation may not be its main regulatory mechanism. In fact, some CTAs show an even more restricted expression pattern, tumor type-specific, suggesting that they require the presence of tumor-characteristic transcription factors in addition to demethylation of the promoters to drive expression.\textsuperscript{24-25} Hybrid, global and targeted, individual cell assays may be of help to elucidate the true weight of methylation in each CTA expression dynamics.\textsuperscript{26} In addition, having whole transcriptome RNA-seq data in 11 T-RNA-seq cases allowed us to compare, as a secondary aim of the work, this technology with our targeted panel of CTAs. In 81\% of genes not expressed in the global RNA-seq experiment, we found a relevant expression when using the directed approach, confirming the latter as the most suitable next generation sequencing (NGS) technique for low transcripts detection.

TFDP3 and DDX53, both CTAs associated to the X chromosome, emerged as the main candidates for immunotherapeutic targeting, as they showed, according to our results, three requirements: i) significant derepression in those patients who achieved CR with hypomethylating treatment, ii) similar dynamics at the protein level to what was observed at the RNA level, and iii) to elicit significant specific cytotoxic immune responses. Previous studies have demonstrated how TFDP3 is expressed in germ tissue and in certain tumors, including brain, esophagus, breast, and melanoma.\textsuperscript{27-29} In addition, its role in the cell has also been partially defined, behaving as an inhibitor of EZF1, a regulator, in turn, of the mechanisms of apoptosis and cell proliferation, dependent on TP53.\textsuperscript{30} The CTA DDX53, also called CAGE, has been attributed the role of promoting tumor angiogenesis in the embryonic period and also during oncogenesis. The expression of this gene has been described in gastric and endometrial cancer and its absence has been reported in leukemia cell lines.\textsuperscript{31} This gene encodes a protein with a DEAD box domain, which suggests that it has RNA helicase activity. Experiments in which recombiant DDX53 was added to the culture medium of endothelial cells showed an increase in their capacity of invasion. The protein is located mainly in the nucleus, but recent data indicate that it can be secreted into the extracellular space.\textsuperscript{32} We have confirmed in this work its presence in plasma samples. In addition, DDX53 is considered a gene absolutely restricted to germine or tumor tissue, which makes it even more promising in terms of its immunogenic potential. The final aim of our line of research would be to target those azacitidine-derepressed CTAs with immunotherapy. Ideally, the correct timing for combination would be when azacitidine derepression peak is manifested at the protein level, with the potential for eliciting the strongest immune response. In our work, the increase of TFDP3 and DDX53 in plasma was reached after only one cycle of azacitidine in CR patients. This finding would mean a significant advantage for therapeutic purposes as they suggest that the correct timing for combination with an anti-CTA vaccine would be early during the azacitidine treatment.

Interest on CTAs relies, in the therapeutic setting, on their capacity to trigger a significant immune response. To be significant in the clinical setting, it is likely that frequencies of antigen-specific T-cells need to reach over 1\% of circulating CD8 T-cells.\textsuperscript{33} Very recently Griffiths et al. showed that an HLA unrestricted NY-ESO-1 vaccine was able to induce an antigen specific immune response in MDS cases.\textsuperscript{34} The robust immune responses against DDX53 and TFDP3 showed by two patients with high-risk MDS provides a building block for considering further steps in the development of CTAs-based vaccination. In addition, the longitudinal study allowed us to delineate that specific cytotoxic response dynamics: with a repetitive variation through the first cycles: reaching its maximum on day +28 and its minimum on day +15. One could speculate that, as the weakest immune response are found at each cytopenic nadir, the reduction of myeloid MDS cells in PB determines less CTA protein circulating and less capacity to provoke the cytotoxic response. The physiological specific response to derepressed TFDP3 and DDX53 reached its peak after three cycles in the CR patient and after one cycle for TFDP3 in the PR patient. In the latter no increase in the specific response was observed for TFDP3 from baseline. Our CR patient whose specific CD8 + T cell tetramer response was assessed, follows the same longitudinal pattern described in the longitudinal T-RNA-seq CR patients TFDP3 and DDX53 gene expression: very low levels at baseline and a significant increase after one cycle. Besides, containing with CD57 allowed us to described how TFDP3 and DDX53-specific memory CD8 +/CD57+ cells accumulated through cycles of azacitidine suggesting an exhaustion of this physiological response. Regarding a potential priming of this response, current clinical trials involving the use of vaccines directed against CTAs have evolved substantially in their design compared to a decade ago, both in the discovery and synthesis of peptides and in the formulation of vaccines. The use of customized peptides, bacterial vectors, or combined treatment, both with mTOR inhibitors and demethylating agents, is now available.\textsuperscript{35} In addition, the generation of dendritic cells ex vivo and their loading with CTAs has successfully improved the anti-cancer immunity in patients.\textsuperscript{36} However, the prominent positioning of the immunotherapy in the treatment of cancer in the last decade has been prompted by the genetic engineering of T lymphocytes. Regarding CTAs, T lymphocytes modified against specific epitopes of NY-ESO-1 have been shown to improve clinical responses and overall survival in patients with refractory melanoma and synovial sarcoma.\textsuperscript{37,38} At present, a clinical trial has been launched, recruiting patients with AML and MDS to be treated with modified T lymphocytes against PRAME and NY-ESO-1. (NCT02494167 H-36346).

Three are the main limitations of our study. First, the T-RNA-seq design did not allow us to include NY-ESO-1 in the panel. NY-ESO-1, PRAME, and the MAGE family are those CTAs against which cytotoxic T cells mediated immune responses have been documented in hematological malignancies.\textsuperscript{413} We did not find a significance change in NY-ESO-1 levels after one cycle of azacitidine using data from our 11 patients with a longitudinal whole RNA-seq analysis. However, as we have shown in this study, the scarce sensitivity of this NGS technique for low transcripts together with the lower number of patients preclude us of reaching robust conclusions about NY-ESO-1 dynamics. Second, the number of patients followed with T-RNA-seq longitudinally (n = 19) and those whose HLA typing made them suitable for the tetramer specific assay (n = 2) are only enough to postulate TFDP3 and DDX53 fold change levels after one cycle of azacitidine as response biomarkers and to confirm the existence of a relevant specific response, but not for establishing definitively the specific CD8 + T-cell dynamics. The inclusion of high-grade disease (where azacitidine is approved) and CMML cases explains, in part,
that the unfractonated PB samples used for CTAs targeted RNA-seq were composed mainly by the granulomonocytic compartment (81%), but we acknowledge the potential bias of the RNA amount from the lymphoid subset as a limitation of our work.

Conclusions

In summary, the prospective nature of patient recruitment and clinical events annotation, and the use of an ad hoc high-throughput RNA-Seq experiment allowed us to reach translational outcome-driven associations within a comprehensive CTA list. Our study addressed the unmet landscape of CTAs expression in MDS and CMML and revealed a previously unrecognized increase of TFDP3 and DDX53 mRNA and protein levels following one azacitidine cycle in MDS and CMML. An increase that elicited a specific cytotoxic response against MDS cells which TFDP3 and DDX53 peptides, pointing us toward a new combinatorial immunotherapeutic approach to explore in these patients.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Funding

This work was supported by grants from the BBVA Foundation (PI 25414/15), Fundación Séneca de la Región de Murcia (19454/PI/14) and Instituto Nacional de Salud Carlos III (PI16/01302) (PI19/00374), and co-funded by the European Union.

ORCID

Ana María Hurtado López http://orcid.org/0000-0002-3235-3690
Manuel Muro http://orcid.org/0000-0001-9987-0994

Author contributions

Contribution: A.M.H.L., T.Z.CH., M.Z., N.NV., R.C., A.I.A., and R.T. performed experiments; A.M.H.L., J.P., I.C., and A.I. analyzed results; A.M.H.L and A.I designed the research and wrote the paper; E.I.S., V.B., B.M., M.L.A., and V.V. provided clinical data and samples; A.M.H.L., T.H.C.L., M.Z., J.P., E.I.S., V.B., B.M., M.L.A., N.NV., R.C., I.C., A.I.A., R.T., M.M., V.V., and A.I gathered data, edited, and approved the final manuscript.

Availability of data and materials

The RNA-seq datasets supporting the conclusions of this article are accessible through accession number GSE102312. For further information on original data, contact the Corresponding Author.

Competing interests

The authors declare that they have no competing interests.

Ethics approval

Informed consent was obtained in accordance with protocols approved by University Hospital Morales Meseguer Institutional Review Board (C.P. - C.I. EST: 47/15) and with the Declaration of Helsinki.

Abbreviations

| Abbreviation | Description |
|--------------|-------------|
| CPM          | counts per million reads |
| CR           | complete response |
| CR+PR        | any type of response; CTA: Cancer-Testis antigens |
| ELISA        | Enzyme-linked immunosorbent assay |
| FC           | fold change |
| FPKM         | fragment per kilobase of transcript per million |
| GESMD        | Spanish group of MDS |
| MDS          | Myelodysplastic syndromes |
| MDS-E/B      | MDS with excess blasts |
| MLD          | MDS with multilineage dysplasia |
| NGS          | next generation sequencing |
| NR           | non-responders |
| PB           | Peripheral blood |
| PR           | partial response |
| RNA-seq      | RNA sequencing |
| SD           | stable disease |
| SDS-PAGE     | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| SSPP         | Statistical Package for Social Science |
| T-RNA-seq    | Targeted RNA sequencing |
| W-RNA-seq    | Whole RNA-sequencing |

References

1. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, Bloomfield CD, Cazzola M, Vardiman JW. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016 May;127(20):2391–2405. doi:10.1182/blood-2016-03-643544.
2. Platzbecker U. Treatment of MDS. Blood. 2019 Mar;133(10):1096–1107. doi:10.1182/blood-2018-10-844696.
3. Bart J, Groen HJM, van der Graaf WTA, Hollema H, Hendriks NH, Vaalburg W, Sleijfer DT, de Vries EG. An oncological view on the blood-tests barrier. Lancet Oncol. 2002 Jun;3(6):357–363. doi:10.1016/S1470-2045(02)00776-3.
4. Nishikawa H, Maeda Y, Iishida T, Gnjatic S, Sato E, Mori F, Sugiyama D, Ito A, Fukumori Y, Utsumiomiya A, et al. Cancer/testis antigens are novel targets of immunotherapy for adult T-cell leukemia/lymphoma. Blood. 2012 Mar;119(13):3097–3104. doi:10.1182/blood-2011-09-379982.
5. Rezvani K, Yong ASM, Tawab A, Jafarpour B, Eniafe R, Mielke S, Savani BN, Keyvanfar K, Li Y, Kurlander R, et al. Ex vivo polarization of polyclonal memory CD8+ T-cell responses to PRAME-specific peptides in patients with acute lymphoblastic leukemia and acute and chronic myeloid leukemia. Blood. 2009 Mar;113(10):2245–2255. doi:10.1182/blood-2008-03-144071.
6. Goodyear O, Piper K, Khan N, Starczynski J, Mahendra P, Pratt G, Moss P. CD8+ T cells specific for cancer germline gene antigens are found in many patients with multiple myeloma, and their frequency correlates with disease burden. Blood. 2005 Dec;106(13):4217–4224. doi:10.1182/blood-2005-02-0563.
7. Wang C, Gu Y, Zhang K, Xie K, Zhu M, Dai N, Jiang Y, Guo X, Liu M, Dai J, et al. Systematic identification of genes with a cancer-testis expression pattern in 19 cancer types. Nat Commun. 2016 Jan;7(1):10499. doi:10.1038/ncomms10499.
8. Junghbluth AA, Chen Y-T, Stockert E, Busam KJ, Kolb D, Iversen K, Coplan K, Williamson B, Altorki N, Old LJ, et al. Immunohistochemical analysis of NY-ESO-1 antigen expression in normal and malignant human tissues. Int J Cancer. 2001 Jun;92(6):856–860. doi:10.1002/ijc.1282.
9. Mercer TR, Clark MB, Crawford J, Brunck ME, Gerhardt DJ, Taft RJ, Nielsen JK, Dinger ME, Mattick JS. Targeted sequencing for gene discovery and quantification using RNA CaptureSeq. Nat Protoc. 2014 May;9(5):1009–1009. doi:10.1038/nprot.2014.058.
10. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009 Jan;10(1):57–63. NIH Public Access. doi:10.1038/nrg2484.
Int Hofmann Cheson síndromes GESMD.loid A, Porwit valproate et McSkeane regimens Giagounidis active Mercer J T, Nov;14(1):778. R, 2005 of in the BD, and Peters antigens, Sette Español and E, al. Lübbert Y. Approach doi: gametogenesis Statistical Organization A, N, myeloid the acute ME, and azacitidine False to doi: cancer. Borowitz Y-T, Cohen JR, Feldman W, Ma Death doi: cancer/testis antigens to microtubule-Targeting drugs. J Biol Chem. 2013;288(51):36502–36518. doi:10.1074/jbc.M113.202047.

Baumgaertner F, Rufer N, Devere E, Derre L, Rimoldi D, Geldhof C, Voelter V, Liéard N, Romero P, Speiser DE, et al. Ex vivo detectable human CD8 T-cell responses to cancer-testis antigens. Cancer Res. 2006;66(4):1912–1918. doi:10.1158/0008-5472.CAN-05-3793.

Griffiths EA, Srivastava P, Matsuoka J, Brunnermeier Z, Wang ES, Kocent J, Miller A, Roloff GW, Wong HY, Paluch BE, et al. NY-ESO-1 vaccination in combination with decitabine induces antigen-specific T-lymphocyte responses in patients with myelodysplastic syndrome. Clin Cancer Res. 2018;24(5):1019–1029. NIH Public Access.

Thomas R, Al-Khadairi G, Roelands J, Hendrickx W, Dermine S, Bedognetti D, Decock J. NY-ESO-1 based immunotherapy of cancer: current perspectives. Front Immunol. 2018;9: doi:10.3389/fimmu.2018.00947.

Sabado RL, Bhardwaj N. Cancer immunotherapy: dendritic-cell vaccines on the move. Nature. 2015;519(7543):300–301. doi:10.1038/nature14211.

Robbins PF, Morgan RA, Feldman SA, Yang JC, Sherry RM, Dudley ME, Wunderlich JR, Nahvi AV, Helman LJ, Mackall CL, et al. Tumor regression in patients with metastatic synovial cell sarcoma and melanoma using genetically engineered lymphocytes reactive with NY-ESO-1. J Clin Oncol. 2011;29(7):917–924. doi:10.1200/JCO.2010.32.2537.

Robbins PF, Kassim SH, Tran TNL, Crystal JS, Morgan RA, Feldman SA, Yang JC, Dudley ME, Wunderlich JR, Sherry RM, et al. A pilot trial using lymphocytes genetically engineered with an NY-ESO-1-reactive T-cell receptor: long-term follow-up and correlates with response. Clin Cancer Res. 2015;21(5):1019–1027. doi:10.1158/1078-0432.CCR-14-2708.