Tristetraprolin/ZFP36 regulates the turnover of autoimmune-associated HLA-DQ mRNAs

Laura Pisapia [a1], Russell Hamilton [b1], Vito D'Agostino [c], Pasquale Barba [a], Maria Strazzullo [a], Alessandro Provenzano [c], Carmen Gianfrani [d] and Giovanna Del Pozzo [a2]

[a] Institute of Genetics and Biophysics "Adriano Buzzati Traverso" CNR, Via Pietro Castellino, 111, 80131, Naples, Italy

[b] Centre for Trophoblast Research, Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Site, Cambridge, CB2 3DY

[c] Centre for Integrative Biology, University of Trento, via Sommarive 9, 38123, Trento, Italy

[d] Institute of Protein Biochemistry-CNR, Via Pietro Castellino, 111, 80131, Naples, Italy

1 Co-first author
2 Corresponding author

Abstract

We have previously demonstrated that the expression of HLA class II transcripts is regulated by the binding of a ribonucleoprotein complex that affects mRNA processing. We identified protein components of a complex binding transcripts encoding the HLA-DR molecule. Here we aimed to verify if the same RNA binding proteins interact with 3'UTR of messengers encoding the HLA-DQ isotype Specifically, we focused on the HLA-DQ2.5 molecule, expressed on the surface of antigen presenting cells, and representing the main susceptibility factor for celiac disease. This molecule, encoded by DQA1*05 and DQB1*02 alleles, presents the antigenic gluten peptides to $CD4^+$ T lymphocytes, activating the autoimmune response.

Here, we identified an additional component of the RNP complex, Tristetraprolin (TTP) or ZFP36, a zinc- finger protein, widely described as a factor modulating mRNA stability. TTP shows high affinity binding to 3'UTR of CD-associated DQA1*05 and DQB1*02 alleles, in contrast to lower affinity binding to DQA1*01 and DQB1*05 non-CD associated alleles. Our *in silico* analysis, confirmed by molecular depletion of protein and HLA mRNA quantification, demonstrates that TTP specifically modulates the stability of the transcripts associated with celiac disease. Our work demonstrates, for the first time, that proteins of the RNP complex, affecting the processing of transcripts, interact with allele-specific transcripts.

Introduction

The expression of HLA class II genes is strictly regulated at the transcriptional level by a highly conserved regulatory module, situated 150–300 base pairs upstream of the transcription-initiation site in all HLA class II genes [1]. This regulatory module interacts with several DNA binding factors which in turn bind the HLA class II trans-activator CIITA [2]. We have previously demonstrated that the expression of HLA class II molecules is regulated through a mechanism that coordinates transcription and processing in the context of an "MHCII RNA operon", a functional unit through which the UTRs of different MHC class II transcripts bind the same RNP complex [3, 4]. We identified two RNA binding proteins, EBP1, the ErbB3 binding protein [5], and NF90, the Nuclear Factor 90 [6], interacting with stem-loop secondary structures at 5' and 3' UTRs of HLA-DRA and HLA-DRB1 mRNAs, transcripts coding for HLA-DR heterodimer, whose expression is downregulated by the depletion of both proteins. The finding that a specific protein complex regulates the expression of HLA DR transcripts, allow us to extend our interest to another isotype, HLA-DQ molecule, frequently involved in the self-antigen presentation, typical of many autoimmune pathologies. Specifically, we aim to verify if the differential expression of HLA class II transcripts with disease association[7], might be regulated by specific proteins in the RNP complex.

The HLA-DQ2.5 molecule is strongly associated with celiac disease (CD), a gluten-triggered disorder [8] and type 1 diabetes (T1D) [9], two autoimmune diseases in which the main genetic predisposing factor is represented by HLA-DQ2.5 encoding genes. More than 90% of CD patients express the HLA-DQ2.5 molecule, while the majority of remaining patients carry the HLA-DQ8 molecule. In T1D patients, the presence of the HLA-DQ2.5 molecule is less frequent than HLA-DQ8, and about 3.5 to 10 % of individuals are affected by co-morbidities of CD and T1D. The surface expression of DQ2 and/or DQ8 molecules, on antigen presenting cells (APC), determines the level of self/gluten-antigen presentation and, as consequence, the magnitude of the pathogenic autoimmune $CD4^+$ T cell response that causing organ damage. We recently demonstrated that DQA1*05 and DQB1* 02 genes are more highly expressed than other non-CD-associated alleles in APC from CD patients carrying the DQ2-DR3 genotype in heterozygosis [7]. Furthermore, we found that both DQA1*05 and DQB1*02 mRNAs show a similar stability, although lower with respect to other DQA1*01 and DQB1*05 non-disease associated transcripts. In this work, we aim to unravel this apparent incongruity between high expression

and low stability focusing on the protein components of the ribonucleoprotein complex that specifically bind autoimmune-associated transcripts.

TTP (ZFP36) is a RNA binding protein that preferentially binds to AU-rich regions contained in the 3'UTR of target genes. TTP functions by destabilizing mRNAs encoding for oncogenes for cytokines (as TNF) and chemokines involved in the inflammatory processes, by favouring their degradation and/or preventing their efficient translation[10, 11]. Studies in TTP KO mice demonstrate that TTP not only regulates the primary immune cellular response to innate stimuli, but also regulates the expression of transcripts involved in the secondary response of fibroblasts/macrophages to the TNF stimulation[12, 13].

Here, we show that TTP binds to HLA class II mRNAs and is a functional player of the MHC operon as its downregulation induces the selective increase of CD-associated alleles.

Results

TTP binds to CD-associated and not associated transcripts

To analyze the interaction of RNA binding proteins with 3'UTR of DQA1* and DQB1* transcripts we synthetized DNA templates that will be used for in vitro transcription to obtain the different riboprobes. We used the cDNA of immortalized B cells (B-LCL#5) obtained from a celiac patient carrying DQA1*01-DQA1*05/DQB1*02-DQB1*05 genotype [7] to synthetize the two riboprobes corresponding to 3'UTR of DQA1*05 (3DQA105 riboprobe) and DQB1*02 (3DQB102 riboprobe) mRNAs, encoding the DQ2.5 molecule associated with CD. The cDNA from HOM-2 cell line, carrying the homozygous DQA1*01/ DQB1*05 genotype, was used for the synthesis of two riboprobes corresponding to the 3'UTR of DQA1*01 (3DQA101) and DQB1*05 (3DQB105) mRNAs, encoding DQ5 molecule not associated with disease.

We previously demonstrated [7] that DQA1*05 and DQB1*02 mRNAs, have a similar half-life, but shorter than that of DQA1*01 and DQB1* messengers, transcribed by alleles on the other chromosome. Moreover we have already demonstrated that the MHC class II transcripts turnover is regulated by a RNP complex interacting with 3'UTR, including NF90 and EBP1 RNA-binding proteins. Here we first performed REMSA experiments using cytoplasmic S100 extracts of M14 and B-LCL#5 and new riboprobes. Fig 1 panel A shows the interaction with 3DQA101 (lanes 3 and 9) and 3DQA105 (lanes 6 and 12) riboprobes to the protein complex in M14 and B-LCL#5 extracts , respectively. Then we assessed if two riboprobes bind to recombinant TTP . We observed that 1ug of recombinant TTP is enough to show a clear band of interaction with either 3DQA101 (lanes 4 and 5) and 3DQA105 (lanes 10 and 11) riboprobes. To

confirm that this complex includes the same proteins of the "MHC Operon" we performed a pull-down assay (Fig 1, panel B). *In vitro* transcribed biotinylated riboprobes, 3DQA101, 3DQA105, 3DQB102 and 3DQB105, were incubated with S100 cytoplasmic extract of B-LCL#5 and after pull-down with streptavidin coated beads, proteins interacting with biotinylated riboprobes were detected by western blot analysis. By using anti-NF90, anti-EBP1 and anti-TTP, we demonstrated that all three proteins bind 3'UTR of DQA1* and DQB1* mRNAs, In conclusion, we confirmed that NF90 and EBP1, previously identified as component of complex interacting with 3'UTRs of DRA and DRB1 mRNA [3, 4], interact with DQ riboprobes while TTP a protein modulating messenger turnover, has been first time recognized as part of RNP complex binding the HLA-DQA1*01, DQA1*05, DQB1*02 and DQB1*05 mRNAs,

Analysis of DQA1* and DQB1* 3'UTR sequences

TTP interacts with transcripts via the specific consensus ARE in the 3'UTR of target genes and is influenced by the folding of the mRNA itself. Therefore, we investigated whether differences in the nucleotide sequences may affect RNA folding and AREs distribution along the 3'UTR. We compared the nucleotide sequences of the 3'UTRs of DQA1* and DQB1* mRNAs for CD-associated and non-CD-associated pairs. We first compared the nucleotide sequences of the 3'UTRs of DQA1* 05 and DQB1*02, CD-associated transcripts, with the 3'UTRs of DQA1* 01 and DQB1*05, non-CD-associated transcripts and highlight the differences in Fig 2A.

In the DQA1* alleles we observed 94.1% identity that caused by a small increase in the GC content with the DQA1*01 allele at 0.46 as compared to DQA1*05 at 0.44 (Fig 2A). In the DQB1* alleles, the sequence differences between 3DQB102 and 3DQB105 riboprobes were higher with 90.1% identity, despite the sequence differences the GC content is identical (0.55) (Fig 2A). The sequence differences between alleles are likely to influence the secondary structures of the RNAs, so we performed minimum-free energy structure predictions (Fig 2C), and mapped on the canonical ARE (AUUUA) and half ARE (UAUU) motifs associated with TPP binding [11, 14]. The canonical ARE and half-ARE motifs are present in both 3DQA101 and 3DQA105 riboprobes, in contrast the ARE motif is only found in the 3DQB102 riboprobe (Fig 2B), suggesting the potentiality for TTP to bind to CD-associated transcripts.

To assess RNA secondary structure stability between the DQA1 and DQB1 pairs, we calculated ensembles of statistically sampled structures using Sfold [15] (Supplementary Figure 1). The ensembles represent the full range of structures a sequence can adopt, highlighting substructures that remain constant across all structure conformations. Minimum-free energy structure prediction makes the assumptions that there is a single rigid structure and that the energy parameters are complete. However, it is likely that messenger RNAs are more dynamic and can adopt multiple conformations depending on their environment and RNA binding proteins. These conserved ensemble substructures co-localise with the ARE and half-ARE sequence motif positions in the DQA1* and 3DQB105 mRNAs. The 3DQB102 ARE is outside of a predicted stable substructure, indicating it is likely to be more flexible and adopt a range of structural conformations. The proportion of the 3DQB102 structure with predicted conserved substructures between ensembles and minimum-free energy predictions, and those from FOLDALIGN (see below), is lower than in the other alleles and importantly, does not contain AU–rich elements indicating that TTP may bind in regions with low structural stability.

We also performed structural motif searches between pairs of 3'UTRs with FOLDALIGN [16] to determine if there are any common structural motifs (defined by sequence and secondary structure) that could confer an RNA Operon-like module as previously found [4]. The presence of a common motif present in different structures could indication a common RNA-binding protein binding site. The identified common structural motifs have been mapped onto the structures in Fig 1C. We found structures common to the DQA1 and DQB1 pairs and we postulate this could enable them to co-bind in an RNA Operon as we previously demonstrated for HLA-DR alleles [4]

We performed a more extensive AU/U rich motif search to determine if there are further possible, weaker TPP binding sites in riboprobes and whether there are differences between the CD and non-CD associated allele pairs. Our kmer analysis at a range of motif lengths (di, tri, tetra, penta and hexa nucleotides) shows a clear pattern of enrichment of AU/U containing motifs in the CD associated alleles as compared to the non-disease associated for all kmer lengths (Fig 3A). Mapping the AU/U rich kmers onto the riboprobe sequences revealed the CD associated alleles contain more AU/U rich sites located across the full length of the sequence and therefore are likely to confer more TPP binding through these AU/U motifs (Fig 3B).
In conclusion, our analysis predicts that TTP can bind to HLA-DRA and HLA-DRB1 mRNAs and that the secondary structures assumed by the alleles facilitate the interaction of TTP with the

3'UTR of CD-associated mRNAs. According to the canonical role of TTP, this binding should confer lower stability to transcripts respect to non-CD associated mRNAs.

Knockdown of TTP increases the expression level of CD-associated transcripts

We previously showed a differential expression of CD-associated transcripts, DQA1*05 and DQB1*02, and their lower half-lives in respect to non-CD associated DQ transcripts in B-LCL#5 cells [7]. In order to investigate if the TTP protein is responsible for the DQ mRNAs turnover we depleted the protein by specific siRNAs transfection in M14 and B-LCL#5 cell lines. The knockdown of TTP was first assayed in M14, 48 h after transfection by western blot with anti-TTP (data not shown) where we observed a depletion of 80%. We then determined, by qRT-PCR, the DRA, DRB1* and DQA1* mRNAs level variation and found a 2 to 3-fold increase after TTP depletion (Fig 4) for the three transcripts. As a control for the specificity of the knockdown, we examined the expression of two groups of genes related to our system, the class I HLA-A, B and C genes and the HLA class II transcriptional activator CIITA and observed them to be unaffected. In parallel, we performed TTP knock-down in B-LCL#5 by specific siRNA nucleofection and assessed protein depletion (Fig 5 panel A) after 48 h by immunoblot. Then we assayed the effect of depletion on the HLA-DQ surface expression by flow cytometry analysis and we observed a 70% increase (Fig 5panel B). To evaluate whether the surface DQ variation corresponds to the HLA-DQ mRNAs increase we carried out qRT-PCR. We found a 2.7 and 2.9-fold increase of DRA and DRB1* mRNAs, respectively, while total DQA1* and DQB1* mRNAs increased 3-fold compared to the control (siCtlr) and in respect to the HLA class I mRNAs (Fig 5 panel C). To investigate if TTP affects the expression of all mRNAs in a comparable way, after knock-down, we specifically quantified DQA1*01, DQA1*05, DQB1*02 and DQB1*05 mRNAs. First, we confirmed that the B-LCL#5, transfected with control siRNA, express higher amounts of DQA1*05 and DQB1*02 mRNAs with respect to DQA1*01 and DQB1*05 mRNAs (Fig 5 panel D and C), as previously found [7]. We then measured the amount of all messengers after nucleofection with siTTP and observed differences in the levels of transcripts. DQA1*05 and DQB1*02 mRNA showed a greater increment (3,5-fold increase), as compared to the control, and when compared to DQA1*01 and DQB1*05 mRNA (2-fold increase), either representing the results as mRNAs copy number, in fig 4 panel D, either showing the mRNAs fold variation, in fig 5 panel E.

In conclusion, we demonstrate that the depletion of TTP protein, interacting with 3'UTR, determines an increment of the CD-associated DQA1*05 and DQB1*02 mRNAs greater than that of the non-CD-associated DQA1*01 and DQB1*05 mRNAs.

Discussion

HLA class II molecules are expressed on the surface of professional antigen presenting cells (APC) such as monocytes, dendritic cells, B lymphocytes, macrophages, and of non-professional APC, such as cancer cells, with the role of presenting antigen epitopes to $CD4^+$ T cells. The magnitude of the $CD4^+$ T cell activation and proliferation is related not only to the nature of cognate epitopes, but also to the amount of the antigen-HLA class II complexes expressed on the surface of APC [17]. In this respect, the expression level of HLA class II molecules that restrict the antigenic responses is very important, particularly in autoimmune diseases as the HLA class II encoding genes represent the main genetic risk factor associated with these pathologies.

Previously, we demonstrated similar amounts of DQ2.5 molecules in APC of CD patients either heterozygous or homozygous for DQ2.5 risk genes [7]. As consequence, DQ2.5 heterozygous or homozygous APC induced a similar strength of functional immune response by gluten-reactive $CD4^+$ T cells. The equivalent surface density of DQ2.5 heterodimer on cells from homozygous and heterozygous CD patients was due to the marked expression of DQA1*05 and DQB1*02 risk genes, in respect to the expression of non-disease associated DQA1*01 and DQB1*05 alleles in DQ2.5 heterozygous genotype. This difference in the quantity of transcripts might be explained by haplotype-specific transcriptional regulation or allele-specific mRNA processing [17]. Indeed, the transcripts of both CD-associated DQA1*05 and DQB1*02 genes showed a similar decay kinetic (3hr half-live), but much lower in respect to non-CD-associated DQA1*01, DQB1*03 and DQB1*05 mRNAs (4h half-live) [7]. This difference in transcript turnover led us to hypothesis that the protein components of the ribonucleoprotein complex might differentially bind the autoimmune-associated messengers, and therefore modulate their decay.

Using cell lines carrying appropriate HLA class II genotypes we amplified the 3'UTR sequences of different alleles, specifically CD-associated DQA1*05 and DQB1*02, and non-CD associated DQA1*01 and DQB1*05 alleles. After sequencing, these templates revealed allele specific differences in their 3'UTRs. In comparing the sequences of alleles, we found several differences in GC content and in canonical ARE (AUUUA) and half ARE (UAUU) motifs, and base differences

that influence the secondary structures and binding sites of the RNAs. The predicted RNA structures show no differences in the presences motif containing substructures between DQA1* alleles, but lower structural stability for DQB1*02. The lower stability of this allele influences also the half-life of DQA1*05 allele, as the turnover of two mRNA is co-regulated. In other words, the amount of a messenger establishes the quantity of its partner messenger coupled in the same RNP complex [17]. To confirm these *in silico* findings we carried out a REMSA revealing the interaction of M14 melanoma and B-LCL cytoplasmic extracts with 3'UTR of DQA1*01 and DQA1*05 transcripts. Importantly, this indicates that the RNA binding proteins included in the complex are constitutively expressed by either professional (B-LCL) and/or non-professional antigen presenting cells (M14). Moreover, in these experimental conditions, the 3'UTR of DQB1*02 and DQB1*05 RNAs do not show binding with cytoplasmic extracts. The absence of binding with beta riboprobes may be caused by the radiolabelling of the 3'UTRs affecting their secondary structures. For this reason, we investigated the RNA-protein interaction by an end labelled desthiobiotin based pull-down method that, does not interfere with the RNA structure. Indeed, following the efficient enrichment of the protein-RNA complexes, we assessed by immunoblot, the interaction of two RNA binding proteins, EBP1 and NF90, previously identified in the complex with of DRA and DRB1* 3'UTR [3, 4], with four 3DQA105, 3DQA101, 3DQB102 and 3DQB105 riboprobes analysed in this work.

Based on previous findings that demonstrated a differential turnover of messengers, we evaluated the interaction of TTP (ZNFP36), a zinc finger protein interacting with AU-rich elements (ARE)-containing mRNAs promoting rapid cytoplasmic decay through acceleration of deadenylation rate [18]. We first assessed the interaction of recombinant TTP to 3DQA101 and 3DQA105 riboprobes by REMSA and later demonstrated, by pull-down, the interaction of TTP with 3'UTR of DQA1*05 and DQB1*02 CD-associated mRNAs (3DQA105 and 3DQB102 riboprobes) and with DQA1*01 and 3DQB105 non-CD-associated mRNA (3DQA101 and 3DQB105) in B-LCL. To unravel the TTP protein function, we performed its depletion by specific siRNA and measured the amount of all HLA class II transcripts by qRT-PCR. We observed, as expected, an increase of DRA, DRB1*, DQA1* and DQB1* mRNAs, while HLA class I and CIITA mRNAs were unaffected. Notably, when we quantified the transcripts of different alleles, we found that TTP depletion differently affects each specific messenger. The DQA1*05 and DQB1*02 mRNAs showed 3.5-fold increase while DQA1*01 and DQB1*05 mRNAs showed the lower 2-fold increase. This result clearly indicates that TTP has a role in the modulation of HLA class II transcripts stability and its depletion shows a relative higher increase for the CD-associated respect to non-CD associated HLA messengers. The lower stability of DQA1*05 and

DQB1*02 may be a biological mechanism to balance their high expression, probably caused by events occurring at transcriptional or epigenetic level.

Tristetraprolin has previously been shown to bind via canonical ARE (AUUUA) motifs, present in the upstream region of 3DQA101 and 3DQA105 riboprobes, in addition to the ARE motifs there are also half-ARE (UAUU) motifs present in the 3DQA101 (two copies) and 3DQA105 riboprobes (Figure 1). More recently, TTP has been demonstrated to bind to AU-rich sequences [14] [11], and through our sequence analysis we have demonstrated there is an enrichment of AU-rich motifs with TPP binding potential in the CD-associated alleles as compared to the non-CD associated. These motifs are located throughout the sequences of the CD-associated alleles in contrast the non-CD-associated alleles where they form a cluster between positions 90-160 (approximate) (Figure 2). The sequence and AU/U-rich motif distribution differences between the DQA1 and DQB1 alleles is also likely to have an impact on the secondary structures of the RNAs and their accessibility to TPP binding. Our analysis of the ensembles of predicted structures indicates that, in addition to the sequence motifs, there are differences in the structure stabilities contributing to the observed differences of TPP binding affinities (Supplemental Figure 1). Across all the secondary structure conformations (ensembles), that can be adopted by each of the RNA alleles, we observed that ARE motifs are co-located with the substructures present across all ensemble structures, thus indicating that they are present in the more stable sections of the structures. In the case of 3DQB102, there is an ARE not present in the DQB105 allele, indicating that TPP binding may higher affinity in DQB102, although this ARE is outside of the more stable substructures. In conclusion, our bioinformatics analysis confirmes that CD-associated mRNAs can strongly interact with TTP through their 3'UTR, conferring a more rapid mRNA turnover.

Our work demonstrates for the first time that TTP, by decreasing the stability of transcripts encoding the HLA-DQ2 molecule in professional human APC, may affect the CD- and T1D-related antigen presentation. As both autoimmune diseases are characterized by the production of inflammatory cytokines, such as IFNγ in the target organs upon the activation of pathogenic $CD4^+$ T cells [19], the TTP binding to the mRNA encoding for the HLA DQ2.5 restriction molecules may counterbalance the high expression of this molecule, by downregulating antigen presentation and activation of pro-inflammatory routes. As consequence the function of TTP protein is very relevant to confirm the relationship between autoimmunity and inflammation, either in M14 cell lines in which the low MHC class II genes

expression is related to the role of non-professional APC and in B-LCL in which the high expression is related to the function or professional APC.

Materials and Methods

Cell lines

M14 human melanoma (DQA1*01/ DQB1*05 genotype) and human HOM cell lines (DQA1*01/DQB1*05 genotype) were obtained by ECACC. B-LCL#1 and B-LCL#5 were EBV-transformed B lymphoblastoid cell lines immortalized from PBMCs of celiac patients carrying DQA1*05/DQB1*02 haplotype and DQA1*01-DQA1*05/DQB1*02-DQB1*05 haplotype, respectively [7]. All cell lines were cultured in RPMI 1640 medium supplemented with 10% Foetal Calf Serum (FCS).

DNA templates synthesis and mRNA quantification

The sequences of specific HLA alleles were obtained from the HLA database (http://www.ebi.ac.uk/ipd/imgt/hla/index.html) and the primers used for PCR and qRT-PCR were synthesized by Eurofins and are listed in Table 1. The template for 3DQA105 riboprobe synthesis was previously prepared [3], while the others obtained from retro-transcribed RNA. Specifically, we used cDNA from B-LCL#1 to obtain 3DB102 template and cDNA from HOM cell to synthetize 3DQA101 and DQB105 templates. Total RNA was prepared with the AurumTM Total RNA kit (BIORAD), and 1 ug of RNA was used for reverse transcriptase reactions, performed using an iScriptTM cDNA Synthesis kit (BIORAD). To quantify specific transcripts, we performed qRT-PCR using the Quanti Tect SYBR Green PCR Kit (BIORAD) through the DNA Engine Opticon Real-Time PCR Detection System (BIORAD). Each reaction was run in triplicate in the presence of 0.2 mM primers, and each experiment performed four times. The relative amount of specific transcripts was calculated by the comparative cycle threshold method [20], and GAPDH and b-actin transcripts were used for normalization. All results shown are the mean of at least three independent experiments. Statistical analysis was performed using the unpaired Student's t-test with two-tailed distribution and assuming two sample equal variance parameters. In the figures, a single asterisk corresponds to $P < 0.05$ and double asterisks correspond to $P < 0.01$.

RNA electrophoretic mobility shift assay (REMSA) and pull-down

The riboprobes synthesis and REMSA were performed according to the published protocol [3,

4]. Briefly, the transcription reactions were performed using T7 *in vitro* transcription system (Ambion) in presence of [^{32}P] UTP and riboprobes obtained were used in binding experiments with M14 and B-LCL#5 S100 extract. TTP recombinant proteins was produced as described [21].

For pull-down experiments, riboprobes were end-labeled with desthiobiotin cytidine and used in binding experiments with 60 ug of B-LCL#5 cytoplasmic extract with the Thermo Scientific Pierce Magnetic RNA-protein pull down kit. The riboprobe used as negative controls was the 3'UTR of androgen receptor RNA poly(A) $_{25}$ RNA, provided by the kit.

Desthiobiotinylated target RNAs bound to proteins were captured using streptavidin magnetic beads and following washing and elution, the proteins interacting with RNA were separated by SDS-PAGE and analysed by western blot. We used three different antibodies: anti-DRBP76 (anti-double stranded RNA binding protein 76 or anti-NF90) antibody (BD Biosciences), anti-EBP1 (Abcam), anti-TTP (Tristetraprolin, Millipore), to reveal the presence of proteins in the RNP complex binding 3'UTR.

TTP protein expression, gene silencing and phenotype analysis

The plasmid for recombinant wild-type and mutant (AA) His-tagged TTP proteins (kindly provided by Dr. Tiedje) have been used to transfect HEK293T cells and protein purification using nickel-chelate agarose beads and following a protocol already described[21] (D'Agostino et al., 2015). After imidazole elution, samples were dialyzed and stored at -80°C in a solution made of 20 mM HEPES pH 8, 100 mM NaCl, 3 mM $MgCl_2$, 8% Glycerol.
For TTP depletion we performed gene silencing using a pool of 4 different siRNA provided by Santa Cruz Biothecnology. We used HiPerFect Transfection Reagent (QIAGEN) to transfect M14 cells and $5x10^5$ cells were harvested after 48 h for RNA extraction and after 24 h for protein extraction. B-LCL were transfected by Nucleofector technology with Lonza kit and $5x10^5$ cells were harvested after 24 hrs for RNA extraction and after 48 for protein extraction and flow cytometry analysis. TTP depletion of both transfected cells was assessed by western blot using an anti-TTP antibody. The HLA-DQ cell surface expression was performed by cytofluorimetric analysis using the FACSAria III and DIVA software with FITC mouse anti-human HLA-DQ antibody (BD Biosciences). The quantitation of specific transcripts was performed by qRT-PCR as described in [7], using primers listed in Table 1.

Bioinformatics Analysis

Sequence alignments and identities were calculated using the global Needleman-Wunch algorithm [22]. Minimum Free Energy (MFE) RNA secondary structures were predicted using RNAfold [23] from single sequences. In addition to MFE structure calculations, an ensemble of statistically sampled structures was generated using Sfold [15], for each riboprobe sequence indicating the most structured parts of the molecules common to the predicted ensemble clusters. To identify common structural motifs between the DQA1 and DQB1 sequences we ran Foldalign [24], in a pairwise all-against-all search. Bracket notation secondary structure from RNAfold were rendered into images with selected motifs mapped onto them using FORNA [25].

AU rich motif mapping and enrichment analysis was performed with custom scripts, freely available at https://github.com/darogan/2018_Pisapia_DelPozzo including R code to recreate Figure 2. All possible kmers, for lengths 2,3,4,5 and 6, were created for each of the riboprobe sequences, and the observed versus expected ratios calculated. Subsets of AU rich motifs are highlighted in the plots. We then mapped the same AU motifs onto each of the riboprobe sequences. Nested structures, i.e. overlapping, are indicated by the score on the y-axis of Figure 2.

Acknowledgements

This work was supported by CNR-DSB Progetto Bandiera "InterOmics" 2017 to CG, MS and GDP. RSH is funded by the Centre for Trophoblast Research , University of Cambridge.

Legends

Fig. 1. RNA binding proteins interaction. A. REMSAs experiments performed using 3DQA101 (lane 1) and 3DQA105 (lane 7) riboprobes with M14 (lane 3 and 9) and B-LCL#5 extract (lanes 6 and 12) respectively. The binding of 3DQA101 (lanes 4 and 5) and 3DQA105 (lanes 10 and 11) with 5 and 1 ug of rTTP is showed. **B**. Western blot analysis of biotin pull-down assay carried by using 3DQA101, 3DQA105, 3DQB102 and 3DQB105 riboprobes. The antibodies used for the immunoblot were anti-NF90, anti-TTP, anti-EBP1. Molecular weights are indicated.

Figure 2. Sequence and structure comparison between 3'UTR of DQA1* and DQB1* genes
A. Comparison of the 3'UTR sequences for the DQA1*01 and DQA1*05 alleles indicates a 94.1% sequence identify. There is a slightly higher proportion of GC content in the DQA1*01 allele (0.46) as compared to DQA1*05 (0.44). **B.** The DQB1*05 and DQB1*02 alleles show a lower sequence identity (90.9%), but identical GC proportions (0.55). **C.** Secondary structure prediction reveals the sequence differences between alleles impacts the minimum free energy folding (RNAfold). Common structural motifs from FOLDALIGN are mapped onto the structures as well as canonical ARE and half-ARE motifs.

Figure 3. AU-rich motif analysis of the DQA1* and DBQ1* genes reveals key differences.
A. All possible di-, tri-, tetra-, penta- and hexa-mers are assessed for their observed frequencies against their expected frequencies. AU-rich motifs are coloured, non-AU rich in grey. All AU-rich kmer motifs have higher enrichments in the CD associated DQA1*05 and DQB1*02 alleles as compared to the non-CD associated DQA1*01 and DQB1*05 alleles. **B.** Mapping the AU-rich motifs to their position within the 3'UTR sequences reveals a clear enrichment in AU-motifs throughout the CD associated alleles DQA1*05 and DQB1*02 (Blue) as compared to the non-CD associated DQA1*01 and DQB1*05 alleles (Red). Nested motifs are indicated on the y-axis, for example two overlapping motifs will show a score of 2.

Fig 4. TTP knockdown in M14.Quantization of DRA, DRB1, DQA1 mRNAs fold increase 24 h after transfection with siTTP or siCtrl transfection in M14. HLA class I, and CIITA mRNA do not show variation.

Fig 5. TTP knockdown in B-LCL#5

A. Western blot performed with anti-TTP for the assessment of protein depletion 48 h after silencing with siCtlr or siTTP nucleofection in B-LCL#5 cell line. **B.** Cytofluorimetric analysis of HLA-DQ surface expression, reported as fold change of MFI (Mean Fluorescence Intensity), 48 h after nucleofection with siCtlr or siTTP **C**. Quantization of DRA, DRB1, DQA1, DQB1 and HLA class I mRNAs fold variation 24 h after siTTP or siCtrl nucleofection. **D.** Quantization of DQA1*01, DQA1*05, DQB1*02 and DQB1*05 mRNA 24 h after siTTP or siCtrl nucleofection reported as copy number or **(E)** fold variation.

Figure S1. Sfold Structure Comparison of 3'UTR of DQA1* and DQB1*. Each of the riboprobe sequences was analysed with Sfold to predict statistical ensembles of structures. The ensembles permit many different conformations of the structures and clustering indicating the most likely conformations adopted by each sequence. The minimum-free energy (MFE) structure is also predicted and its parent cluster identified. **A.** 3DQA101, **B.** 3DQB105, **C.** 3DQA105 and **D.** 3DQB102 show the ensemble structures to contain fewer conserved base pairings (grey boxes) than the MFE structure indicating the structure is more dynamic/flexible outside of these regions. This is particularly evident in **D.** 3DQB102 where there is little conservation of base pairings in the large stem as seen in the other structures.

References

1 Ting, J. P. Y., Trowsdale, J. 2002 Genetic control of MHC class II expression. *Cell*. **109**, S21-S33. (Doi 10.1016/S0092-8674(02)00696-7)

2 Reith, W., LeibundGut-Landmann, S., Waldburger, J. M. 2005 Regulation of MHC class II gene expression by the class II transactivator. *Nat Rev Immunol*. **5**, 793-806. (10.1038/nri1708)

3 Corso, C., Pisapia, L., Citro, A., Cicatiello, V., Barba, P., Cigliano, L., Abrescia, P., Maffei, A., Manco, G., Del Pozzo, G. 2011 EBP1 and DRBP76/NF90 binding proteins are included in the major histocompatibility complex class II RNA operon. *Nucleic acids research*. **39**, 7263-7275. (10.1093/nar/gkr278)

4 Pisapia, L., Cicatiello, V., Barba, P., Malanga, D., Maffei, A., Hamilton, R. S., Del Pozzo, G. 2013 Co-regulated expression of alpha and beta mRNAs encoding HLA-DR surface heterodimers is mediated by the MHCII RNA operon. *Nucleic acids research*. **41**, 3772-3786. (10.1093/nar/gkt059)

5 Ko, H. R., Chang, Y. S., Park, W. S., Ahn, J. Y. 2016 Opposing roles of the two isoforms of ErbB3 binding protein 1 in human cancer cells. *Int J Cancer*. **139**, 1202-1208. (10.1002/ijc.30165)

6 Masuda, K., Kuwano, Y., Nishida, K., Rokutan, K., Imoto, I. 2013 NF90 in posttranscriptional gene regulation and microRNA biogenesis. *Int J Mol Sci*. **14**, 17111-17121. (10.3390/ijms140817111)

7 Pisapia L, C. A., Picascia S, Bassi V, Barba P, Del Pozzo G and Gianfrani C. . 2016 HLA-DQ2.5 genes associated to celiac disease risk are preferentially expressed respect to non-predisposing HLA genes: implication for anti-gluten T cell response. *J of Autoimmunity*.

8 Abadie, V., Sollid, L. M., Barreiro, L. B., Jabri, B. 2011 Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annual review of immunology*. **29**, 493-525. (10.1146/annurev-immunol-040210-092915)

9 Noble, J. A. 2015 Immunogenetics of type 1 diabetes: A comprehensive review. *J Autoimmun*. **64**, 101-112. (10.1016/j.jaut.2015.07.014)

10 Brooks, S. A., Blackshear, P. J. 2013 Tristetraprolin (TTP): interactions with mRNA and proteins, and current thoughts on mechanisms of action. *Biochim Biophys Acta*. **1829**, 666-679. (10.1016/j.bbagrm.2013.02.003)

11 Mukherjee, N., Jacobs, N. C., Hafner, M., Kennington, E. A., Nusbaum, J. D., Tuschl, T., Blackshear, P. J., Ohler, U. 2014 Global target mRNA specification and regulation by the RNA-binding protein ZFP36. *Genome Biol*. **15**, R12. (10.1186/gb-2014-15-1-r12)

12 Patial, S., Curtis, A. D., 2nd, Lai, W. S., Stumpo, D. J., Hill, G. D., Flake, G. P., Mannie, M. D., Blackshear, P. J. 2016 Enhanced stability of tristetraprolin mRNA protects mice against immune-mediated inflammatory pathologies. *Proc Natl Acad Sci U S A*. **113**, 1865-1870. (10.1073/pnas.1519906113)

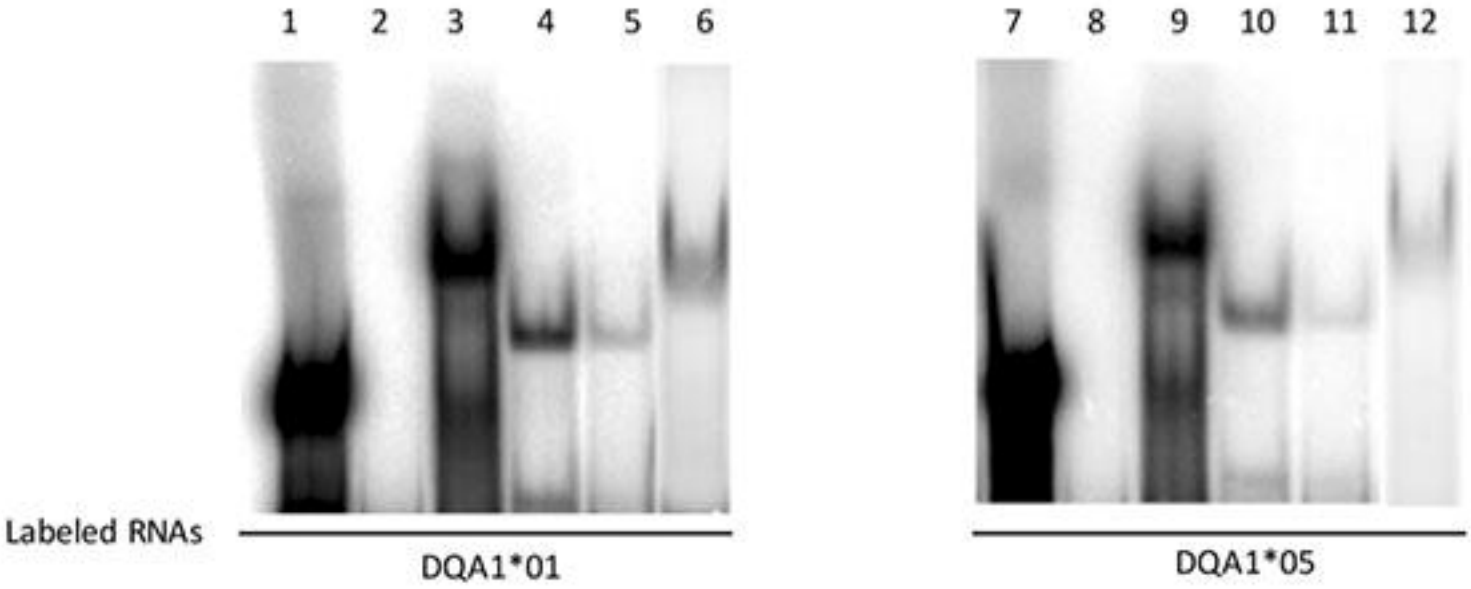
13 Qiu, L. Q., Lai, W. S., Bradbury, A., Zeldin, D. C., Blackshear, P. J. 2015 Tristetraprolin (TTP) coordinately regulates primary and secondary cellular responses to proinflammatory stimuli. *J Leukoc Biol*. **97**, 723-736. (10.1189/jlb.3A0214-106R)

14 Bhandare, S., Goldberg, D. S., Dowell, R. 2017 Discriminating between HuR and TTP binding sites using the k-spectrum kernel method. *PLoS One*. **12**, e0174052. (10.1371/journal.pone.0174052)

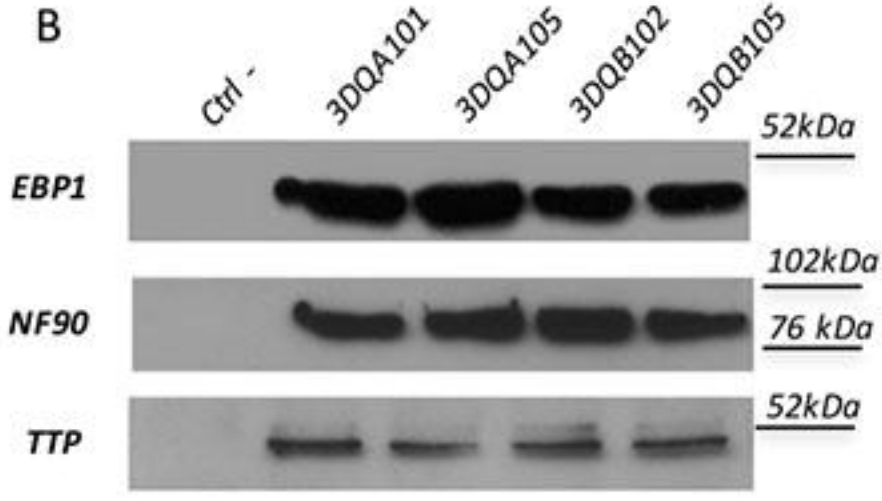
15 Ding, Y., Chan, C. Y., Lawrence, C. E. 2004 Sfold web server for statistical folding and rational design of nucleic acids. *Nucleic acids research*. **32**, W135-141. (10.1093/nar/gkh449)

16 Sundfeld, D., Havgaard, J. H., de Melo, A. C., Gorodkin, J. 2016 Foldalign 2.5: multithreaded implementation for pairwise structural RNA alignment. *Bioinformatics*. **32**, 1238-1240. (10.1093/bioinformatics/btv748)
17 Gianfrani, C., Pisapia, L., Picascia, S., Strazzullo, M., Del Pozzo, G. 2018 Expression level of risk genes of MHC class II is a susceptibility factor for autoimmunity: New insights. *J Autoimmun*. (10.1016/j.jaut.2017.12.016)
18 Patial, S., Blackshear, P. J. 2016 Tristetraprolin as a Therapeutic Target in Inflammatory Disease. *Trends Pharmacol Sci*. **37**, 811-821. (10.1016/j.tips.2016.07.002)
19 Pugliese, A. 2017 Autoreactive T cells in type 1 diabetes. *J Clin Invest*. **127**, 2881-2891. (10.1172/JCI94549)
20 Livak, K. J., Schmittgen, T. D. 2001 Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. **25**, 402-408.
21 D'Agostino, V. G., Adami, V., Provenzani, A. 2013 A novel high throughput biochemical assay to evaluate the HuR protein-RNA complex formation. *PLoS One*. **8**, e72426. (10.1371/journal.pone.0072426)
22 Needleman, S. B., Wunsch, C. D. 1970 A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol*. **48**, 443-453.
23 Lorenz, R., Bernhart, S. H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., Hofacker, I. L. 2011 ViennaRNA Package 2.0. *Algorithms Mol Biol*. **6**, 26. (10.1186/1748-7188-6-26)
24 Havgaard, J. H., Torarinsson, E., Gorodkin, J. 2007 Fast pairwise structural RNA alignments by pruning of the dynamical programming matrix. *PLoS Comput Biol*. **3**, 1896-1908. (10.1371/journal.pcbi.0030193)
25 Kerpedjiev, P., Hammer, S., Hofacker, I. L. 2015 Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams. *Bioinformatics*. **31**, 3377-3379. (10.1093/bioinformatics/btv372)

A


1
2
3
4
5
6
7
8
9
10
11
12
Labeled RNAs
DQA1*01
DQA1*05

B


Ctrl -
3DQA101
3DQA105
3DQB102
3DQB105
EBP1
52kDa
NF90
102kDa
76 kDa
TTP
52kDa

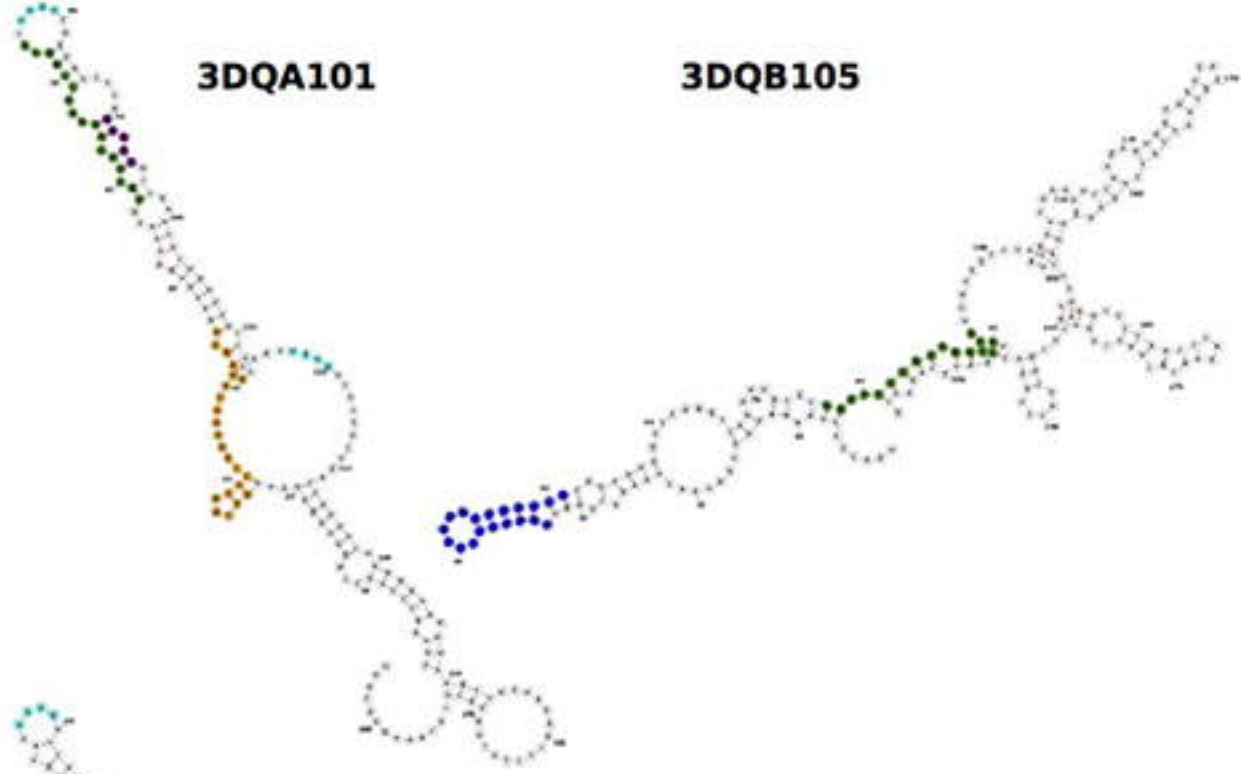
A.

```
3DQA101   UGAAUCCCAUCCUGGAAGGGAAGUGCAUCGCCAUCUACAGGAGCAGAAGAGUGGACUUGC
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
3DQA105   UGAAUCCCAUCCUGGAAUGGAAGUGCAUCGCCAUCUACAGGAGCAGAAGAGUGGACUUGC

3DQA101   UACAUGACCUAGCACUAUUCUCUGGCCCGAUUUAUCAUAUCCCUUUUCUCCUCCAAAUAU
          |||||||||||||| |||| ||||||| ||||||||||||||||||||||||||||| |
3DQA105   UACAUGACCUAGCAUUAUUUUCUGGCCCCAUUUAUCAUAUCCCUUUUCUCCUCCAAAUGU

3DQA101   UUCUCCUCUCACCUUUUCUGUGGGACUUAAGCUGCUAUAUCCCCUCAGAGCUCACAAAUG
          ||||||||||||||| ||||||||||||||  |||||||||  ||||||||||||||||||
3DQA105   UUCUCCUCUCACCUCUUCUGUGGGACUUAAAUUGCUAUAUCUGCUCAGAGCUCACAAAUG

3DQA101   UCUUU
           ||||
3DQA105   CCUUU
```

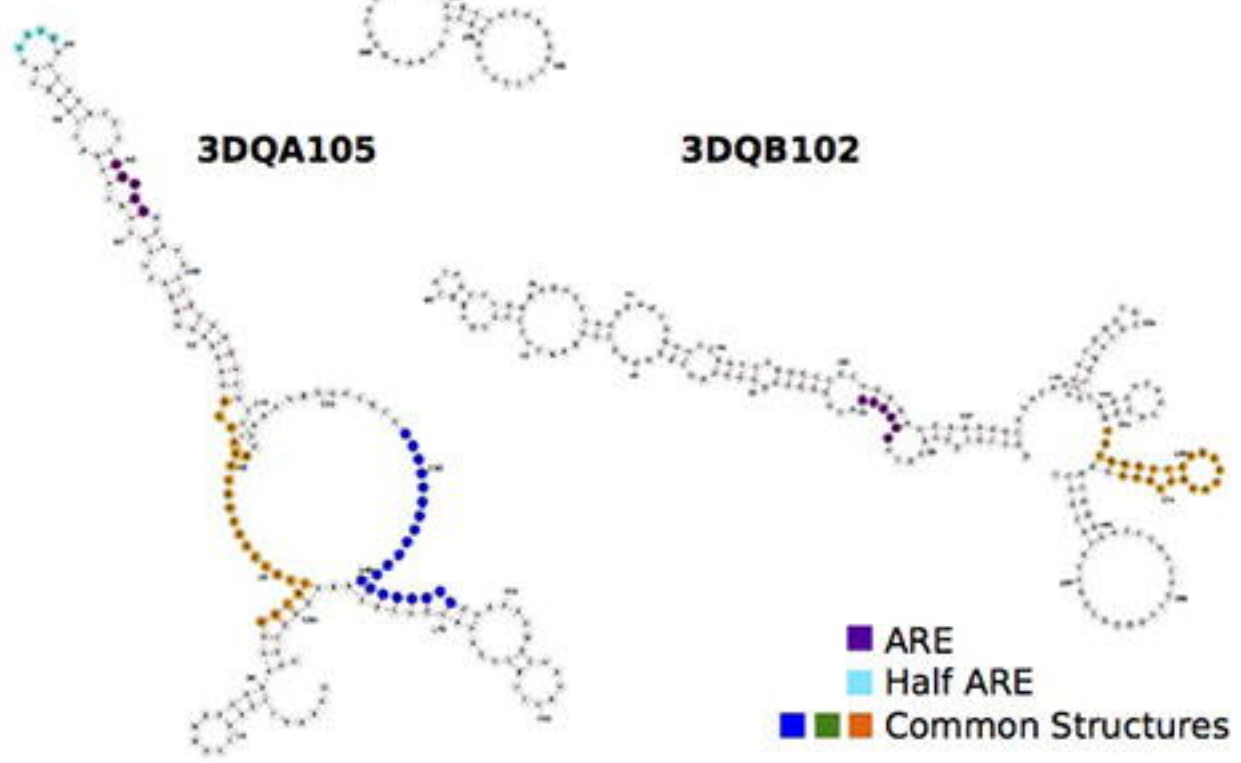
B.

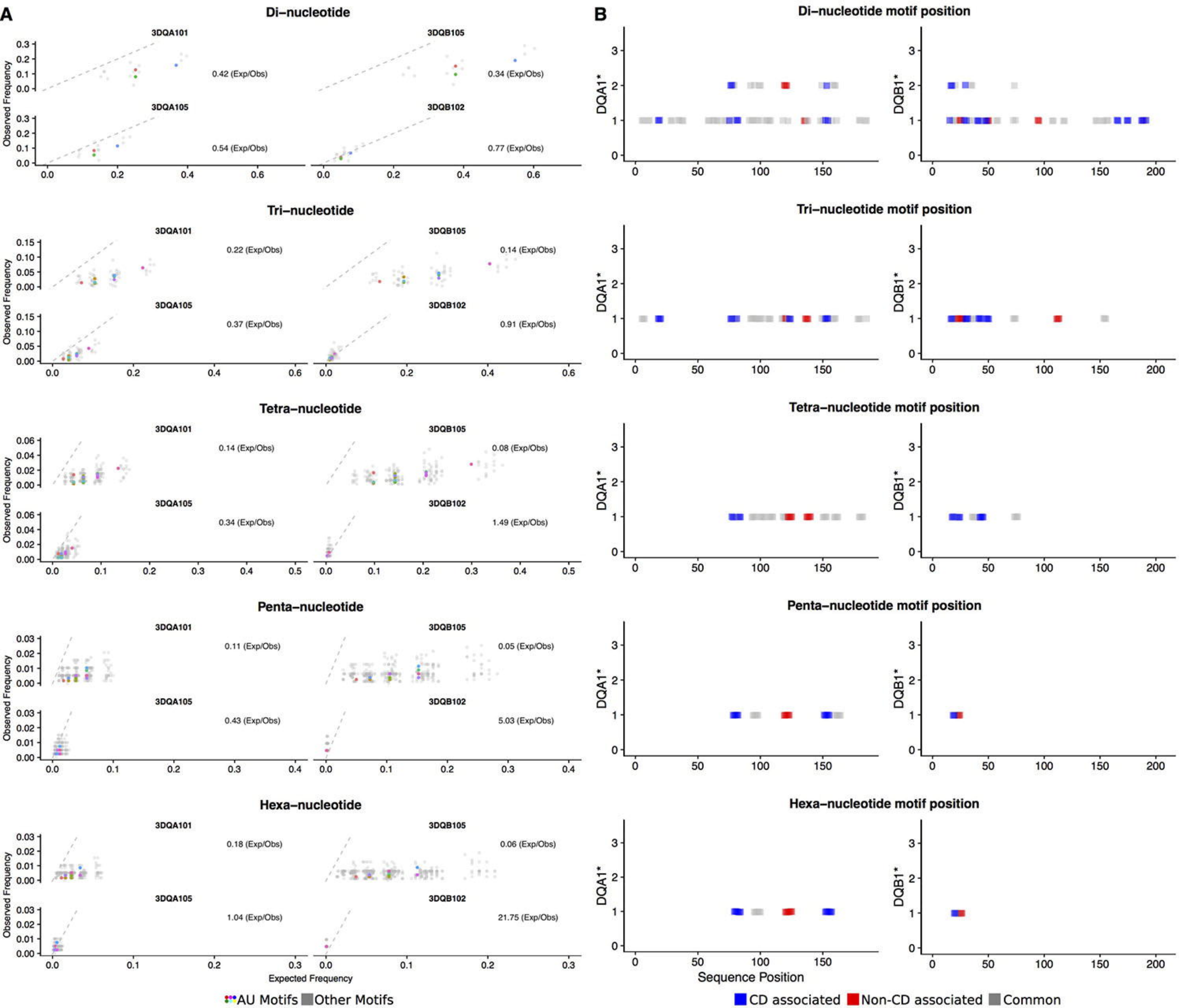
```
3DQB102   UGACUCCUGAGACUAUUUUAACUGGGAUUGGUUAUCACUUUUCUGUAACGCCUGCUUGUC
          ||||||||||||||| ||||||||  || |||||||||||| |||||| | ||||||||||||
3DQB105   UGACUCCUGAGACUGUUUUAACUAAGACUGGUUAUCACUCUUCUGUGAUGCCUGCUUGUC

3DQB102   CCUGCCCAGAAUUCCCAGCUGUCUGUGUCAGCCUGUCCCCCUGAGAUCAGAGUCCUACAG
          |||||||||||||||||||||| ||||||||| |||||||||||||||| |||||||||
3DQB105   CCUGCCCAGAAUUCCCAGCUGCCUGUGUCAGCUUGUCCCCCUGAGAUCAAAGUCCUACAG

3DQB102   UGGCUGUCACGCAGCCACCAGGUCAUCUCCUUUCAUCCCCACCUUGAGGCGGAUGGCUGU
          |||||||||||||| |||||||||||||||||||||||||||||   |||||  |||||||
3DQB105   UGGCUGUCACGCAACCACCAGGUCAUCUCCUUUCAUCCCCACCCCAAGGCGC-UGGCUGU
```

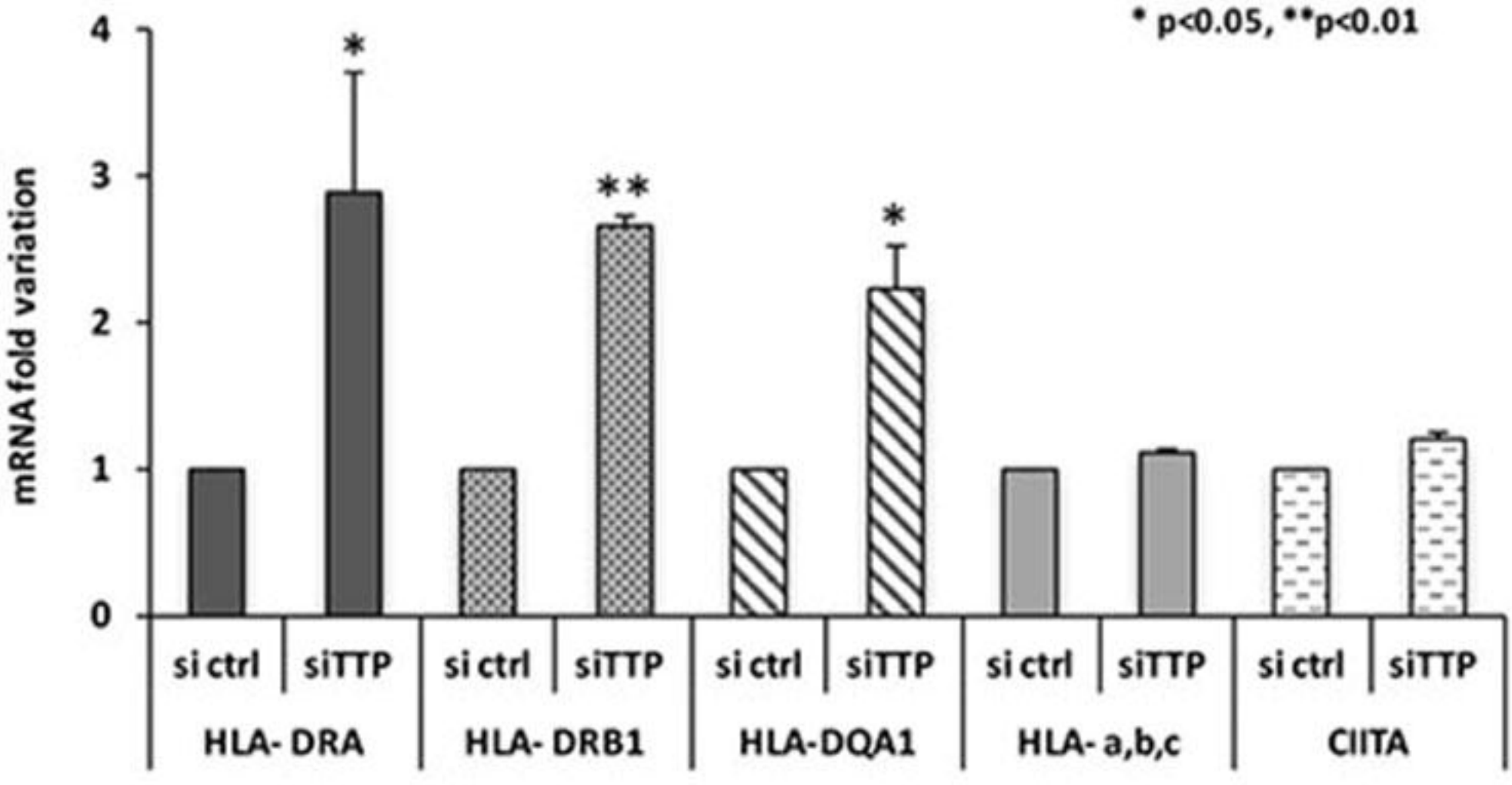
C.

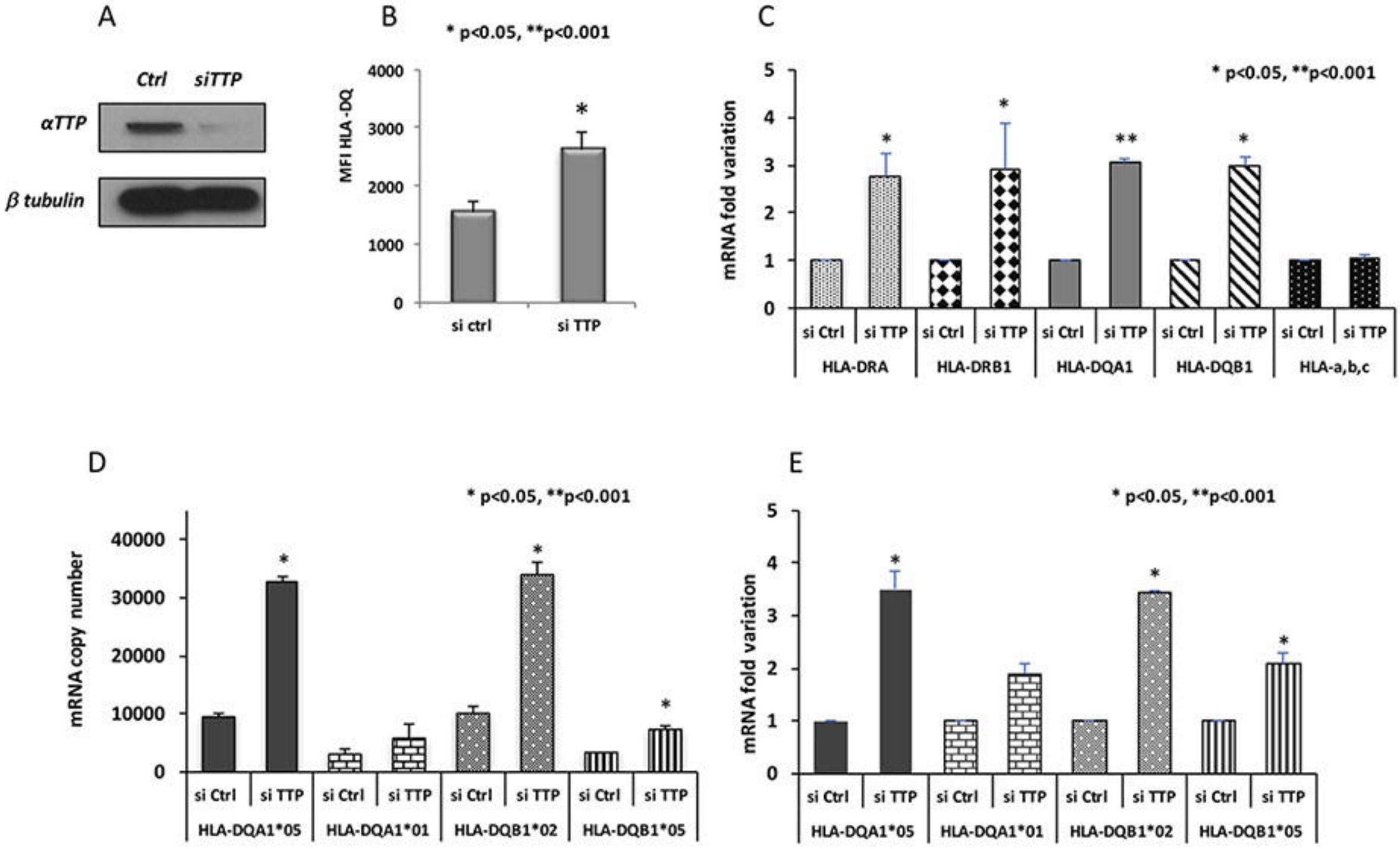


D.




A
Di−nucleotide
3DQA101
0.42 (Exp/Obs)
3DQB105
0.34 (Exp/Obs)
3DQA105
0.54 (Exp/Obs)
3DQB102
0.77 (Exp/Obs)
Tri−nucleotide
3DQA101
0.22 (Exp/Obs)
3DQB105
0.14 (Exp/Obs)
3DQA105
0.37 (Exp/Obs)
3DQB102
0.91 (Exp/Obs)
Tetra−nucleotide
3DQA101
0.14 (Exp/Obs)
3DQB105
0.08 (Exp/Obs)
3DQA105
0.34 (Exp/Obs)
3DQB102
1.49 (Exp/Obs)
Penta−nucleotide
3DQA101
0.11 (Exp/Obs)
3DQB105
0.05 (Exp/Obs)
3DQA105
0.43 (Exp/Obs)
3DQB102
5.03 (Exp/Obs)
Hexa−nucleotide
3DQA101
0.18 (Exp/Obs)
3DQB105
0.06 (Exp/Obs)
3DQA105
1.04 (Exp/Obs)
3DQB102
21.75 (Exp/Obs)
Observed Frequency
Expected Frequency
AU Motifs
Other Motifs
B
Di−nucleotide motif position
Tri−nucleotide motif position
Tetra−nucleotide motif position
Penta−nucleotide motif position
Hexa−nucleotide motif position
DQA1*
DQB1*
Sequence Position
CD associated
Non-CD associated
Common


* p<0.05, **p<0.01
mRNA fold variation
4
3
2
1
0
*
**
*
si ctrl
siTTP
si ctrl
siTTP
si ctrl
siTTP
si ctrl
siTTP
si ctrl
siTTP
HLA- DRA
HLA- DRB1
HLA-DQA1
HLA- a,b,c
CIITA


A
Ctrl
siTTP
αTTP
β tubulin
B
* p<0.05, **p<0.001
MFI HLA-DQ
4000
3000
2000
1000
0
si ctrl
si TTP
C
* p<0.05, **p<0.001
mRNA fold variation
5
4
3
2
1
0
si Ctrl
si TTP
HLA-DRA
HLA-DRB1
HLA-DQA1
HLA-DQB1
HLA-a,b,c
D
* p<0.05, **p<0.001
mRNA copy number
40000
30000
20000
10000
0
HLA-DQA1*05
HLA-DQA1*01
HLA-DQB1*02
HLA-DQB1*05
E
* p<0.05, **p<0.001
mRNA fold variation