

Pan-Cancer chromatin analysis of the human vtRNA genes uncovers their association with cancer biology.

Authors:

Rafael Sebastián Fort ^{1,3} and María Ana Duhagon ^{1,2,*}

¹ Laboratorio de Interacciones Moleculares, Facultad de Ciencias, Universidad de la República, Montevideo 11400, Uruguay; rfort@fcien.edu.uy (R.S.F.).

² Departamento de Genética, Facultad de Medicina, Universidad de la República, Montevideo 11800, Uruguay.

³ Departamento de Genómica, Instituto de Investigaciones Biológicas Clemente Estable, Montevideo 11600, Uruguay.

* Correspondence: mduhagon@fcien.edu.uy/mduhagon@fmed.edu.uy (M.A.D.); Tel.: +598-2-525-8618 (ext. 7237)

Abstract

The vault RNAs (vtRNAs) are a class of 84-141 nt eukaryotic non-coding RNAs transcribed by RNA polymerase III, named for their association with the conserved vault particle, a riboprotein complex whose function remains poorly understood. Of the 4 human vtRNA genes, the three clustered at locus 1, i.e. vtRNA1-1, vtRNA1-2 and vtRNA1-3, are integral components of the vault particle, while vtRNA2-1 is a more divergent homologue located in a second locus. Gene expression studies of vtRNAs in large cancer cohorts have been hindered by the failure of vtRNA sequencing using conventional transcriptomic approaches. However, since the vtRNAs transcription is regulated by DNA methylation, the analysis of the chromatin status of their promoters is a suitable surrogate approach to study their expression. Here we infer the landscape of vtRNA expression in cancer from the genome-wide DNA methylation (Illumina Infinium Human Methylation 450 K BeadChip) and chromatin accessibility (ATAC-seq) data of The Cancer Genome Atlas (TCGA). On average, vtRNA1-1 has the most accessible chromatin, followed by vtRNA1-2, vtRNA2-1 and vtRNA1-3. The correlation of the chromatin status of the vtRNA promoters and the binding sites of a common core of transcription factors stands for their transcriptional co-regulation by factors related to viral infection. Yet, vtRNA2-1 is the most independently regulated vtRNA homologue across tissue types. VtRNA1-1 and vtRNA1-3 chromatin status does not significantly change in cancer, though vtRNA1-3 promoter has repressive chromatin marks in a few cancer types. However, vtRNA2-1 and vtRNA1-2 expression are widely deregulated in neoplastic tissues and is compatible with a broad oncogenic role of vtRNA1-2, and both tumor suppressor and oncogenic functions of vtRNA2-1 depending of tissue contexts. Yet, vtRNA1-1, vtRNA1-2 and vtRNA2-1 promoter DNA methylation predicts a shorter patient overall survival cancer-wide. In addition, gene ontology analyses of co-regulated genes identifies a chromosome 5 regulatory domain controlling vtRNA1-1 and neighboring genes, and epithelial differentiation, immune and thyroid cancer gene sets for vtRNA1-2, vtRNA2-1 and vtRNA1-3 respectively. Furthermore, vtRNA expression patterns are associated with cancer immune subtypes. Finally, vtRNA1-2 expression is positively associated with cell proliferation and wound healing, in agreement with its oncogenic expression profile. Overall, our study presents the landscape of vtRNA expression cancer-wide, identifying co-regulated gene networks and ontological pathways associated with the different vtRNA genes that may account for their diverse roles in cancer.

Keywords: vault RNA; vtRNA1-1; vtRNA1-2, vtRNA1-3, vtRNA2-1; nc886; cancer; TCGA; DNA methylation; chromatin accessibility.

Introduction

The vault RNAs (vtRNAs) are a class of eukaryotic small non-coding RNAs (84-141 nt) transcribed by the RNA polymerase III, which associates to the vault particle (<5 % of the total mass of the vault particle) (Kedersha and Rome, 1986; Kedersha et al., 1991). Although this particle is the largest cellular nucleoprotein complex, its function remains scarcely understood. Most mammals have a single vtRNA gene (i.e., mouse and rat), but the human genome has three copies, annotated as vtRNA1-1 (98pb), vtRNA1-2 and vtRNA1-3 (88pb), which are located in the chromosome region 5q33.1 (the vtRNA1 locus) and transcribe the RNAs that are integral components of the vault particle. Another vault RNA gene in chromosome X, was classified as a pseudogene (vtRNA3-1P) due to the absence of expression in several cell lines and the presence of silencing mutations at the B box element of its promoter (van Zon et al., 2001). Lately, a transcript initially annotated as the human microRNA precursor hsa-mir-886 (Landgraf et al., 2007), was re-classified as another human vtRNA homologue and consequently renamed as vtRNA2-1 (Stadler et al., 2009). VtRNA2-1 is also located in chromosome 5 (5q31.1, locus vtRNA2) at 0.5Mb distance from the vtRNA1 cluster, where is placed between SMAD5 and TGFB1 genes in an antisense direction. Existing evidence indicates that vtRNA2-1 is neither associated with the vault particle nor co-regulated with the vtRNA1 locus (Lee et al., 2011; Stadler et al., 2009). Nevertheless, the realization that only 5-20% of all the cellular vtRNA transcripts are associated to the vault particle suggests additional roles for the majority of vtRNA transcripts, independent of the vault particle (Kickhoefer et al., 1998; Nandy et al., 2009; van Zon et al., 2001).

Most of the current knowledge of vtRNA function focuses on vtRNA1-1 and vtRNA2-1 roles in viral infection and cancer. Functional studies found that vtRNA1-1 and vtRNA1-2 (but not vtRNA1-3) can interact directly with drugs as mitoxantrone, doxorubicin and etoposide (Gopinath et al., 2005, 2010; Mashima et al., 2008). A recent report demonstrated that p62 protein interacts with vtRNA1-1 and inhibits the p62 dependent autophagy *in vitro* and *in vivo* (Horos et al., 2019).

The vault RNAs have also been linked to native immune response, because of their strong upregulation during the infection of Influenza A virus (IAV) and Epstein-Barr virus (EBV) (Amort et al., 2015; Li et al., 2015; Nandy et al., 2009). These studies proposed a viral activation of host vtRNAs as a mechanism to maintain the inhibition of PKR signaling pathway, representing a viral strategy to circumvent host innate immunity.

Initial reports about vtRNA2-1 (locus vtRNA2) revealed that it is abundant in normal tissues while lowly expressed in cancer cell lines from different tissue origin (breast, melanoma, cervix, lung, oral and prostate), which is consistent with a tumor suppressive role in cancer (Lee et al., 2011; Treppendahl et al., 2012). Moreover, vtRNA2-1 was proposed as a new type of ncRNA functioning as a tumor suppressor gene (TSG) that inhibits PKR, and was consequently renamed as “nc886” (Golec et al., 2019; Jeon et al., 2012a, 2012b; Lee et al., 2011). Then, its anti-proliferative and TSG function was described in prostate (Aakula et al., 2015; Fort et al., 2018; Ma et al., 2020), skin (Lee et al., 2019a), gastric (Lee et al., 2014b) and esophageal (Im et al., 2020; Lee et al., 2014a) and cholangiocarcinoma cells (Kunkeaw et al., 2012). Conversely, a pro-proliferative and anti-apoptotic oncogenic role was proposed for vtRNA2-1 in renal (Lei et al., 2017), ovarian (Ahn et al., 2018), thyroid (Lee et al., 2016), cervical (Li et al., 2017) and endometrial (Hu et al., 2017) tissues. A recent finding proposing a vtRNA2-1/PKR loss mediated doxorubicin cytotoxicity introduced a novel view about its contribution to chemotherapy response (Kunkeaw et al., 2019). In addition, a potential TSG role has been put forward for vtRNA1-3 in Myelodysplastic syndrome (MDS) (Helbo et al., 2015).

Apart from their role as full length RNAs, vtRNAs are precursors of small RNAs. Indeed, small RNAs with microRNA like function were demonstrated to derive from vtRNA1-1 (svRNAs) and to repress the expression CYP3A4, a key enzyme of drug metabolism (Meng et al., 2016; Persson et al., 2009). VtRNA2-1 has also been shown to act as a microRNA precursor in different tissues, serving both as TSG in prostate (Aakula et al., 2015; Fendler et al., 2011; Fort et al., 2020), bladder (Nordentoft et al., 2012), breast (Tahiri et al., 2014), colon (Yu et al., 2011), lung (Bi et al., 2014; Cao et al., 2013; Gao et al., 2011; Shen et al., 2018) and thyroid cancers (Dettmer et al., 2014; Xiong et al., 2011) and as oncogene (OG) in renal (Yu et al., 2014), colorectal (Schou et al., 2014) and esophagus cancer (Okumura et al., 2016) through the repression of specific mRNA transcripts in human cancer.

The vtRNAs transcription is controlled by promoter DNA methylation. Different lines of evidence revealed that the epigenetic control of vtRNA2-1 is complex and owns clinical relevance in several tissues solid tumors including breast, lung, colon, bladder, prostate, esophagus, hepatic and stomach cancer (Cao et al., 2013; Fort et al., 2018; Lee et al., 2014a, 2014b; Romanelli et al., 2014; Treppendahl et al., 2012; Yu et al., 2020). Intriguing aspects of the epigenetic regulation of vtRNA2-1 locus comprise its dependence on the parental origin of the allele (Joo et al., 2018; Paliwal et al., 2013), and its sensitivity to the periconceptional environment ((Silver et al., 2015) and subsequent independent studies (van Dijk et al., 2018; Richmond et al., 2018)). In addition, vtRNA1-3 promoter methylation was associated with significantly poor outcome in lower risk myelodysplastic syndrome patients (Helbo et al., 2015).

Currently, transcriptomic sequencing is the benchmark technique to study global RNA expression (Stark et al., 2019). Nonetheless, some classes of RNAs are elusive to the standard transcriptomics due to their stable RNA structure and the presence of modified bases or ends, which impair the cDNA synthesis and/or adapter ligation during the sequencing library preparation (Sendler et al., 2011; Zheng et al., 2015). The vtRNAs belong to this group due to their conserved stable stem/hairpin loop secondary structure. The latter, together with the lack of sequencing of 40-200bp-long RNAs in the conventional transcriptomic studies, which are mostly intended for small and long RNAs, probably delayed their study in comparison with other regulatory RNAs. Nonetheless, since chromatin accessibility plays a critical role in the regulation of gene expression, at some extent the transcription of an RNA can be inferred from the chromatin status of its promoter (Corces et al., 2018; Duren et al., 2017; Liu et al., 2018). Since mounting evidence has shown that vtRNAs expression is tightly controlled by chromatin accessibility, dependent on nucleosome positioning and promoter DNA methylation (Ahn et al., 2018; Fort et al., 2018; Helbo et al., 2015, 2017; Lee et al., 2014a, 2014b; Park et al., 2017; Sallustio et al., 2016; Treppendahl et al., 2012), the chromatin structure is a suitable surrogate marker of vtRNA transcription and could be used as a proxy of their expression.

The expression, regulation, and role of the four human vtRNAs in normal and disease conditions are still poorly understood, and available knowledge indicates that they hold diverse tissue specific activities. The availability of genomic data of large sets of human tissues provided by The Cancer Genome Atlas (TCGA), allows the study of the human vtRNAs across 16 tissue

types and normal/cancer conditions. Here, we analyze vtRNA genes chromatin of the TCGA patient cohort withdrawn from two approaches: the assay for transposase-accessible chromatin followed by NGS sequencing (ATAC-seq) to analyze chromatin accessibility (Buenrostro et al., 2013, 2015; Corces et al., 2018; Thurman et al., 2012) and the Illumina Infinium Human Methylation 450K BeadChip to analyze the CpG methylation of DNA (Moran et al., 2016). This led us to determine the patterns of transcriptional regulation of the four human vtRNAs throughout cancer tissues. We also evaluate the association between vtRNA expression and the patient clinical outcome in all the available cancer types. Finally, seeking for functional discoveries, we analyze vtRNA relations with immune subtypes and transcriptionally co-regulated gene programs. Our study reveals specific patterns of expression for each vtRNA that support previous evidence and poses potential new roles and molecular programs in which they may participate across and intra cancer types, increasing the comprehension of the role of the human vtRNAs in health and disease.

Results

VtRNA genes have different chromatin accessibility at the promoter region.

The Cancer Genome Atlas consortium recently incorporated ATAC-seq data for tumor samples of the GDC Pan-Cancer dataset (385 primary tumors samples across 23 cancer types (Corces et al., 2018)). The ATAC-seq strategy is a genome-wide sequencing profiling approximation to chromatin accessibility using the hyperactive Tn5 transposase that simultaneously cut and ligate adapters preferentially to the accessible DNA regions. The counts of sequencing reads in a particular genomic region provides a direct measurement of chromatin accessibility (Buenrostro et al., 2015; Tsompana and Buck, 2014). For those gene promoters regulated by DNA methylation, ATAC-seq is a more direct approximation to chromatin accessibility than promoter CpG DNA methylation measurements because the latter is well correlated but not an exact measurement of chromatin accessibility. Indeed, the effect of DNA methylation on chromatin structure depends on the gene and the distribution of the methylation sites along it. It is also important to mention that gene expression interpretations based solely on chromatin accessibility are blind to post-transcriptional regulatory events such as processing, localization,

and stability of the RNA transcript. Nevertheless, since the chromatin status of the different vtRNA promoters has been positively correlated with the abundance of their transcripts (Ahn et al., 2018; Fort et al., 2018; Helbo et al., 2015; Lee et al., 2014a, 2014b; Park et al., 2017; Sallustio et al., 2016; Treppendahl et al., 2012), ATAC-seq is likely a good surrogate of vtRNA expression. DNA CpG methylation data is available for 746 normal adjacent (across 23 tissues) and 8403 primary tumors (across 32 tissues) tissue samples of the Pan-Cancer TCGA cohort (Supplementary Table 1 and Supplementary Figure 1). The full data is listed in Supplementary Table 1 and the tissue composition of the three datasets of the cohort, comprising tumor ATAC-seq, normal and tumor DNA methylation, is shown in Supplementary Figure 1. The DNA methylation data has 9 more tissue types and a higher number of samples than the others, while the abundance of some tissues is skewed to twice enrichment in 1-4 specific dataset specific tissue types.

Although several reports demonstrated that the DNA methylation of vtRNA promoters is well correlated to vtRNA transcript expression in various tissues (Ahn et al., 2018; Fort et al., 2018; Helbo et al., 2015, 2017; Lee et al., 2014a, 2014b; Sallustio et al., 2016; Treppendahl et al., 2012), the association between the DNA methylation and chromatin accessibility at their promoters was not previously investigated.

To assess the expression of vtRNAs in cancer tissues, we analyzed a 500 bp region, containing the full vtRNA genes, the proximal promoter region and the RNA polymerase III A and B box elements located downstream of the transcription start site (TSS) (Figure 1). Indeed, Vilalta et al., 1994 demonstrated that the deletion of the 300bp 5'-flanking region of the rat vtRNA transcript largely reduced its transcription greater than 30-fold (Vilalta et al., 1994). As depicted in Figure 1, this region bears epigenetic marks of transcriptionally active chromatin (Li et al., 2007), including histone H3K27ac, H3K4me1, H3K4me3 at the promoters and DNaseI hypersensitivity clusters around the TSS in the cell types compiled by UCSC browser (Goldman et al., 2013). Interestingly, vtRNA2-1 is the only vtRNA immersed in a CpG island and vtRNA1-1 displays the strongest epigenetic marks of transcriptional activity (Figure 1). The analysis of the ATAC-seq values of this 500 bp region of all the TCGA primary tumors reveals that the four human vtRNAs have different levels of chromatin accessibility (Figure 2A). Moreover, the average DNA methylation of the promoters (Figure 2B) in primary tumor samples supports this

finding (32 tissues, Supplementary Figure 1 and Supplementary Table 1). VtRNA1-1 promoter has the highest chromatin accessibility and the lowest DNA methylation, showing the smallest dispersion of values among the samples (Ave. 4.1, 0.6 SD) relative to the other three vtRNAs. Meanwhile, vtRNA1-2 (Ave. 2.9, 1.6 SD), vtRNA1-3 (Ave. 1.7, 1.2 SD) and vtRNA2-1 (Ave. 2.5, 1.4 SD) have broader ranges of ATAC-seq and DNA methylation values (Figure 2, Table 1, Supplementary Table 1 and Supplementary Figure 1). A comparison among the vtRNAs shows that the average chromatin accessibility of their promoters is inversely proportional to their average DNA methylation (Table 1 and Figure 2) (relative ATAC-seq values are vtRNA1-1 (4.1) > vtRNA1-2 (2.9) > vtRNA2-1 (2.5) and accordingly, relative DNA methylation are vtRNA1-1 (0.09) < vtRNA1-2 (0.4) < vtRNA2-1 (0.5)). Yet vtRNA1-3 average low chromatin accessibility is not accompanied by a comparatively denser promoter methylation (values 1.7 and 0.22 respectively). Nevertheless, the intra-sample correlation between the ATAC-seq and DNA methylation for the vtRNAs is between -0.24 and -0.78 (see rs-values in Table 1 and Supplementary Figure 2), suggesting that regardless of the relatively smaller average effect of DNA methylation on vtRNA1-3 promoter accessibility, all the vtRNAs are regulated by DNA methylation to some extent (Table 1). In addition, the CpG DNA methylation negatively correlates with ATAC-seq values in the 500bp of the promoters of the four vtRNAs (Table 1 and Supplementary Figure 2). The strength of the correlation for each vtRNA is higher for the two vtRNAs that have a broader spectrum of DNA methylation, i.e. vtRNA1-2 and vtRNA2-1 ($r_s = -0.74$ and $r_s = -0.78$ respectively) (Table 1 and Supplementary Figure 2B and 2D). Meanwhile, vtRNA1-1 and vtRNA1-3 promoters, which have narrower range of DNA methylation (vtRNA1-1 has also a small range of chromatin accessibility) present a smaller association between both values ($r_s = -0.24$ and $r_s = -0.54$ respectively) (Table 1 and Supplementary Figure 2A and 2C). Technical differences in ATAC-seq and DNA methylation array approaches, as well as naturally occurring non-linear correlation between average DNA methylation and chromatin structure may account for these differences (Corces et al., 2018; Liu et al., 2018). The low ATAC-seq value of the vtRNA3-1P pseudogene (Ave. -1.2, 0.6 SD), represents a proof of concept of the analyses (Supplementary Figure 3 and Supplementary Table 1).

Chromatin accessibility, DNA methylation and transcription factor binding at the four vtRNA promoters are correlated.

Aiming to investigate a possible co-regulation of vtRNA transcription, we determined the pairwise correlations in chromatin accessibility and DNA methylation between the four vtRNA promoters. We found that all pairs of vtRNAs, except vtRNA1-2 and vtRNA2-1, shown positive correlations in both datasets (Figure 3A and 3B) (rs-values 0.23-0.41 for ATAC-seq and rs-values 0.20-0.50 for DNA methylation). The unsupervised clustering of these epigenetic marks indicates that the three vtRNAs clustered at vtRNA1 locus (vtRNA1-1, vtRNA1-2 and vtRNA1-3), which are the vault particle associated vtRNAs, are more similar than vtRNA2-1 (Figure 3A and 3B).

Given that the chromatin accessibility of a DNA region is partially controlled by transcription and chromatin remodeler's factors (Corces et al., 2018; Klemm et al., 2019), we analyzed transcription factors (TFs) occupancy at a ± 3000 pb region centered at each vtRNA gene. Since there is no such study of the TCGA samples, we investigated the CHIP-seq data of the K562 chronic myelogenous leukemia cell line, which has the richest TFs data at ENCODE dataset (Goldman et al., 2013). We observe that the 4 vtRNAs share a common core of 23 TFs (26%). The 3 members of the vtRNA1 cluster share 16 additional TFs (18%), reaching 39 common TFs (44%), while only 2-4 TFs with vtRNA2-1 (Figure 3C and 3D and Supplementary Table 2). We finally asked whether this common core of TFs is linked to a specific biological process. The STRING enrichment analysis of the core of 23 TFs common to all vtRNAs identified HTLV-I infection (ATF1, ATF3, E2F1, EP300, NFATC3, TBP), Hepatitis B (E2F1, EP300, NFATC3) and Transcriptional misregulation in cancer (AFF1, ATF1, MAX) as the top 3 enriched KEGG pathway terms (FDR < 0.05) (Figure 3C and Supplementary Table 2).

Taken together these findings indicate that the transcriptional activity of the vtRNAs is gene specific, being vtRNA1-1 the most accessible and possibly the most expressed. In addition, the data suggest that the regulatory status of the vtRNA promoters is coordinately controlled, and the three vtRNA1s are more co-regulated among themselves. Yet, the core of TFs shared by the vtRNAs is associated with viral infection and cancer related terms in the myeloid cell line K562.

The vtRNA promoters present gene and tissue specific patterns of chromatin accessibility in primary tumors.

In order to investigate the expression of vtRNAs in different tissues we analyzed the Pan-Cancer TCGA ATAC-seq data discriminating the tissue of origin. As expected, vtRNA1-1 has the highest promoter chromatin accessibility among tissues and the smallest variation (Figure 4A, Supplementary Table 1 and Supplementary Figure 1). Remarkably, low-grade glioma tumor samples (LGG) show a global reduction in chromatin accessibility at the 4 vtRNA promoters (Figure 4A). An opposite pattern is seen in adrenocortical carcinoma tissue (ACC), where the vtRNAs have the highest concerted chromatin accessibility (Figure 4A). Individually, vtRNAs promoter accessibility is maximum and minimum in THCA (4.7) and LGG (3.3) for vtRNA1-1, ACC (4.2) and LGG (0.29) for vtRNA1-2, ACC (3.7) and LGG (0.24) for vtRNA1-3 and KIRC (4.2) and LGG (1.0) for vtRNA2-1 (Figure 4A). VtRNA3-1P reaches a maximum of promoter chromatin accessibility in Testicular Germ Cell Tumors (TGCT) (0.65) with an extremely low average ATAC-seq value and its minimum in PRAD (-1.5) (Supplementary Figure 3C). Remarkably, although vtRNA1-1 and vtRNA1-3 have the highest and lowest promoter accessibility in the majority of tissues respectively, vtRNA1-2 and vtRNA2-1 are the most variable (vtRNA1-1 SD = 0.34, vtRNA1-2 SD = 1.4, vtRNA1-3 SD = 0.75 and vtRNA2-1 SD = 0.78) (Figure 4A). In addition, in 12 tissues the chromatin accessibility of vtRNA1-2 is higher than vtRNA2-1 (57%) whereas in 9 tissues the chromatin accessibility of vtRNA1-2 is lower than vtRNA2-1 (43%) (Figure 4A).

Using the vtRNAs promoter DNA methylation data, we extended the analysis to 32 primary tumor tissues, incorporating nine more tissues than ATAC-seq samples (DLBC, KICH, OV, PAAD, READ, SARC, THYM, UCS, UVM) (Figure 4B, Supplementary Figure 1) and 16 normal adjacent tissues (Figure 4C, Supplementary Figure 1). Again, the DNA methylation profile mirrors the chromatin accessibility in the individual tissue types (Figure 4). Remarkably, the association between the average vtRNA promoter chromatin accessibility and DNA methylation of the vtRNA1 cluster in each tissue type is higher than that observed for the average of all tissues (vtRNA1-1 $r_s = -0.28$, vtRNA1-2 $r_s = -0.90$ vtRNA1-3 $r_s = -0.74$ and vtRNA2-1 $r_s = -0.58$) (Table 1 and Supplementary Figure 4). The opposite finding for vtRNA2-1 may be explained by the

impact of the previously recognized chromatin polymorphisms in the locus (Silver et al., 2015), which may prevail over the tissue specific variation for this gene.

In primary tumor tissues, promoter DNA methylation of vtRNA2-1 is higher than vtRNA1-2 in 17 tissues (53%) and the opposite is observed in 15 tissues (47%) (Figure 4B), while in normal tissues, promoter DNA methylation of vtRNA2-1 is higher than vtRNA1-2 in 4 tissues (25%) and the opposite is observed in 12 tissues (75%) (Figure 4C), indicating that although vtRNA1-2 is more methylated in normal samples, vtRNA2-1 gains methylation in neoplastic tissue, becoming more methylated than vtRNA1-2 (in 7 tissue types: BLCA, BRCA, CHOL, HNCSC, LUAD, LUSC, UCEC). Indeed, the Fisher's exact test identifies a mirrored profile of chromatin accessibility between normal (4 tissues with high vtRNA1-2 and 12 tissues with high vtRNA2-1 from 16 total tissues) and tumor tissues (17 tissues with high vtRNA1-2 and 15 tissues with high vtRNA2-1 from 32 total tissues) for vtRNA1-2 and vtRNA2-1 (p-value = 0.07). Despite this deregulation in transformed tissues, we wondered if the tissue specific differences in vtRNA promoters were explained by their pre-existing status in the normal tissues. The correlation among average vtRNA promoter DNA methylation in normal and tumor tissues suggest that the variation in chromatin accessibility among tissue types is already established in the normal tissue counterparts (vtRNA1-1 rs = 0.53, vtRNA1-2 rs = 0.84, vtRNA1-3 rs = 0.47 and vtRNA2-1 rs = 0.75, Supplementary Figure 5).

VtRNA promoter's DNA methylation is associated with tumor stage and tissue of origin.

The assessment of differential vtRNA expression from normal to tumor condition at the TCGA cohort can only be performed using DNA methylation dataset since no ATAC-seq analyses were performed in the normal tissues. The average beta-value of CpG sites in the 500bp vtRNAs promoter of normal and tumor tissue evidenced gene specific patterns of deregulation in neoplastic tissues (Figure 5). As depicted in Figure 4, vtRNA1-1 and vtRNA1-3 promoter DNA methylation and chromatin accessibility is low and high tissue-wide respectively, in both normal and tumor tissues (yet, there are highly methylated tumor outliers as LGG, that lack normal counterpart) (Figure 4 and Figures 5A and 5C). Conversely, vtRNA1-2 and vtRNA2-1 revealed significant differences in average promoter DNA methylation among the tissue categories (Figures 5B and 5D). In agreement with previous results the methylation of vtRNA1-2 promoter

decreases from normal to tumor samples, while vtRNA2-1's increases (Figures 5B and 5D and Figure 4).

We next asked whether the promoter status of the vtRNAs is deregulated during the normal and tumor transformation of different tissue types. This comparison was restricted to the 16 tissues with normal samples (with data of at least five samples, Supplementary Figure 1, Supplementary Table 1 and Figure 6). The results revealed that vtRNA1-1 is globally unmethylated in normal and tumor from different tissue origins, but presents small but statistically significant changes in 4 tissue types (OG profile in BLCA, THCA UCEC and TSG in LUSC), whose biological impact remains to be determined (Figure 6A). Similarly, vtRNA1-3 is globally unmethylated, showing larger variability and significant upregulation in BRCA and PRAD, compatible with a TSG function (TSG trends for LIHC p-value = 0.08 and LUSC p-value = 0.09, Figure 6C). In agreement with the literature, vtRNA2-1 presents a TSG pattern of deregulation in PRAD, LUSC and BRCA (Cao et al., 2013; Fort et al., 2018; Romanelli et al., 2014), and an OG pattern in KIRP that has not been previously described (Figure 6D). Finally, a statistically significant dysregulation of vtRNA1-2 promoter DNA methylation across almost all cancer types (13 out of 16 tissues analyzed) poses it as a candidate OG in cancer (Figure 6B).

Overall, the data shows different DNA methylation changes in the vtRNAs promoters upon malignant transformation in several tissues. In addition, the deregulation of each vtRNA occurs mostly in a gene specific direction, where vtRNA1-2 has an OG like epigenetic de-repression while vtRNA2-1 (and to a lesser extent vtRNA1-1, vtRNA1-3) has a TSG like repressive methylation in tumor tissues.

High chromatin accessibility at the promoter of vtRNA1-1, vtRNA1-3 and vtRNA2-1 is associated with low patient overall survival.

Seeking for a possible clinical significance of the chromatin accessibility changes of the vtRNA promoters, we studied its association with the overall survival of patients (Figure 7 and Supplementary Table 3). The analysis of patients stratified by vtRNA promoter accessibility quartiles did not show differences in overall survival (data available at Supplementary Table 3). However, a significantly lower patient survival probability when patient tumors have lower

promoter DNA methylation is observed for vtRNA1-1 (p-value = 0.003), vtRNA1-2 (p-value = 0.004) and vtRNA2-1 (p-value = 0.002) (patient stratified in quartiles of the promoter DNA methylation cohort values, Figure 7). Therefore, a lower promoter DNA methylation of vtRNA1-1, vtRNA1-2 and vtRNA2-1, a surrogate of their high expression, might be associated with poor patient overall survival cancer-wide.

Genome-wide chromatin accessibility correlations with vtRNA promoters reveals their link to specific cancer related functions.

Seeking to get insight into the function of the vtRNAs from their transcription regulatory marks, we searched for the genes co-regulated at chromatin level. We calculated the correlation of the ATAC-seq promoter values of each vtRNA and all the individual genes in the 385 primary tumor cohort and selected those showing a Spearman correlation $r_s \geq 0.4$. Using STRING software analysis, we then investigated their connection with Biological Process categories and KEGG pathways (Szklarczyk et al., 2017). A few enriched terms for the genes correlated with vtRNA1-2, vtRNA1-3 and vtRNA2-1 is identified ($FDR < 1 \times 10^{-3}$) (Supplementary Table 4). The 303 genes correlated with vtRNA1-2 are enriched in the Biological Process “Epithelial cell differentiation” ($FDR < 1 \times 10^{-4}$) (Supplementary Table 4). Meanwhile, the 550 genes correlated with vtRNA2-1 are enriched in the Biological Processes “Immune system process”, “Inflammatory response” and “Immune response” ($FDR < 1 \times 10^{-5}$) and the KEGG pathways term “Cytokine-cytokine receptor interaction” ($FDR < 1 \times 10^{-4}$) (Supplementary Table 4). The only 11 genes correlated with vtRNA1-3 are enriched in the “Thyroid cancer” KEGG pathway term ($FDR < 1 \times 10^{-4}$). Remarkably, although STRING analysis of the 46 vtRNA1-1 associated genes did not find enriched process or pathways, 45 of the 46 correlated genes are located at chromosome 5q region. Using the Cluster Locator algorithm (Pazos Obregón et al., 2018), we found a statistically significant non-random clustering behavior for the chromosomal location of the genes co-regulated with vtRNA1-1 (p-value $< 1 \times 10^{-10}$). Indeed, when the analysis was extended to 213 genes with Spearman correlation $r_s \geq 0.3$, 186 genes were situated at chromosome 5 (p-value $< 1 \times 10^{-10}$). These genes are positioned in particular regions at chr5q31, chr5q35, chr5q23, chr5q14, chr5q32, chr5q34 ($FDR < 1 \times 10^{-4}$) (Kuleshov et al., 2016) (Supplementary Table 4). We did not find a similar phenomenon for the other 3 vtRNAs. In summary, vtRNA1-3, vtRNA1-2

and vtRNA2-1 transcription may be co-regulated with genes belonging to specific biological processes located elsewhere, while vtRNA1-1 transcription may be controlled by a locally specialized chromatin domain at chromosome 5q region lacking apparent functional relatedness.

Immune Subtypes associated with the vtRNAs profile.

Taking into account that vtRNAs have been previously related to the immune response (Golec et al., 2019; Li et al., 2015; Nandy et al., 2009), and that we found an association of vtRNA2-1 with immune related terms, we sought to investigate a putative relation between the vtRNAs expression and the six Immune Subtypes defined by Thorsson et al. (Thorsson et al., 2018), compiled at Pan-Cancer TCGA (Supplementary Table 5). These six Immune Subtypes are named because of the foremost immune characteristic, comprising wound healing, IFN-g dominant, lymphocyte depleted, inflammatory, immunologically quiet and TGF- β dominant types (Thorsson et al., 2018). As was expected, mirrored patterns were observed for promoter ATAC-seq and DNA methylation data of the six groups (Figure 8). The immunologically quiet subtype shows a concerted shutdown of all vtRNA promoters (Figure 8), which is largely composed by LGG samples showing low vtRNA ATAC-seq and high promoter methylation average values in LGG tumors (Figures 4A and 4B). The immunological quiet subtype exhibits the lowest leukocyte fraction (very low Th17 and low Th2) and the highest macrophage:lymphocyte ratio (high M2 macrophages) (Thorsson et al., 2018). The remaining subtypes present a similar vtRNA profile except for the inversion of vtRNA1-2/vtRNA2-1 accessibility ratio, which is higher than 1 for the wound healing and IFN-g dominant subtypes while is lower than 1 for the lymphocyte depleted and inflammatory subtypes (Figure 8A). As expected, a mirrored pattern was observed for vtRNA1-2 and vtRNA2-1 at DNA methylation data (Figure 8B). Since the wound healing and IFN-g dominant subtype have high proliferation rate opposite to the lymphocyte depleted and inflammatory subtypes (Thorsson et al., 2018), we evaluated the correlation between vtRNA1-2 and vtRNA2-1 chromatin status and proliferation rate or wound healing status defined by Thorsson et al. (Supplementary Table 5) (Thorsson et al., 2018). A negative correlation between vtRNA1-2 promoter DNA methylation and the proliferation rate ($r_s = -0.44$) and wound healing features of the tumors is revealed ($r_s = -0.48$) (Supplementary Table 5). These findings suggest that vtRNA transcriptional regulation may be associated with tumor immunity.

Discussion

Although important insight into the vtRNA function has been recently gained, the landscape of vtRNA expression in human tissues and cancer was unknown, perhaps for the absence of vtRNA data in transcriptomic studies. Here we used ATAC-seq and methylation arrays data from the Pan-Cancer TCGA consortium (Weinstein et al., 2013), as surrogate variables to uncover the expression of vtRNAs across multiple cancer types. A limitation of our study is its blindness to post-transcriptional controls of vtRNA expression, such as RNA processing, transport, localization and/or stability. Indeed, there is evidence of vtRNA transcript regulation via RNA methylation by NSUN2, or via its association to other proteins like DUSP11, DIS3L2, SSB, SRSF2 and DICER (Burke et al., 2016; Hussain et al., 2013; Kickhoefer et al., 2002; Łabno et al., 2016; Persson et al., 2009; Sajini et al., 2019). Nonetheless, the chromatin status of the vtRNA promoters has been correlated with the expression of their transcripts in several reports (Ahn et al., 2018; Fort et al., 2018; Helbo et al., 2015, 2017; Lee et al., 2014a, 2014b; Sallustio et al., 2016; Treppendahl et al., 2012). These evidences unequivocally stand for the use of chromatin structure as a suitable proxy for vtRNA expression.

An integrated genomic view of the regulatory features of the vtRNA genes compiled in UCSC genome browser indicates that the 4 vtRNAs have chromatin signs of active transcription in different cell lines and defines a proximal promoter regulatory region of approximately 500pb (Nikitina et al., 2011; Vilalta et al., 1994). High resolution chromatin accessibility mapping studies by NOMe-Seq of the vtRNA promoters in cell lines validate this assumption (Helbo et al., 2017). Particularly, vtRNA2-1 is singular for being immersed in a CpG island, for lacking the TATA box, DSE or PSE elements characteristic of the others vtRNA promoters, and for bearing a cAMP-response (CRE) element (Canella et al., 2010; Helbo et al., 2017; Nandy et al., 2009). Since ATAC-seq approach is currently the best high throughput method to determine the chromatin accessibility of a DNA region directly (Chen et al., 2016; Sun et al., 2019; Tsompana and Buck, 2014), we interrogated the ATAC-seq data of TCGA samples as a surrogate marker of vtRNA expression (Corces et al., 2018). Although DNA methylation is a less direct and less accurate measurement of chromatin accessibility compared to ATAC-seq, 25 times more TCGA samples have been analyzed by DNA Methylation Arrays, including matched normal tissues samples,

thus this dataset is of great significance for the current study. Our analyses revealed that vtRNA1-1 has accessible chromatin and small variation in all the evaluated tissues, which is consistent with its identity as the ancestor vtRNA gene and its significance as the major RNA component of the ubiquitous vault particle (Stadler et al., 2009; van Zon et al., 2001). In addition, vtRNA1-2, vtRNA1-3 and vtRNA2-1, have lower chromatin accessibility at their promoters and a larger variation among the samples. As expected, the average DNA methylation assessed by microarrays is negatively correlated with the chromatin accessibility assessed by ATAC-seq. Therefore, the landscape of vtRNA chromatin status was corroborated by both approaches. Yet, the low chromatin accessibility of vtRNA1-3 promoter is not reflected by a concomitant high DNA methylation, suggesting that additional regulatory features of this gene, acting in trans or cis, play a major influence in its chromatin structure. As was reported by Helbo et. al. using NOME-seq assay (Nucleosome Positioning), the chromatin at vtRNA1-3 may be poised for transcription, and still require TFs to achieve physiological levels of transcription (Helbo et al., 2017).

The relative abundance of the 4 vtRNAs determined in the current study has been previously observed in various cell lines, including HSC CD34+, HL60 (Helbo et al., 2015), T24, colorectal RKO, human diploid fibroblasts IMR90 (Helbo et al., 2017) and MCF7 (Chen et al., 2018). Yet, others found different levels of the vtRNAs in MCF7, SW1573, GLC4, 8226, KB, SW620, MCF-7, HeLa, HEK-293 and L88/5 (van Zon et al., 2001), CBL, BL2 and BL41 (Nandy et al., 2009), EBV-negative Burkitt's lymphoma cell lines and A549 (Li et al., 2015) and in LNCaP, PC3, DU145, RWPE-1, HEK, HeLa and MCF-7 (Stadler et al., 2009) cell lines, suggesting that the expression of vtRNAs may differ in laboratory cell lines compared to tissues. Alternatively, the methods used to quantify the vtRNAs are not comparable among the studies.

Seeking to investigate regulatory connections between the vtRNAs as a valuable tool to study their function, we determine the pairwise correlations between the promoters of the vtRNAs. The positive correlations found are compatible with a coordinated transcriptional control of the vtRNAs, that is greater for the three members of the vtRNA1 cluster (specially vtRNA1-1 and vtRNA1-3). The exception is the pair vtRNA1-2 and vtRNA2-1, whose chromatin accessibility and promoter methylation are not linked in the cohort and are indeed negatively associated in some comparisons.

As an alternative approach to investigate the vtRNA transcriptional co-regulation, we analyzed transcription and chromatin remodeling factors associated with their proximal promoter region. In agreement with their higher chromatin correlation, the vtRNA1 cluster have a larger core of common TFs (39) in comparison with vtRNA2-1 (23 TFs). The divergence of the transcriptional control of vtRNA2-1 was previously recognized given its unique regulatory elements and the complex developmental methylation it undergoes (Canella et al., 2010; Carpenter et al., 2018; Helbo et al., 2017; Nandy et al., 2009). Remarkably, the 23 TFs common to all vtRNAs are related to viral infections and cancer. The upregulation of the four vtRNAs upon viral infection has been reported for several viruses, so it is possible that this core of TFs participate in the coordinated expression of vtRNA during viral infection and related responses (Amort et al., 2015; Li et al., 2015; Nandy et al., 2009). These TFs regulate cell cycle and translational arrest as well as the inhibition of host innate immune response (Attar and Kurdistani, 2017; Ertosun et al., 2016; Horsley and Pavlath, 2002; Jean et al., 2000; Persengiev and Green, 2003; Rohini et al., 2018). Since these processes are also hallmarks of cancer, the dysregulation of the normal function of vtRNAs in viral response may provide a selective advantage for cancer cell to overcome the many sources of stress faced during neoplastic evolution. If that holds true, the core vtRNA TFs may be co-opted in the malignant context to tune vtRNA expression favoring tumor progression. In agreement with that hypothesis, ATF1 and ATF3 were associated with CREB response and increased cell viability (Persengiev and Green, 2003). Likewise, Golec et al. 2019 showed that altered levels of vtRNA2-1 modulate PKR activation and consequently altered the levels of CREB phosphorylation during T cell activation, which is a prerequisite for IFN-g activity (Golec et al., 2019). Furthermore, E2F1, a cell cycle regulator, was described as a direct transcriptional regulator of vtRNA2-1 in cervical cancer cells (Li et al., 2017). Moreover, it was shown that MYC binding to vtRNA2-1 promoter raises its expression and could be the explanation to the increased levels of vtRNA2-1 in some tumors (Park et al., 2017). Similarly, TGFB1 provokes the demethylation of vtRNA2-1 promoter and consequently increases its expression in ovarian cancer (Ahn et al., 2018).

Due to its importance for cancer biology, vtRNA2-1 transcriptional regulation was recently more investigated (Yeganeh and Hernandez, 2020). Various studies reported tissue specific roles of

vtRNA2-1 in normal (Golec et al., 2019; Lee et al., 2019a; Miñones-Moyano et al., 2013; Sallustio et al., 2016; Suojalehto et al., 2014) and cancer tissue (Ahn et al., 2018; Fort et al., 2018; Hu et al., 2017; Im et al., 2020; Jeon et al., 2012a; Kunkeaw et al., 2012; Lee et al., 2011; Lee, 2015; Lee et al., 2016, 2014a, 2014b; Lei et al., 2017; Li et al., 2017; Ma et al., 2020; Treppendahl et al., 2012) and epigenetic alterations of the locus have been described in different malignancies (Ahn et al., 2018; Cao et al., 2013; Fort et al., 2018; Helbo et al., 2015, 2017; Joo et al., 2018; Lee et al., 2014a, 2014b; Park et al., 2017; Romanelli et al., 2014; Treppendahl et al., 2012). Meanwhile, except for Kirsten Grønbaek group's contributions on the chromatin characterization of vtRNA1-3 and vtRNA1-2 promoters (Helbo et al., 2015, 2017), little is known about the transcriptional regulation and expression of the other vtRNAs. Likewise, to our knowledge, the global landscape of vtRNA expression across different cancer types has not been investigated. Our study shows that, except for vtRNA1-1, which has a high promoter chromatin accessibility and low DNA methylation in all the tissues, the other vtRNA promoters exhibit different levels of chromatin accessibility and DNA methylation among the tissue types. The relative status of the chromatin in different tissues is accurately mirrored by both approaches in the tissues examined by ATAC-seq and Illumina DNA methylation arrays. Interestingly, the same analysis performed in the normal tissue samples reveals that the tissue specific variation in DNA methylation of vtRNA promoters in the neoplastic tissues is already established in their normal tissue counterparts, indicating that the regulation of vtRNA expression is tuned during normal development and cell differentiation.

While there is no apparent concerted modulation of the chromatin accessibility of the vtRNA promoters in most tissues, ACC and LGG tumors seem to be the exception. Concordantly, a global hypomethylation of malignant ACC tumors (Legendre et al., 2016; Rechache et al., 2012), and an aberrantly methylation processes associated to altered DNMTs activity in LGG tumors were reported (Nomura et al., 2019). However, more research is necessary to understand the molecular basis of these observations.

The differential expression of the vtRNAs from normal to primary tumor tissues revealed different patterns depending on tissue origin. The average DNA methylation of the vtRNA1-1 and vtRNA1-3 promoters is not significantly deregulated, whereas vtRNA1-2 and vtRNA2-1 are

decreased and increased respectively in tumor tissues. The latter finding favors a candidate OG and TSG function of vtRNA1-2 and vtRNA2-1 in cancer respectively. Indeed, the epigenetic repression of vtRNA2-1 was previously reported by Romanelli et al. in BLCA, BRCA, COAD and LUSC analyzing the same data from TCGA (Romanelli et al., 2014). Furthermore, functional studies in LUSC, PRAD, ESCA and AML provided experimental support to that hypothesis (Cao et al., 2013; Fort et al., 2018; Lee et al., 2014a; Treppendahl et al., 2012). Yet, vtRNA2-1 has been also proposed as an OG in ovarian, thyroid, endometrial, cervical and renal cancer (Ahn et al., 2018; Hu et al., 2017; Lee et al., 2016; Lei et al., 2017; Li et al., 2017; Yeganeh and Hernandez, 2020). In agreement with the functional data found by Lei et. al. (Lei et al., 2017), we found an epigenetic de-repression of vtRNA2-1 promoter DNA methylation in renal carcinoma (KIRP). Likewise, a closer look at THCA samples, shows an average decrease in the DNA methylation of vtRNA2-1 promoter that supports the OG function described in this tissue (Lee et al., 2016). Unfortunately, normal ovarian, cervical and endometrial tissues are not available at the TCGA. Finally, vtRNA1-2 has an oncogenic pattern of inferred expression between normal and tumor tissues for 13 of 16 tissues analyzed, raising a possible oncogenic function for this RNA.

Our study found also an association of DNA methylation at the promoter with a shorter overall survival for vtRNA1-1, vtRNA1-2 and vtRNA2-1. This is surprising for vtRNA2-1, since it has both TSG and OG roles and TSG compatible expression profiles in 3 cancer types, inferred by DNA methylation of its promoter in tumor vs normal tissues and the average promoter DNA methylation is increased in tumor compared to normal tissue. A higher impact of vtRNA2-1 expression in patient survival in the oncogenic context may justify this discrepancy. On the contrary, the association between vtRNA1-2 low promoter DNA methylation and low patient survival is in agreement with its OG expression in cancer. The lack of survival association with the ATAC-seq values may be due to the small number of patients.

VtRNA expression in individual cancer types, inferred from the epigenetic status of their promoters, has been associated with patient survival in several studies of vtRNA2-1, one of vtRNA1-3 and none for vtRNA1-1 or vtRNA1-2. Low methylation or high expression of vtRNA2-1 promoter was associated with good prognosis or overall survival in lung (Cao et al., 2013),

esophageal (Lee et al., 2014a), prostate (Fort et al., 2018), AML (Treppendahl et al., 2012), gastric (Lee et al., 2014b) and liver (Yu et al., 2020). Conversely, a worse prognosis or overall survival association of vtRNA2-1 was reported in thyroid (Lee et al., 2016) and ovarian (Ahn et al., 2018). From these reports, only prostate (Fort et al., 2018) and gastric (Lee et al., 2014b) studies used the TCGA data. Likewise, the methylation status of the vtRNA1-3 promoter associates with overall survival in the lower risk Myelodysplastic Syndrome patients (Helbo et al., 2015). Additionally, vtRNA1-1 and vtRNA1-2 correlate to chemotherapeutic resistance by direct interaction with drugs (doxorubicin, etoposide and mitoxantrone) and the modulation of vtRNA1-1 confirmed this finding in osteosarcoma cell lines (Gopinath et al., 2005, 2010; Mashima et al., 2008; van Zon et al., 2001). Furthermore, increased levels of vtRNA1-1 were associated with increased proliferation and chemoresistance due to GAGE6 induction in MCF-7 cells (Chen et al., 2018). Besides, Norbert Polacek group showed that vtRNA1-1 expression confers apoptosis resistance in several human cell lines (BL2, BL41, HS578T, HEK293, A549 and HeLa) and revealed its capacity to repress intrinsic and extrinsic apoptosis pathway (Amort et al., 2015; Bracher et al., 2020). The later findings are in agreement with the lower patient survival associated with high levels of vtRNA1-1 expression cancer-wide observed in our analysis.

Seeking to get insights into the cancer related function of the vtRNAs we performed a genome wide search for genes co-regulated at the level of promoter chromatin accessibility. Remarkably, we identified immune and cytokine related terms for the genes co-regulated with vtRNA2-1. This association agrees with its proposed roles in innate immune modulation via PKR repression or OAS1 regulation and in cytokine production (Calderon and Conn, 2018; Golec et al., 2019; Lee et al., 2019b; Li et al., 2015). Furthermore, vtRNA2-1 has been associated with autoimmune disorders (Renauer et al., 2015; Weeding and Sawalha, 2018) and tumor engraftment in prostate cancer (Ma et al., 2020). This enrichment in viral infection pathways, together with the viral infection involvement of the core TFs common to all vtRNAs, reinforces the hypothesis of the re-utilization of a regulatory RNA induced upon viral response in favor of cancer development at upstream and downstream regulation steps. Additionally, we found a non-random clustering localization at chromosome 5 for genes co-regulated with vtRNA1-1. The chromosome arm 5q and regions chr5q31 and chr5q32-q33 were found frequently deleted in myelodysplastic

syndrome (MDS) and acute myeloid leukemia (AML) (Fuchs, 2012; Treppendahl et al., 2012). Furthermore, the modulation of vtRNA2-1 and vtRNA1-3 was previously associated with human bone marrow CD34+ cells, AML and MDS (Helbo et al., 2015; Treppendahl et al., 2012). Yet, our analysis points to vtRNA2-1 and vtRNA1-2, and to a lesser extent to vtRNA1-1 and vtRNA1-3, as the most relevant players across cancer types.

Support for the involvement of the vtRNAs in the immune responses in the context of cancer came from the association of their chromatin status with the immune subtype categories defined by Thorsson et al. (Thorsson et al., 2018). Our findings suggest a possible role of the vtRNAs in lymphocyte response in tumor samples, since the immunologically quiet subtype shows the lowest leukocyte fraction and vtRNAs promoter chromatin accessibility (Thorsson et al., 2018). Meanwhile, the vtRNA2-1 and vtRNA1-2 expression distinguishes 2 groups among the remaining 4 subtypes. The participation of vtRNA2-1 in the induction of the IFN- γ and IL-2 expression in activated T cells through PKR modulation (Golec et al., 2019) goes in agreement with this hypothesis. Although the same study demonstrated that vtRNA1-1 expression was unchanged during T cell activation, it is remarkable that the 5q31-q33 region, comprising the vtRNA1 locus and chromatin co-regulated genes, encodes several cytokines that regulate the differentiation of Th1 and Th2 lymphocytes and has been linked with susceptibility to infections (Jeronimo et al., 2007; Lacy et al., 2000; Naka et al., 2009; Rodrigues et al., 1999). Furthermore, vtRNA2-1 levels were recently associated with macrophages M1/M2 fates in prostate PC3 cell line mice xenografts via TGF- β (Ma et al., 2020), pathways that define the six immune subtypes (Thorsson et al., 2018). Finally, we found that the proliferation rate and wound healing of the tumors are associated with the levels of vtRNA1-2 chromatin accessibility, which reinforces a OG role of vtRNA1-2 in cancer. Nonetheless, the participation of the vtRNAs in immune cells response inside the neoplastic niche needs to be further investigated.

Taken together our analysis reveals the pattern of chromatin accessibility and DNA methylation at the four vtRNA promoters, analyzed by tissue of origin in the TCGA cohort. The agreement between both dataset and previous related literature endorse the use of these variables as surrogates of the vtRNA transcripts expression. The four vtRNAs seem to be co-regulated at the transcriptional level, and TFs involved in viral infection are likely to take part in their coordinated

transcription. The pattern of expression in normal vs tumor tissue and the association with cancer cell pathways and patient survival suggest that vtRNA2-1 and vtRNA1-2 are possibly the more relevant contributors to cancer. The results favor tissue specific TSG and OG roles for vtRNA2-1, a still not investigated oncogenic role of vtRNA1-2 and a limited cancer driver effect of vtRNA1-1/3 in specific tissue types. Lastly, we uncovered new evidence linking the vtRNAs with the immune response, cell proliferation and overall survival in cancer, which guarantees further investigation.

Materials and Methods.

ATAC-seq data

The ATAC-seq data obtained by The Cancer Genome Atlas (TCGA) consortium were retrieved from UCSC Xena Browser (Goldman et al., 2020) on date 03/10/2018. It comprises the genomic matrix TCGA_ATAC_peak_Log2Counts_dedup_promoter with normalized count values for 404 samples (385 primary solid tumor samples across 23 tissues). As is described in UCSC Xena Browser to calculate the average ATAC-seq values a prior count of 5 was added to the raw counts, then put into a "counts per million", then log2 transformed, then quantile normalized; the result is the average value in the file (log2(count+5)-qn values) across all technical replicates and all biospecimens belonging to the same TCGA sample group. The gene promoter for ATAC-seq is defined as a region within -1000 to +100bp from the TSS site. Assignment of promoter peak to gene mapping information derives from the peak summit within the promoter region. Peak location information of the ATAC-seq values for 500 bp gene promoter regions was retrieved from the file TCGA_ATAC_Log2Counts_Matrix.180608.txt.gz. All the analyses were performed for tissues with at least five primary tumor samples (21 tissues), which resulted in the exclusion of CESC (2 primary tumor samples) and SKCM (4 primary tumor samples). Correlations between ATAC-seq and DNA promoter methylation were performed for the 329 primary tumors that are studied by both strategies (Supplementary table 1).

DNA methylation data

The DNA methylation data obtained from The Cancer Genome Atlas (TCGA) consortium was retrieved from UCSC Xena Browser (Goldman et al., 2020) on date 03/10/2018. It comprises the normalized beta-value of DNA methylation obtained using Illumina Infinium Human Methylation 450 BeadChip arrays of the total (9149 samples) normal adjacent tissues (746 samples across 23 tissues) and primary solid tumors (8403 samples across 32 tissues) of the Pan-Cancer TCGA cohort. The current study is limited to solid tumors since only AML DNA methylation is available at the TCGA. All the analysis was performed for tissues with at least five samples. Normal adjacent tissue comprises 16 tissues, excluding CESC (2 samples), GBM (2 samples), PCPG (3 samples), SARC (4 samples), SKCM (2 samples), STAD (2 samples) and THYM (2 samples). Primary tumors comprise 32 tissues. The normalized promoter average beta-values (500 bp bin) comprise the following CpG sites for vtRNA1-1: cg12532653, cg05913451, cg14633504, cg16615348, cg25602765, cg13323902, cg18296956; for vtRNA1-2: cg21161173, cg13303313, cg00500100, cg05174942, cg25984996, cg11807153, cg15697852; for vtRNA1-3: cg23910413, cg19065177, cg02053188, cg01063759, cg07379832, cg07741016; and for vtRNA2-1: cg16615357, cg08745965, cg00124993, cg26896946, cg25340688, cg26328633, cg06536614, cg18678645, cg04481923 (Supplementary table 1).

Immune Subtypes data.

The data of Immune Subtypes defined by Thorsson et al. (Thorsson et al., 2018) is available at the Pan-Cancer TCGA and was retrieved from UCSC Xena Browser (Goldman et al., 2020) on date 03/10/2018 and from the supplementary material of Thorsson et al. article (Thorsson et al., 2018). ATAC-seq data of vtRNA promoters for the six Immune Subtype comprised a total of 364 samples: wound healing (105 samples), IFN-g dominant (93 samples), inflammatory (110 samples), lymphocyte depleted (43 samples), immunologically quiet (9 samples), TGF-b dominant (4 samples). DNA methylation average normalized beta-values of vtRNA promoters for the six Immune Subtype comprised 7693 samples including wound healing (1963 samples), IFN-g dominant (2137 samples), inflammatory (2126 samples), lymphocyte depleted (933 samples), immunologically quiet (383 samples), TGF-b dominant (151 samples) (Supplementary table 5).

Overall survival analysis.

The Kaplan Meier curve analysis was performed with software GraphPad Prism 6 using Log-ranked Mantel-Cox test. We analyzed the ATAC-seq normalized values and DNA methylation average normalized beta-values for vtRNA promoters and survival data until 4000 days (99% of data). The log-rank statistics was calculated for quartiles of vtRNAs promoter chromatin accessibility comparison groups (25th and 75th) (Supplementary Table 4).

Pathway enrichment and cluster chromosome localization analyses.

The genome wide correlation analysis of promoter chromatin accessibility (ATAC-seq normalized values) for each vtRNA in comparison with all the annotated genes was performed using the genomic matrix TCGA_ATAC_peak_Log2Counts_dedup_promoter retrieved from UCSC Xena Browser (385 primary tumor samples across 23 tissues) (Goldman et al., 2020) on date 03/10/2018. The genes showing a Spearman correlation $r \geq 0.4$ were selected for pathway and cluster location analysis. The pathway enrichment analysis was done using STRING software for the Gene Ontology Biological Process and KEGG pathways categories (Szklarczyk et al., 2017). The cluster localization analysis was performed with two approaches, the Cluster Locator algorithm (Pazos Obregón et al., 2018) and the Enrichr software (Kuleshov et al., 2016).

Statistical Analysis.

Unless specified all the variables are expressed as average value \pm standard deviation (SD). Statistical analyses were performed with one-way ANOVA for multiple comparison tests, including Sidak's Honest Significant Difference test as a post-hoc test, as referred in the legend of the Figures. Spearman equations were applied to determine the correlations. All the analyses were done in R software (version 3.6) using libraries ggcorrplot, corrplot, xlsx, heatmap.2 (clustering distance measured by Euclidean and Ward clustering algorithms) and software GraphPad Prism 6. The statistical significance of the observed differences was expressed using the p-value (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** p -value < 0.0001). Differences with a p-value of < 0.05 were considered significant.

Figures and Tables legends:

Figure 1. Genomic view of human vtRNA genes epigenetic features. Genomic view of 1.5 kb region of human vtRNA genes in UCSC Genome browser (GRCh37/hg19) centered at the 500bp bin highlighted in *yellow*, which was used for ATAC-seq and CpG methylation analyses. VtRNAs of locus 1 (vtRNA1-1, vtRNA1-2, vtRNA1-3) and of locus 2 (vtRNA2-1) clusters are in the sense and antisense orientation respectively. Several Gene annotation and ENCODE Project tracks for seven cell lines (GM12878, H1-hESC, HSMM, HUVEC, K562, NHEK, NHL) are displayed: DNA accessibility (DNaseI hypersensitivity clusters (color intensity is proportional to the maximum signal strength)), DNA methylation (CpG islands length greater than 200 bp), histone modification (H3K27Ac, H3K4me1, H3K4me3 marks), conservation of the region in 100 Vertebrates (log-odds PhyloP scores). The vertical viewing range of the tracks displays the default settings of the browser for each variable.

Figure 2. Chromatin accessibility and DNA methylation of the vtRNA promoters. Analysis of vtRNA 500pb promoter region of **A.** Average chromatin accessibility ATAC-seq values for the 385 primary tumors available at the Pan-Cancer TCGA dataset (23 tissues types) and **B.** Average DNA methylation beta-values for the 8403 primary tumors (32 tissues types). Dashed horizontal lines denote unmethylated (bottom gray, average beta-value ≤ 0.2), 50% methylated (middle red, average beta-value = 0.5) and highly methylated (top gray, average beta-value ≥ 0.8) promoter. The box plots show the median and the lower and upper quartile, and the whiskers the 2.5 and 97.5 percentile of the distribution.

Figure 3. Comparative chromatin accessibility and transcription factor occupancy at the vtRNA loci. **A-B** Matrix of pairwise Spearman correlations (two-way hierarchical clustering distance measured by Euclidean and Ward clustering algorithms) of vtRNA 500pb promoter region for ATAC-seq data (385 primary tumors samples across 23 tissues) (A) and DNA methylation average beta-values data (8403 primary tumors samples across 32 tissues) (B). **C.** Venn diagram of transcription factors identified as ChIP-seq Peaks by ENCODE 3 project in the cell line K562 (Venny 2.1; <https://bioinfogp.cnb.csic.es/tools/venny/index.html>). The region for TFs assignment was defined as ± 3000 pb from the vtRNA transcript sequence boundaries. **D.** Interaction cluster of the core 23 TFs common to all vtRNAs performed with STRING (Szklarczyk et al., 2017). The colored labels of the top 3 enriched KEGG pathway terms (FDR < 0.05) are indicated.

Figure 4. Chromatin accessibility and DNA methylation of the vtRNA promoters in tumor and normal tissues. **A.** ATAC-seq values for vtRNA promoters in 21 primary tumor tissues. **B.** Average beta-values of promoter DNA methylation for vtRNAs in 32 primary tumors tissues. **C.** Average beta-values of promoter DNA methylation for vtRNAs in 16 normal adjacent tissues. For **A**, **B** and **C**, the charts show the average value and standard deviation of each vtRNAs for the tissues with at least five samples available at Pan-Cancer TCGA dataset.

Figure 5. Promoter DNA methylation of vtRNAs in Normal and Tumor tissues of Pan-Cancer TCGA dataset. Average beta-values of promoter DNA methylation of vtRNA1-1 (**A**), vtRNA1-2 (**B**), vtRNA1-3 (**C**) and vtRNA2-1 (**D**), assessed in 746 normal and 8403 tumor tissues respectively. The box plots show the median line and lower and upper quartile and the whiskers the 2.5 and 97.5 percentile. Horizontal grey striped and red dotted lines denote unmethylated (average beta-value ≤ 0.2), 50% methylated (average beta-value = 0.5) and highly methylated (average beta-value ≥ 0.8) promoters. One-way ANOVA multiple test analysis with Sidak as posthoc was performed. **** p-value < 0.0001.

Figure 6. VtRNA promoters DNA methylation in Normal vs. Tumor samples. Average beta-values of promoter DNA methylation for vtRNA1-1 (**A**) vtRNA1-2 (**B**), vtRNA1-3 (**C**) and vtRNA2-1 (**D**). Acronyms indicate the tissue

condition (normal (N) and tumor (T)). *Blue* and *red colors* indicate an increase or reduction in promoter methylation in tumor vs their normal tissues counterparts, compatible with a TSG and OG function respectively. The box plots show the median and the lower and upper quartile, and the whiskers the 2.5 and 97.5 percentile of the distribution. Horizontal lines denote the methylation level of the promoters: grey striped bottom and top for unmethylated (average beta-value ≤ 0.2) or highly methylated (average beta-value ≥ 0.8) respectively, and red dotted for 50% methylated (average beta-value = 0.5) promoters. One-way ANOVA multiple test analysis with Sidak as posthoc was performed. * p-value < 0.05; ** p-value < 0.01; **** p-value < 0.0001.

Figure 7. Patient Overall survival relative to vtRNA promoters DNA methylation status in Pan-Cancer TCGA dataset. Kaplan-Meier curves of overall patient survival probability over the time (4000 days) discriminating the primary tumors into two cohorts with relatively higher (75th) and lower (25th) DNA methylation at each vtRNA promoter. Promoter DNA methylation data of the 8060 primary tumors was used to stratify the patients in two quartiles, based on the expression of vtRNA1-1 (A) vtRNA1-2 (B), vtRNA1-3 (C) and vtRNA2-1 (D). Patient survival probability of the two groups was analyzed using Log-rank (Mantel-Cox) test. The dotted lines represent the 95% confidence interval for each curve.

Figure 8. Chromatin accessibility and DNA methylation of vtRNA promoters in the Six Immune Subtypes analyzed in the PANCAN TCGA dataset. A. ATAC-seq values expressed as log2 normalized values for vtRNA promoters grouped in six Immune Subtype: wound healing (105 samples), IFN-g dominant (93 samples), inflammatory (110 samples), lymphocyte depleted (43 samples), immunologically quiet (9 samples), TGF-b dominant (4 samples). B. DNA methylation average beta-values for vtRNA promoters grouped in six Immune Subtype: wound healing (1963 samples), IFN-g dominant (2137 samples), inflammatory (2126 samples), lymphocyte depleted (933 samples), immunologically quiet (383 samples), TGF-b dominant (151 samples) (Thorsson et al., 2018). The charts show the average and standard deviation for each vtRNA promoter in each Immune Subtype group.

Table 1. Correlation between vtRNA promoters ATAC-seq and DNA methylation in the Pan-Cancer TCGA dataset. The Spearman correlation was calculated for the 329 primary tumors studied by ATAC-seq and DNA methylation.

Supplementary Materials: Supplementary Tables and Figure Legends

Supplementary Figure 1. Data type and tissue sample distribution. Sample distribution per tissue type in the ATAC-seq, DNA methylation, ATAC-seq & DNA methylation and Normal adjacent tissues sample groups. A. Number of samples. B. Percentage of samples. The color reference, the number of samples of this tissue and the accounted percentage are presented to each tissue. Names references: ACC (Adrenocortical carcinoma), BLCA (Bladder Urothelial Carcinoma), BRCA (Breast invasive carcinoma), CESC (Cervical squamous cell carcinoma and endocervical adenocarcinoma), CHOL (Cholangiocarcinoma), COAD (Colon adenocarcinoma), DLBC (Lymphoid Neoplasm Diffuse Large B-cell Lymphoma), ESCA (Esophageal carcinoma), GBM (Glioblastoma multiforme), HNSC (Head and Neck squamous cell carcinoma), KICH (Kidney Chromophobe), KIRC (Kidney renal clear cell carcinoma), KIRP (Kidney renal papillary cell carcinoma), LGG (Brain Lower Grade Glioma), LIHC (Liver hepatocellular carcinoma), LUAD (Lung adenocarcinoma), LUSC (Lung squamous cell carcinoma), MESO (Mesothelioma), OV (Ovarian serous cystadenocarcinoma), PAAD (Pancreatic adenocarcinoma), PCPG (Pheochromocytoma and Paraganglioma), PRAD (Prostate adenocarcinoma), READ (Rectum adenocarcinoma),

SARC (Sarcoma), SKCM (Skin Cutaneous Melanoma), STAD (Stomach adenocarcinoma), TGCT (Testicular Germ Cell Tumors), THCA (Thyroid carcinoma), THYM (Thymoma), UCEC (Uterine Corpus Endometrial Carcinoma), UCS (Uterine Carcinosarcoma) and UVM (Uveal Melanoma).

Supplementary Figure 2. Correlation between ATAC-seq values and promoter methylation values of vtRNAs in Pan-Cancer TCGA dataset. A-D. The correlation was calculated between ATAC-seq data and DNA methylation average normalized beta values data for vtRNAs (vtRNAs1-1 (A), vtRNA1-2 (B), vtRNA1-3 (C) and vtRNA2-1 (D)), established in 329 primary tumors samples. The box plots show the median line and lower and upper quartile and the whiskers the 2.5 and 97.5 percentile. Horizontal grey striped and red dotted lines are shown to denote unmethylated (average beta-value ≤ 0.2), 50% methylated (average beta-value = 0.5) and highly methylated (average beta-value ≥ 0.8). Spearman correlation value and the best-fit line (red line) with 95% confidence bands (black dot lines) are shown.

Supplementary Figure 3. Genomic context of human vtRNA3-1P and the evaluated region of ATAC-seq.

A. Genomic view of the 1.5 kb region of human vtRNA3-1P gene in UCSC Genome browser (GRCh37/hg19) centered at the 500bp of the ATAC-seq. Highlighted in yellow is the 500 bp ATAC-seq region used in the posterior analyses. The following Gene annotation and ENCODE Project tracks for seven cell lines (GM12878, H1-hESC, HSMM, HUVEC, K562, NHEK, NHL) are displayed: DNA accessibility (DNaseI hypersensitivity clusters (color intensity is proportional to the maximum signal strength)), DNA methylation (CpG islands length greater than 200 bp), histone modification (H3K27Ac, H3K4me1, H3K4me3 marks), conservation of the region in 100 Vertebrates (log-odds PhyloP scores). The vertical viewing range of the tracks displays the default settings of the browser for each variable. **B.** The assay for transposase-accessible chromatin using sequencing (ATAC-seq) values for vtRNAs in different tumors present in PANCAN TCGA dataset (385 tumors samples across 23 cancer types) expressed as log2 normalized values. The box plots show the median and the lower and upper quartile, and the whiskers the 2.5 and 97.5 percentile of the distribution. **C.** The assay for transposase-accessible chromatin using sequencing (ATAC-seq) values for vtRNAs in 21 different tumors with at least 5 tumor samples present in Pan-Cancer TCGA dataset (385 tumors samples) expressed as log2 normalized values. The chart shows the average and standard error for each vtRNAs in each tumor.

Supplementary Figure 4. VtRNA promoter's chromatin accessibility (ATAC-seq and DNA methylation) average values for tumor tissues from Pan-Cancer TCGA dataset. The average vtRNA promoters ATAC-seq and DNA methylation values for each primary tumor tissue type (21 tissues). VtRNA1-1 (A), vtRNA1-2 (B), vtRNA1-3 (C) and vtRNA2-1 (D). The Spearman correlation between averages ATAC-seq and DNA methylation values was calculated for each vtRNA. The chart shows the average and standard deviation for each tissue with at least five samples available at Pan-Cancer TCGA.

Supplementary Figure 5. VtRNA promoter's DNA methylation values for each normal and tumor tissues from Pan-Cancer TCGA dataset. The average promoter DNA methylation beta-values of normal adjacent (Normal) and all primary tumor (Tumor) tissues (16 tissue types) for vtRNA1-1 (A), vtRNA1-2 (B), vtRNA1-3 (C) and vtRNA2-1 (D) is shown. The Spearman correlation between tissue averages DNA methylation normalized beta values for each vtRNA was calculated. The chart shows the average and standard deviation for each tissue with at least five samples available at Pan-Cancer TCGA.

Supplementary Table 1. ATAC-seq and DNA methylation data for vtRNA promoters in primary tumors and normal adjacent tissue samples. Excel spreadsheets: ATAC-seq_data_500bp: All ATAC-seq data of vtRNAs promoter (500bp) data for primary tumor samples; DNA_methylation_500bp: All DNA methylation data and ATAC-

seq data of vtRNAs promoter (500bp) data for total primary tumor and normal adjacent samples; DNA_methylation_NORMAL: All DNA methylation data of vtRNAs promoter (500bp) data for normal adjacent samples; DNA_methylation_TUMOR: All DNA methylation data of vtRNAs promoter (500bp) data for primary tumor samples; Normal_&_Tumor_matched: All DNA methylation data of vtRNAs promoter (500bp) data for primary tumor and normal adjacent samples.

Supplementary Table 2. VtRNAs Transcription Factors Binding and KEGG enriched terms. Excel spreadsheets: Binding_Factors: Transcription factors identified in the cell line K562 as ChIP-seq Peaks by ENCODE 3 project; KEGG_terms: enriched KEGG pathway terms (FDR < 0.05).

Supplementary Table 3. DNA methylation, ATAC-seq data and associated survival data for primary tumors. Excel spreadsheets: DNA-methylation_Survival_data: All DNA methylation data of vtRNAs promoter (500bp) and survival data for primary tumor samples; ATAC-seq_Survival_data: ATAC-seq data of vtRNAs promoter (500bp) and survival data for primary tumor samples.

Supplementary Table 4. Correlation of ATAC-seq values between vtRNA and all genome promoters in primary tumor samples. Excel spreadsheets: Spearman_correlation: Spearman correlation values of all promoter genes and vtRNAs in primary tumors samples; vtRNA1-1_cluster: Pathway enrichment and cluster chromosome localization data of vtRNA1-1; vtRNA1-2_cluster: Pathway enrichment and cluster chromosome localization data of vtRNA1-2; vtRNA1-3_cluster: Pathway enrichment and cluster chromosome localization data of vtRNA1-3; vtRNA2-1_cluster: Pathway enrichment and cluster chromosome localization data of vtRNA2-1.

Supplementary Table 5. ATAC-seq and DNA methylation data for vtRNA promoters in primary tumors and the associated Immune Subtypes data. Excel spreadsheets: DNA_methylation_data: All DNA methylation data of vtRNAs promoter (500bp) and Immune Subtypes data for primary tumor samples; DNA_methylation_Spearman_corr: Spearman correlation values of DNA methylation data of vtRNAs promoter (500bp) and Immune Subtypes data for primary tumor samples; ATAC-seq_data: All ATAC-seq data of vtRNAs promoter (500bp) and Immune Subtypes data for primary tumor samples; ATAC-seq_Spearman_corr: Spearman correlation values of ATAC-seq data of vtRNAs promoter (500bp) and Immune Subtypes data for primary tumor samples.

Abbreviations:

vtRNA	vault RNA
TCGA	The Cancer Genome Atlas
ATAC-seq	Assay for Transposase Accessible Chromatin with high-throughput sequencing
NoMe-seq	Nucleosome Occupancy and Methylome sequencing
OG	Oncogene
TSG	Tumor Suppressor Gene
TCGA	The Cancer Genome Atlas
TSS	Transcription Start Site
Ave	Average
SD	Standard Deviation
ACC	Adrenocortical carcinoma
BLCA	Bladder Urothelial Carcinoma

BRCA	Breast invasive carcinoma
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	Cholangiocarcinoma
COAD	Colon adenocarcinoma
DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma
ESCA	Esophageal carcinoma
GBM	Glioblastoma multiforme
HNSC	Head and Neck squamous cell carcinoma
KICH	Kidney Chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
LGG	Brain Lower Grade Glioma
LIHC	Liver hepatocellular carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
MESO	Mesothelioma
OV	Ovarian serous cystadenocarcinoma
PAAD	Pancreatic adenocarcinoma
PCPG	Pheochromocytoma and Paraganglioma
PRAD	Prostate adenocarcinoma
READ	Rectum adenocarcinoma
SARC	Sarcoma
SKCM	Skin Cutaneous Melanoma
STAD	Stomach adenocarcinoma
TGCT	Testicular Germ Cell Tumors
THCA	Thyroid carcinoma
THYM	Thymoma
UCEC	Uterine Corpus Endometrial Carcinoma
UCS	Uterine Carcinosarcoma
UVM	Uveal Melanoma

Author Contributions: Conceptualization, M.A.D.; methodology, R.S.F. and M.A.D.; software, R.S.F.; formal analysis, R.S.F.; investigation, R.S.F. and M.A.D.; resources, M.A.D.; data curation, M.A.D. and R.S.F.; writing—original draft preparation, R.S.F.; writing—review and editing, M.A.D.; visualization, R.S.F.; supervision, M.A.D.; project administration, M.A.D.; funding acquisition, M.A.D.

All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by Agencia Nacional de Investigación e Innovación (ANII), Comisión Sectorial de Investigación Científica (CSIC), Comisión Académica de posgrado (CAP) and Programa de Desarrollo de las Ciencias Básicas (PEDECIBA) from Uruguay.

Conflicts of Interest: The authors declare no conflict of interest.

References:

- Aakula, A., Kohonen, P., Leivonen, S.K., Mäkelä, R., Hintsanen, P., Mpindi, J.P., Martens-Uzunova, E., Aittokallio, T., Jenster, G., Perälä, M., et al. (2015). Systematic Identification of MicroRNAs That Impact on Proliferation of Prostate Cancer Cells and Display Changed Expression in Tumor Tissue. *Eur. Urol.* 69, 1120–1128.
- Ahn, J.H., Lee, H.S., Lee, J.S., Lee, Y.S., Park, J.L., Kim, S.Y., Hwang, J.A., Kunkeaw, N., Jung, S.Y., Kim, T.J., et al. (2018). Nc886 is induced by TGF- β and suppresses the microRNA pathway in ovarian cancer. *Nat. Commun.* 9.
- Amort, M., Nachbauer, B., Tuzlak, S., Kieser, A., Schepers, A., Villunger, A., and Polacek, N. (2015). Expression of the vault RNA protects cells from undergoing apoptosis. *Nat. Commun.* 6, 7030.
- Attar, N., and Kurdistani, S.K. (2017). Exploitation of EP300 and CREBBP lysine acetyltransferases by cancer. *Cold Spring Harb. Perspect. Med.* 7.
- Bi, N., Cao, J., Song, Y., Shen, J., Liu, W., Fan, J., He, J., Shi, Y., Zhang, X., Lu, N., et al. (2014). A MicroRNA signature predicts survival in early stage small-cell lung cancer treated with surgery and adjuvant chemotherapy. *PLoS One* 9, 1–8.
- Bracher, L., Ferro, I., Pulido-Quetglas, C., Ruepp, M.D., Johnson, R., and Polacek, N. (2020). Human vtRNA1-1 levels modulate signaling pathways and regulate apoptosis in human cancer cells. *Biomolecules* 10.
- Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218.
- Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* 2015, 21.29.1-21.29.9.
- Burke, J.M., Kincaid, R.P., Nottingham, R.M., Lambowitz, A.M., and Sullivan, C.S. (2016). DUSP11 activity on triphosphorylated transcripts promotes Argonaute association with noncanonical viral microRNAs and regulates steady-state levels of cellular noncoding RNAs. *Genes Dev.* 30, 2076–2092.
- Calderon, B.M., and Conn, G.L. (2018). A human cellular noncoding RNA activates the antiviral protein 2'-5'-oligoadenylate synthetase 1. *J. Biol. Chem.* 293, 16115–16124.
- Canella, D., Praz, V., Reina, J.H., Cousin, P., and Hernandez, N. (2010). Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res.* 20, 710–721.
- Cao, J., Song, Y., Bi, N., Shen, J., Liu, W., Fan, J., Sun, G., Tong, T., He, J., Shi, Y., et al. (2013). DNA methylation-mediated repression of miR-886-3p predicts poor outcome of human small cell lung cancer. *Cancer Res.* 73, 3326–3335.
- Carpenter, B.L., Zhou, W., Madaj, Z., DeWitt, A.K., Ross, J.P., Grønbaek, K., Liang, G., Clark,

- S.J., Molloy, P.L., and Jones, P.A. (2018). Mother–child transmission of epigenetic information by tunable polymorphic imprinting. *Proc. Natl. Acad. Sci.* *115*, E11970–E11977.
- Chen, J., OuYang, H., An, X., and Liu, S. (2018). Vault RNAs partially induces drug resistance of human tumor cells MCF-7 by binding to the RNA/DNA-binding protein PSF and inducing oncogene GAGE6. *PLoS One* *13*, e0191325.
- Chen, X., Shen, Y., Draper, W., Buenrostro, J.D., Litzenburger, U., Cho, S.W., Satpathy, A.T., Carter, A.C., Ghosh, R.P., East-Seletsky, A., et al. (2016). ATAC-se reveals the accessible genome by transposase-mediated imaging and sequencing. *Nat. Methods* *13*, 1013–1020.
- Corces, M.R., Granja, J.M., Shams, S., Louie, B.H., Seoane, J.A., Zhou, W., Silva, T.C., Groeneveld, C., Wong, C.K., Cho, S.W., et al. (2018). The chromatin accessibility landscape of primary human cancers. *Science* (80-.). 362.
- Dettmer, M.S., Perren, A., Moch, H., Komminoth, P., Nikiforov, Y.E., and Nikiforova, M.N. (2014). MicroRNA profile of poorly differentiated thyroid carcinomas: New diagnostic and prognostic insights. *J. Mol. Endocrinol.* *52*, 181–189.
- van Dijk, S.J., Peters, T.J., Buckley, M., Zhou, J., Jones, P.A., Gibson, R.A., Makrides, M., Muhlhauser, B.S., and Molloy, P.L. (2018). DNA methylation in blood from neonatal screening cards and the association with BMI and insulin sensitivity in early childhood. *Int. J. Obes.* *42*, 28–35.
- Duren, Z., Chen, X., Jiang, R., Wang, Y., and Wong, W.H. (2017). Modeling gene regulation from paired expression and chromatin accessibility data. *Proc. Natl. Acad. Sci. U. S. A.* *114*, E4914–E4923.
- Ertosun, M.G., Hapil, F.Z., and Osman Nidai, O.Z.E.S. (2016). E2F1 transcription factor and its impact on growth factor and cytokine signaling. *Cytokine Growth Factor Rev.* *31*, 17–25.
- Fendler, A., Jung, M., Stephan, C., Honey, R.J., Stewart, R.J., Pace, K.T., Erbersdobler, A., Samaan, S., Jung, K., and Yousef, G.M. (2011). miRNAs can predict prostate cancer biochemical relapse and are involved in tumor progression. *Int. J. Oncol.* *39*, 1183–1192.
- Fort, R.S., Mathó, C., Geraldo, M.V., Ottati, M.C., Yamashita, A.S., Saito, K.C., Leite, K.R.M., Méndez, M., Maedo, N., Méndez, L., et al. (2018). Nc886 is epigenetically repressed in prostate cancer and acts as a tumor suppressor through the inhibition of cell growth. *BMC Cancer* *18*, 127.
- Fort, R.S., Garat, B., Sotelo-Silveira, J.R., and Duhagon, M.A. (2020). vtRNA2-1/nc886 Produces a Small RNA That Contributes to Its Tumor Suppression Action through the microRNA Pathway in Prostate Cancer. *Non-Coding RNA* *6*, 7.
- Fuchs, O. (2012). Important genes in the pathogenesis of 5q- syndrome and their connection with ribosomal stress and the innate immune system pathway. *Leuk. Res. Treatment* *2012*, 179402.
- Gao, W., Shen, H., Liu, L., Xu, J., Xu, J., and Shu, Y. (2011). MiR-21 overexpression in human primary squamous cell lung carcinoma is associated with poor patient prognosis. *J. Cancer Res. Clin. Oncol.* *137*, 557–566.
- Goldman, M., Craft, B., Swatloski, T., Ellrott, K., Cline, M., Diekhans, M., Ma, S., Wilks, C., Stuart, J., Haussler, D., et al. (2013). The UCSC Cancer Genomics Browser: update 2013. *Nucleic Acids Res.* *41*, D949–D954.
- Goldman, M.J., Craft, B., Hastie, M., Repečka, K., McDade, F., Kamath, A., Banerjee, A., Luo, Y., Rogers, D., Brooks, A.N., et al. (2020). Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* *38*, 675–678.
- Golec, E., Lind, L., Qayyum, M., Blom, A.M., and King, B.C. (2019). The Noncoding RNA nc886

- Regulates PKR Signaling and Cytokine Production in Human Cells. *J. Immunol.* 202, 131–141.
- Gopinath, S.C.B., Matsugami, A., Katahira, M., and Kumar, P.K.R. (2005). Human vault-associated non-coding RNAs bind to mitoxantrone, a chemotherapeutic compound. *Nucleic Acids Res.* 33, 4874–4881.
- Gopinath, S.C.B., Wadhwa, R., and Kumar, P.K.R. (2010). Expression of noncoding vault RNA in human malignant cells and its importance in mitoxantrone resistance. *Mol. Cancer Res.* 8, 1536–1546.
- Helbo, A.S., Treppendahl, M., Aslan, D., Dimopoulos, K., Nandrup-Bus, C., Holm, M.S., Andersen, M.K., Liang, G., Kristensen, L.S., and Grønbæk, K. (2015). Hypermethylation of the VTRNA1-3 promoter is associated with poor outcome in lower risk myelodysplastic syndrome patients. *Genes (Basel).* 6, 977–990.
- Helbo, A.S., Lay, F.D., Jones, P.A., Liang, G., and Grønbæk, K. (2017). Nucleosome Positioning and NDR Structure at RNA Polymerase III Promoters. *Sci. Rep.* 7.
- Horos, R., Büscher, M., Kleinendorst, R., Alleaume, A.-M., Tarafder, A.K., Schwarzl, T., Dziuba, D., Tischer, C., Zielonka, E.M., Adak, A., et al. (2019). The Small Non-coding Vault RNA1-1 Acts as a Riboregulator of Autophagy. *Cell* 176, 1054-1067.e12.
- Horsley, V., and Pavlath, G.K. (2002). NFAT: Ubiquitous regulator of cell differentiation and adaptation. *J. Cell Biol.* 156, 771–774.
- Hu, Z., Zhang, H., Tang, L., Lou, M., and Geng, Y. (2017). Silencing nc886, a Non-Coding RNA, Induces Apoptosis of Human Endometrial Cancer Cells-1A In Vitro. *Med. Sci. Monit.* 23, 1317–1324.
- Hussain, S., Sajini, A.A., Blanco, S., Dietmann, S., Lombard, P., Sugimoto, Y., Paramor, M., Gleeson, J.G., Odom, D.T., Ule, J., et al. (2013). NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Rep.* 4, 255–261.
- Im, W.R., Lee, H.-S., Lee, Y.-S., Lee, J.-S., Jang, H.-J., Kim, S.-Y., Park, J.-L., Lee, Y., Kim, M.S., Lee, J.M., et al. (2020). A Regulatory Noncoding RNA, nc886, Suppresses Esophageal Cancer by Inhibiting the AKT Pathway and Cell Cycle Progression. *Cells* 9, 801.
- Jean, D., Tellez, C., Huang, S., Davis, D.W., Bruns, C.J., McConkey, D.J., Hinrichs, S.H., and Bar-Eli, M. (2000). Inhibition of tumor growth and metastasis of human melanoma by intracellular anti-ATF-1 single chain Fv fragment. *Oncogene* 19, 2721–2730.
- Jeon, S.H., Johnson, B.H., and Lee, Y.S. (2012a). A tumor surveillance model: A non-coding RNA senses neoplastic cells and its protein partner signals cell death. *Int. J. Mol. Sci.* 13, 13134–13139.
- Jeon, S.H., Lee, K., Lee, K.S., Kunkeaw, N., Johnson, B.H., Holthauzen, L.M.F., Gong, B., Leelayuwat, C., and Lee, Y.S. (2012b). Characterization of the direct physical interaction of nc886, a cellular non-coding RNA, and PKR. *FEBS Lett.* 586, 3477–3484.
- Jeronimo, S.M.B., Holst, A.K.B., Jamieson, S.E., Francis, R., Martins, D.R.A., Bezerra, F.L., Ettinger, N.A., Nascimento, E.T., Monteiro, G.R., Lacerda, H.G., et al. (2007). Genes at human chromosome 5q31.1 regulate delayed-type hypersensitivity responses associated with *Leishmania chagasi* infection. *Genes Immun.* 8, 539–551.
- Joo, J.E., Dowty, J.G., Milne, R.L., Wong, E.M., Dugué, P.A., English, D., Hopper, J.L., Goldgar, D.E., Giles, G.G., Southey, M.C., et al. (2018). Heritable DNA methylation marks associated with susceptibility to breast cancer. *Nat. Commun.* 9.
- Kedersha, N.L., and Rome, L.H. (1986). Isolation and characterization of a novel ribonucleoprotein particle: Large structures contain a single species of small RNA. *J. Cell Biol.*

103, 699–709.

Kedersha, N.L., Heuser, J.E., Chugani, D.C., and Rome, L.H. (1991). Vaults. III. Vault ribonucleoprotein particles open into flower-like structures with octagonal symmetry. *J. Cell Biol.* 112, 225–235.

Kickhoefer, V.A., Rajavel, K.S., Scheffer, G.L., Dalton, W.S., Scheper, R.J., and Rome, L.H. (1998). Vaults are up-regulated in multidrug-resistant cancer cell lines. *J. Biol. Chem.* 273, 8971–8974.

Kickhoefer, V.A., Poderycki, M.J., Chan, E.K.L., and Rome, L.H. (2002). The La RNA-binding protein interacts with the vault RNA and is a vault-associated protein. *J. Biol. Chem.* 277, 41282–41286.

Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* 20, 207–220.

Kuleshov, M. V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44, W90–W97.

Kunkeaw, N., Jeon, S.H., Lee, K., Johnson, B.H., Tanasanvimon, S., Javle, M., Pairojkul, C., Chamgramol, Y., Wongfieng, W., Gong, B., et al. (2012). Cell death/proliferation roles for nc886, a non-coding RNA, in the protein kinase R pathway in cholangiocarcinoma. *Oncogene* 32, 3722–3731.

Kunkeaw, N., Lee, Y.S., Im, W.R., Jang, J.J., Song, M.J., Yang, B., Park, J.L., Kim, S.Y., Ku, Y., Kim, Y., et al. (2019). Mechanism mediated by a noncoding RNA, nc886, in the cytotoxicity of a DNA-reactive compound. *Proc. Natl. Acad. Sci. U. S. A.* 116, 8289–8294.

Łabno, A., Warkocki, Z., Kuliński, T., Krawczyk, P.S., Bijata, K., Tomecki, R., and Dziembowski, A. (2016). Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. *Nucleic Acids Res.* gkw649.

Lacy, D.A., Wang, Z.E., Symula, D.J., McArthur, C.J., Rubin, E.M., Frazer, K.A., and Locksley, R.M. (2000). Faithful expression of the human 5q31 cytokine cluster in transgenic mice. *J. Immunol.* 164, 4569–4574.

Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., et al. (2007). A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 129, 1401–1414.

Lee, Y.S. (2015). A Novel Type of Non-coding RNA, nc886, Implicated in Tumor Sensing and Suppression. *Genomics Inform.* 13, 26–30.

Lee, E.K., Hong, S.H., Shin, S., Lee, H.S., Lee, J.S., Park, E.J., Choi, S.S., Min, J.W., Park, D., Hwang, J.A., et al. (2016). Nc886, a non-coding RNA and suppressor of PKR, exerts an oncogenic function in thyroid cancer. *Oncotarget* 7, 75000–75012.

Lee, H.-S., Lee, K., Jang, H.-J., Lee, G.K., Park, J.-L., Kim, S.-Y., Kim, S.-B., Johnson, B.H., Zo, J.I., Lee, J.-S., et al. (2014a). Epigenetic silencing of the non-coding RNA nc886 provokes oncogenes during human esophageal tumorigenesis. *Oncotarget* 5, 3472–3481.

Lee, K.-S., Park, J.-L., Lee, K., Richardson, L.E., Johnson, B.H., Lee, H.-S., Lee, J.-S., Kim, S.-B., Kwon, O.-H., Song, K.S., et al. (2014b). nc886, a non-coding RNA of anti-proliferative role, is suppressed by CpG DNA methylation in human gastric cancer. *Oncotarget* 5, 3944–3955.

Lee, K., Kunkeaw, N., Jeon, S.H., Lee, I., Johnson, B.H., Kang, G.-Y., Bang, J.Y., Park, H.S., Leelayuwat, C., and Lee, Y.S. (2011). Precursor miR-886, a novel noncoding RNA repressed in cancer, associates with PKR and modulates its activity. *RNA* 17, 1076–1089.

Lee, K.S., Shin, S., Cho, E., Im, W.K., Jeon, S.H., Kim, Y., Park, D., Fréchet, M., Chajra, H., and

- Jung, E. (2019a). nc886, a non-coding RNA, inhibits UVB-induced MMP-9 and COX-2 expression via the PKR pathway in human keratinocytes. *Biochem. Biophys. Res. Commun.* 512, 647–652.
- Lee, Y.S., Kunkeaw, N., and Lee, Y.S. (2019b). Protein kinase R and its cellular regulators in cancer: An active player or a surveillant? *Wiley Interdiscip. Rev. RNA* e1558.
- Legendre, C.R., Demeure, M.J., Whitsett, T.G., Gooden, G.C., Bussey, K.J., Jung, S., Waibhav, T., Kim, S., and Salhia, B. (2016). Pathway Implications of Aberrant Global Methylation in Adrenocortical Cancer. *PLoS One* 11, e0150629.
- Lei, J., Xiao, J.H., Zhang, S.H., Liu, Z.Q., Huang, K., Luo, Z.P., Xiao, X.L., and Hong, Z.D. (2017). Non-coding RNA 886 promotes renal cell carcinoma growth and metastasis through the Janus kinase 2/signal transducer and activator of transcription 3 signaling pathway. *Mol. Med. Rep.* 16, 4273–4278.
- Li, B., Carey, M., and Workman, J.L. (2007). The Role of Chromatin during Transcription. *Cell* 128, 707–719.
- Li, F., Chen, Y., Zhang, Z., Ouyang, J., Wang, Y., Yan, R., Huang, S., Gao, G.F., Guo, G., and Chen, J.-L. (2015). Robust expression of vault RNAs induced by influenza A virus plays a critical role in suppression of PKR-mediated innate immunity. *Nucleic Acids Res.* 43, gkv1078.
- Li, J.-H., Wang, M., Zhang, R., Gao, W.-L., Meng, S.-H., Ma, X.-L., Hou, X.-H., and Feng, L.-M. (2017). E2F1-directed activation of nc886 mediates drug resistance in cervical cancer cells via regulation of major vault protein. *Int. J. Clin. Exp. Pathol.* 10, 9233–9242.
- Liu, G., Wang, W., Hu, S., Wang, X., and Zhang, Y. (2018). Inherited DNA methylation primes the establishment of accessible chromatin during genome activation. *Genome Res.* 28, 998–1007.
- Ma, H., Wang, M., Zhou, Y., Yang, J. jie, Wang, L.Y., Yang, R. hui, Wen, M. jie, and Kong, L. (2020). Noncoding RNA 886 alleviates tumor cellular immunological rejection in host C57BL/C mice. *Cancer Med.* 9, 5258–5271.
- Mashima, T., Kudo, M., Takada, Y., Matsugami, A., Gopinath, S.C.B., Kumar, P.K.R., and Katahira, M. (2008). Interactions between antitumor drugs and vault RNA. *Nucleic Acids Symp. Ser. (Oxf).* 217–218.
- Meng, C., Meng, C., Wei, Z., Wei, Z., Zhang, Y., Zhang, Y., Yan, L., Yan, L., He, H., He, H., et al. (2016). Regulation of cytochrome P450 3A4 by small vault RNA derived from the non-coding vault RNA1 of multidrug resistance-linked vault particle (D.A. Spandidos).
- Miñones-Moyano, E., Friedländer, M.R., Pallares, J., Kagerbauer, B., Porta, S., Escaramís, G., Ferrer, I., Estivill, X., and Martí, E. (2013). Upregulation of a small vault RNA (svtRNA2-1a) is an early event in Parkinson disease and induces neuronal dysfunction. *RNA Biol.* 10, 1093–1106.
- Moran, S., Arribas, C., and Esteller, M. (2016). Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* 8, 389–399.
- Naka, I., Nishida, N., Patarapotikul, J., Nuchnoi, P., Tokunaga, K., Hananantachai, H., Tsuchiya, N., and Ohashi, J. (2009). Identification of a haplotype block in the 5q31 cytokine gene cluster associated with the susceptibility to severe malaria. *Malar. J.* 8, 232.
- Nandy, C., Mrázek, J., Stoiber, H., Grässer, F.A., Hüttenhofer, A., and Polacek, N. (2009). Epstein-Barr Virus-Induced Expression of a Novel Human Vault RNA. *J. Mol. Biol.* 388, 776–784.
- Nikitina, T. V., Tischenko, L.I., and Schulz, W.A. (2011). Recent insights into regulation of transcription by RNA polymerase III and the cellular functions of its transcripts. *Biol. Chem.* 392,

395–404.

Nomura, M., Saito, K., Aihara, K., Nagae, G., Yamamoto, S., Tatsuno, K., Ueda, H., Fukuda, S., Umeda, T., Tanaka, S., et al. (2019). DNA demethylation is associated with malignant progression of lower-grade gliomas. *Sci. Rep.* 9, 1903.

Nordentoft, I., Birkenkamp-Demtroder, K., Agerbæk, M., Theodorescu, D., Ostfeld, M.S., Hartmann, A., Borre, M., Ørntoft, T.F., and Dyrskjød, L. (2012). miRNAs associated with chemosensitivity in cell lines and in advanced bladder cancer. *BMC Med.* ... 5, 40.

Okumura, T., Kojima, H., Miwa, T., Sekine, S., Hashimoto, I., Hojo, S., Nagata, T., and Shimada, Y. (2016). The expression of microRNA 574-3p as a predictor of postoperative outcome in patients with esophageal squamous cell carcinoma. *World J. Surg. Oncol.* 14, 228.

Paliwal, A., Temkin, A.M., Kerkel, K., Yale, A., Yotova, I., Drost, N., Lax, S., Nhan-Chang, C.L., Powell, C., Borczuk, A., et al. (2013). Comparative Anatomy of Chromosomal Domains with Imprinted and Non-Imprinted Allele-Specific DNA Methylation. *PLoS Genet.* 9.

Park, J.-L., Lee, Y.-S., Kunkeaw, N., Kim, S.-Y., Kim, I.-H., and Lee, Y.S. (2017). Epigenetic regulation of noncoding RNA transcription by mammalian RNA polymerase III. *Epigenomics* 9, 171–187.

Pazos Obregón, F., Soto, P., Lavín, J.L., Cortázar, A.R., Barrio, R., Aransay, A.M., and Cantera, R. (2018). Cluster Locator, online analysis and visualization of gene clustering. *Bioinformatics* 34, 3377–3379.

Persengiev, S.P., and Green, M.R. (2003). The role of ATF/CREB family members in cell growth, survival and apoptosis. *Apoptosis* 8, 225–228.

Persson, H., Kvist, A., Vallon-Christersson, J., Medstrand, P., Borg, A., and Rovira, C. (2009). The non-coding RNA of the multidrug resistance-linked vault particle encodes multiple regulatory small RNAs. *Nat. Cell Biol.* 11, 1268–1271.

Rechache, N.S., Wang, Y., Stevenson, H.S., Killian, J.K., Edelman, D.C., Merino, M., Zhang, L., Nilubol, N., Stratakis, C.A., Meltzer, P.S., et al. (2012). DNA methylation profiling identifies global methylation differences and markers of adrenocortical tumors. *J. Clin. Endocrinol. Metab.* 97, E1004-13.

Renauer, P., Coit, P., Jeffries, M.A., Merrill, J.T., McCune, W.J., Maksimowicz-McKinnon, K., and Sawalha, A.H. (2015). DNA methylation patterns in naïve CD4+ T cells identify epigenetic susceptibility loci for malar rash and discoid rash in systemic lupus erythematosus. *Lupus Sci. Med.* 2.

Richmond, R.C., Sharp, G.C., Herbert, G., Atkinson, C., Taylor, C., Bhattacharya, S., Campbell, D., Hall, M., Kazmi, N., Gaunt, T., et al. (2018). The long-term impact of folic acid in pregnancy on offspring DNA methylation: follow-up of the Aberdeen Folic Acid Supplementation Trial (AFASST). *Int. J. Epidemiol.*

Rodrigues, V., Piper, K., Couissinier-Paris, P., Bacelar, O., Dessein, H., and Dessein, A.J. (1999). Genetic control of schistosome infections by the SM1 locus of the 5q31-q33 region is linked to differentiation of type 2 helper T lymphocytes. *Infect. Immun.* 67, 4689–4692.

Rohini, M., Haritha Menon, A., and Selvamurugan, N. (2018). Role of activating transcription factor 3 and its interacting proteins under physiological and pathological conditions. *Int. J. Biol. Macromol.* 120, 310–317.

Romanelli, V., Nakabayashi, K., Vizoso, M., Moran, S., Iglesias-Platas, I., Sugahara, N., Simón, C., Hata, K., Esteller, M., Court, F., et al. (2014). Variable maternal methylation overlapping the nc886/vtRNA2-1 locus is locked between hypermethylated repeats and is frequently altered in cancer. *Epigenetics* 9, 783–790.

- Sajini, A.A., Choudhury, N.R., Wagner, R.E., Bornelöv, S., Selmi, T., Spanos, C., Dietmann, S., Rappsilber, J., Michlewski, G., and Frye, M. (2019). Loss of 5-methylcytosine alters the biogenesis of vault-derived small RNAs to coordinate epidermal differentiation. *Nat. Commun.* 10.
- Sallustio, F., Serino, G., Cox, S.N., Dalla Gassa, A., Curci, C., De Palma, G., Banelli, B., Zaza, G., Romani, M., and Schena, F.P. (2016). Aberrantly methylated DNA regions lead to low activation of CD4+ T-cells in IgA nephropathy. *Clin. Sci. (Lond).* 130, 733–746.
- Schou, J. V., Rossi, S., Jensen, B. V., Nielsen, D.L., Pfeiffer, P., Høgdall, E., Yilmaz, M., Tejpar, S., Delorenzi, M., Kruhøffer, M., et al. (2014). miR-345 in Metastatic Colorectal Cancer: A Non-Invasive Biomarker for Clinical Outcome in Non-KRAS Mutant Patients Treated with 3rd Line Cetuximab and Irinotecan. *PLoS One* 9, e99886.
- Sendler, E., Johnson, G.D., and Krawetz, S.A. (2011). Local and global factors affecting RNA sequencing analysis. *Anal. Biochem.* 419, 317–322.
- Shen, J., Zhou, W., Bi, N., Song, Y.-M., Zhang, F.-Q., Zhan, Q.-M., and Wang, L.-H. (2018). MicroRNA-886-3P functions as a tumor suppressor in small cell lung cancer. *Cancer Biol. Ther.* 19, 1185–1192.
- Silver, M.J., Kessler, N.J., Hennig, B.J., Dominguez-Salas, P., Laritsky, E., Baker, M.S., Coarfa, C., Hernandez-Vargas, H., Castelino, J.M., Routledge, M.N., et al. (2015). Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol* 16, 118.
- Stadler, P.F., Chen, J.J.L., Hackermüller, J., Hoffmann, S., Horn, F., Khaitovich, P., Kretzschmar, A.K., Mosig, A., Prohaska, S.J., Qi, X., et al. (2009). Evolution of vault RNAs. *Mol. Biol. Evol.* 26, 1975–1991.
- Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. *Nat. Rev. Genet.* 20, 631–656.
- Sun, Y., Miao, N., and Sun, T. (2019). Detect accessible chromatin using ATAC-sequencing, from principle to applications. *Hereditas* 156, 29.
- Suojalehto, H., Lindström, I., Majuri, M.L., Mitts, C., Karjalainen, J., Wolff, H., and Alenius, H. (2014). Altered microRNA expression of nasal mucosa in long-term asthma and allergic rhinitis. *Int. Arch. Allergy Immunol.* 163, 168–178.
- Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., et al. (2017). The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* 45, D362–D368.
- Tahiri, A., Leivonen, S.K., Lüders, T., Steinfeld, I., Aure, M.R., Geisler, J., Mäkelä, R., Nord, S., Riis, M.L.H., Yakhini, Z., et al. (2014). Deregulation of cancer-related miRNAs is a common event in both benign and malignant human breast tumors. *Carcinogenesis* 35, 76–85.
- Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.H., Porta-Pardo, E., Gao, G.F., Plaisier, C.L., Eddy, J.A., et al. (2018). The Immune Landscape of Cancer. *Immunity* 48, 812–830.e14.
- Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. *Nature* 489, 75–82.
- Treppendahl, M.B., Qiu, X., Søgaard, A., Yang, X., Nandrup-Bus, C., Hother, C., Andersen, M.K., Kjeldsen, L., Möllgaard, L., Hellström-Lindberg, E., et al. (2012). Allelic methylation levels of the noncoding VTRNA2-1 located on chromosome 5q31.1 predict outcome in AML. *Blood*

119, 206–216.

- Tsompana, M., and Buck, M.J. (2014). Chromatin accessibility: A window into the genome. *Epigenetics and Chromatin* 7, 1–16.
- Vilalta, A., Kickhoefer, V.A., Rome, L.H., and Johnson, D.L. (1994). The rat vault RNA gene contains a unique RNA polymerase III promoter composed of both external and internal elements that function synergistically. *J. Biol. Chem.* 269, 29752–29759.
- Weeding, E., and Sawalha, A.H. (2018). Deoxyribonucleic acid methylation in systemic lupus erythematosus: Implications for future clinical practice. *Front. Immunol.* 9, 1.
- Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R.M., Ozenberger, B.A., Ellrott, K., Sander, C., Stuart, J.M., Chang, K., Creighton, C.J., et al. (2013). The cancer genome atlas pan-cancer analysis project. *Nat. Genet.* 45, 1113–1120.
- Xiong, Y., Zhang, L., Holloway, A.K., Wu, X., Su, L., and Kebebew, E. (2011). MiR-886-3p regulates cell proliferation and migration, and is dysregulated in familial non-medullary thyroid cancer. *PLoS One* 6, 1–11.
- Yeganeh, M., and Hernandez, N. (2020). RNA polymerase III transcription as a disease factor. *Genes Dev.* 34, 865–882.
- Yu, M.C., Lee, C.W., Lin, C.H., Wu, C.H., Lee, Y.S., Tsai, C.L., and Tsai, C.N. (2020). Differential hypermethylation of the VTRNA2-1 promoter in hepatocellular carcinoma as a prognostic factor: Tumor marker prognostic study. *Int. J. Surg.* 79, 282–289.
- Yu, X.F., Zou, J., Bao, Z.J., and Dong, J. (2011). miR-93 suppresses proliferation and colony formation of human colon cancer stem cells. *World J Gastroenterol* 17, 4711–4717.
- Yu, Z., Chen, D., Su, Z., Li, Y., Yu, W., Zhang, Q., Yang, L., Li, C., Yang, S., Ni, L., et al. (2014). miR-886-3p upregulation in clear cell renal cell carcinoma regulates cell migration, proliferation and apoptosis by targeting PITX1. *Int. J. Mol. Med.* 34, 1409–1416.
- Zheng, G., Qin, Y., Clark, W.C., Dai, Q., Yi, C., He, C., Lambowitz, A.M., and Pan, T. (2015). Efficient and quantitative high-throughput tRNA sequencing. *Nat. Methods* 12, 835–837.
- van Zon, A., Mossink, M.H., Schoester, M., Scheffer, G.L., Scheper, R.J., Sonneveld, P., and Wiemer, E.A. (2001). Multiple human vault RNAs. Expression and association with the vault complex. *J. Biol. Chem.* 276, 37715–37721.

FIGURE 1

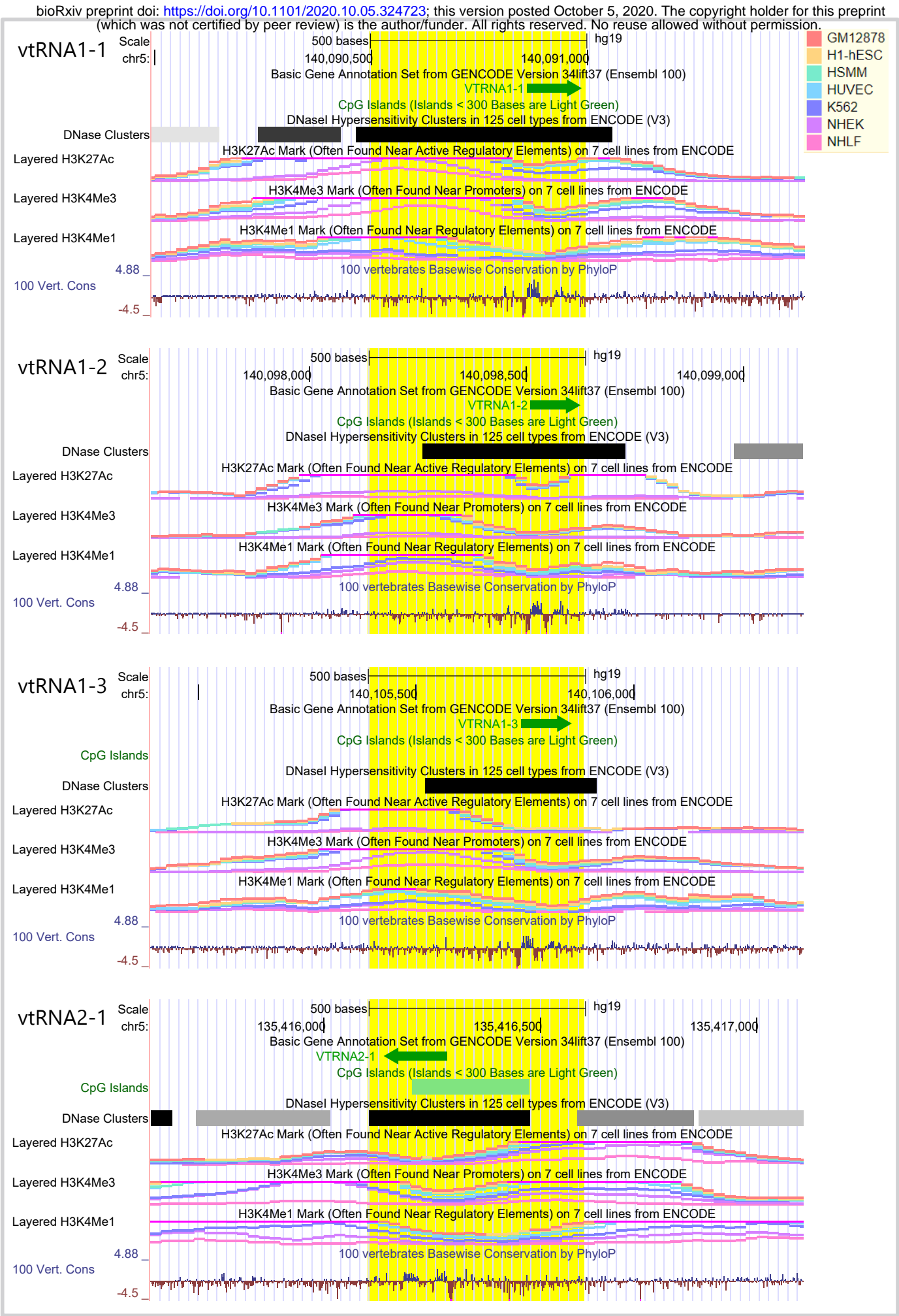


FIGURE 2

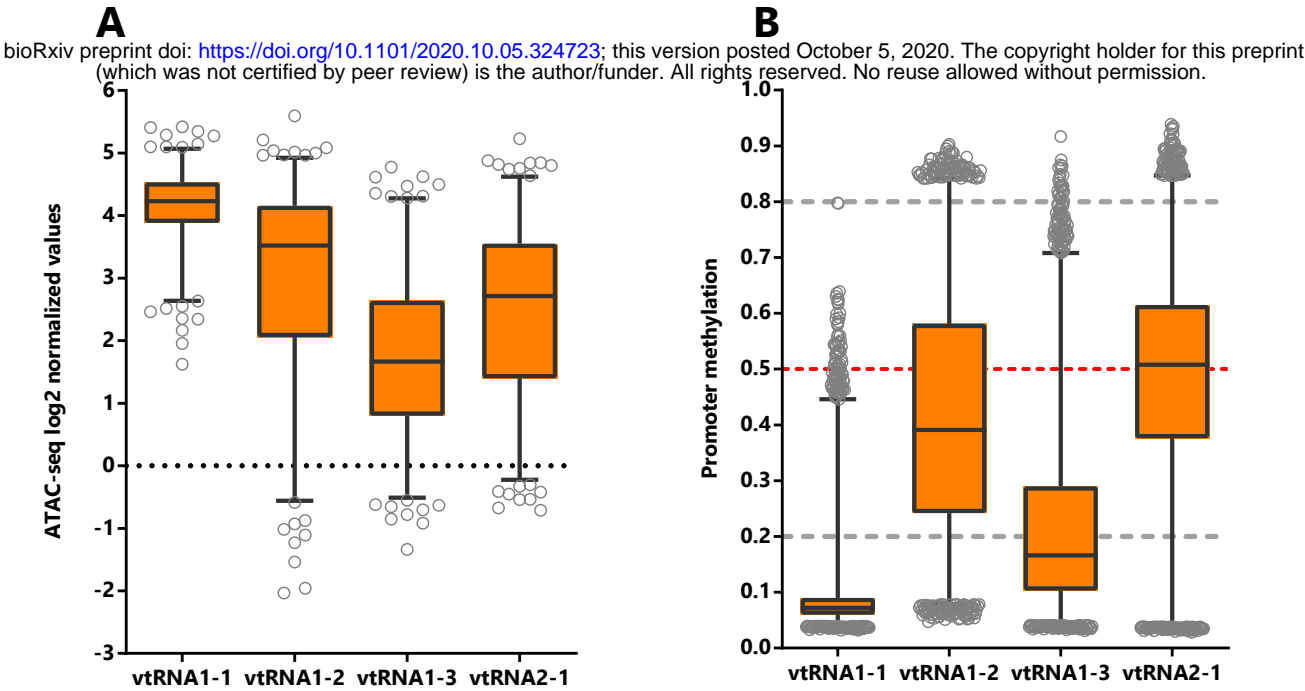


FIGURE 3

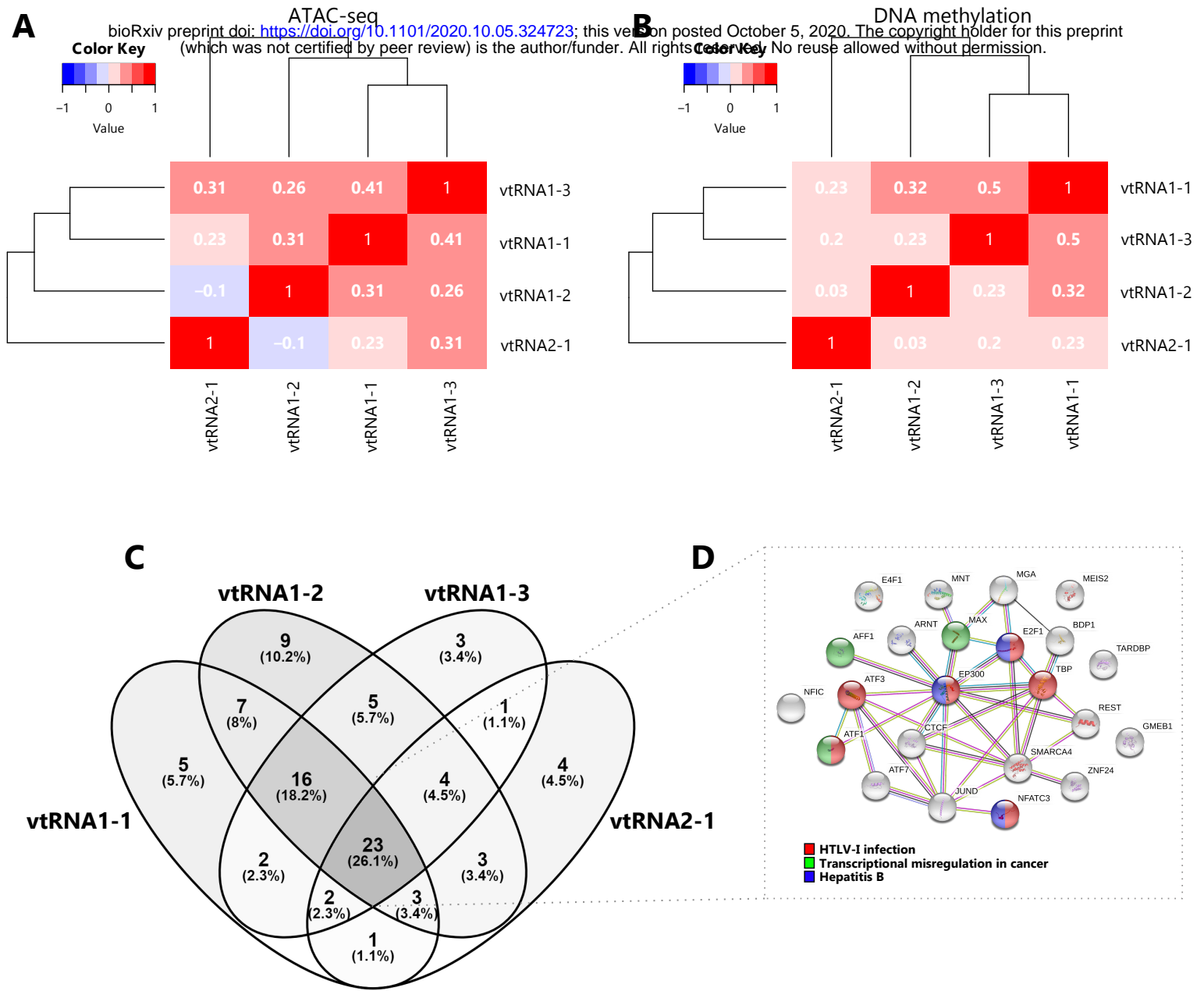


FIGURE 4

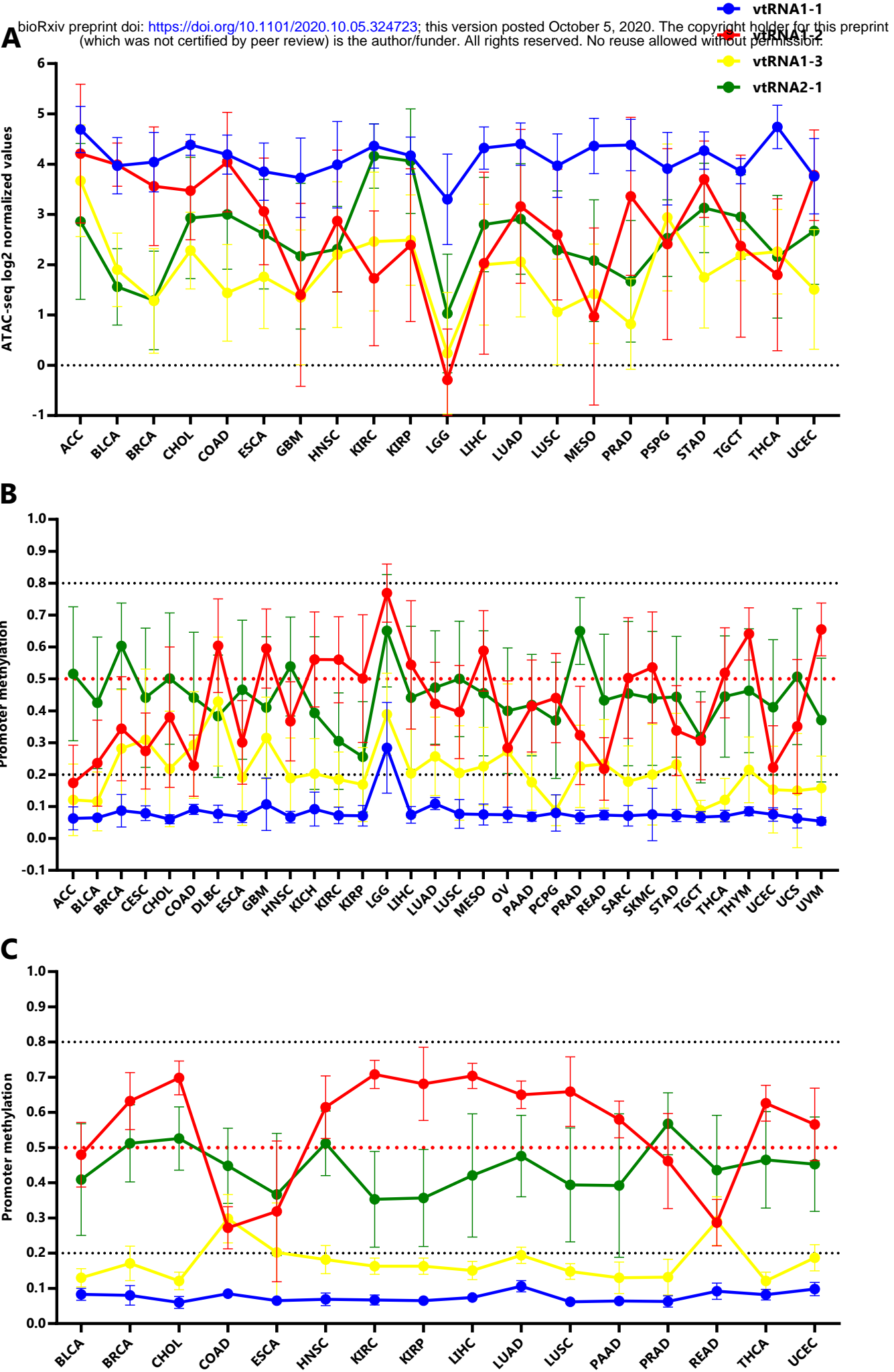


FIGURE 5

bioRxiv preprint doi: <https://doi.org/10.1101/2020.10.05.324723>; this version posted October 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

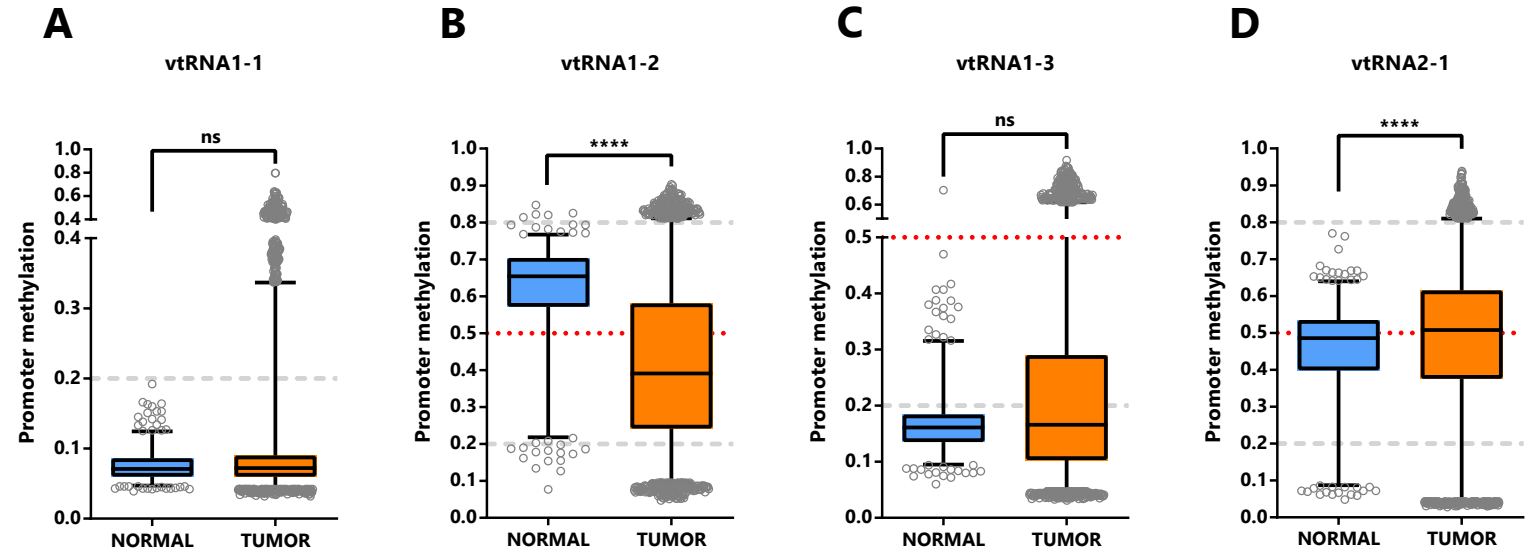


FIGURE 6

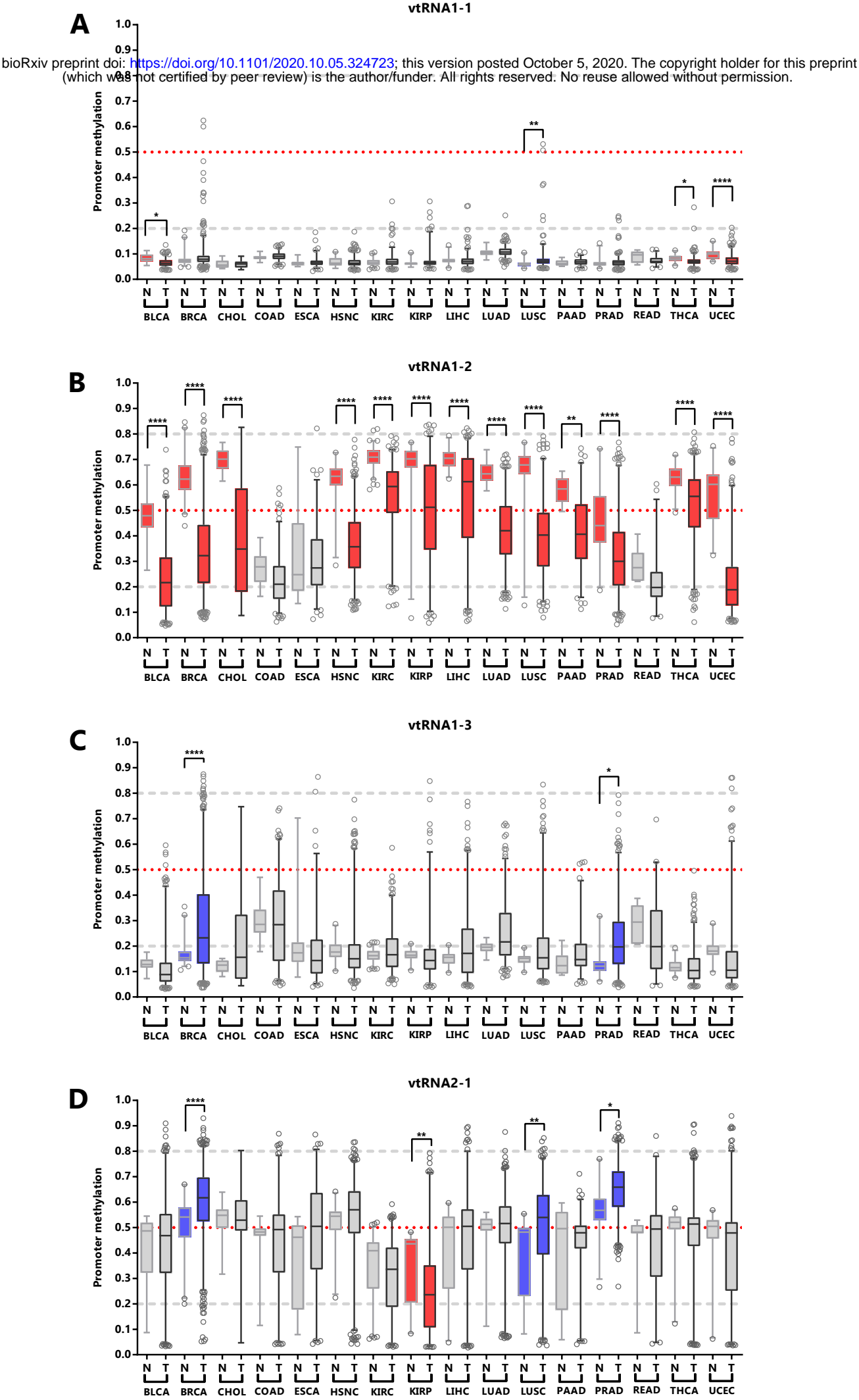


FIGURE 7

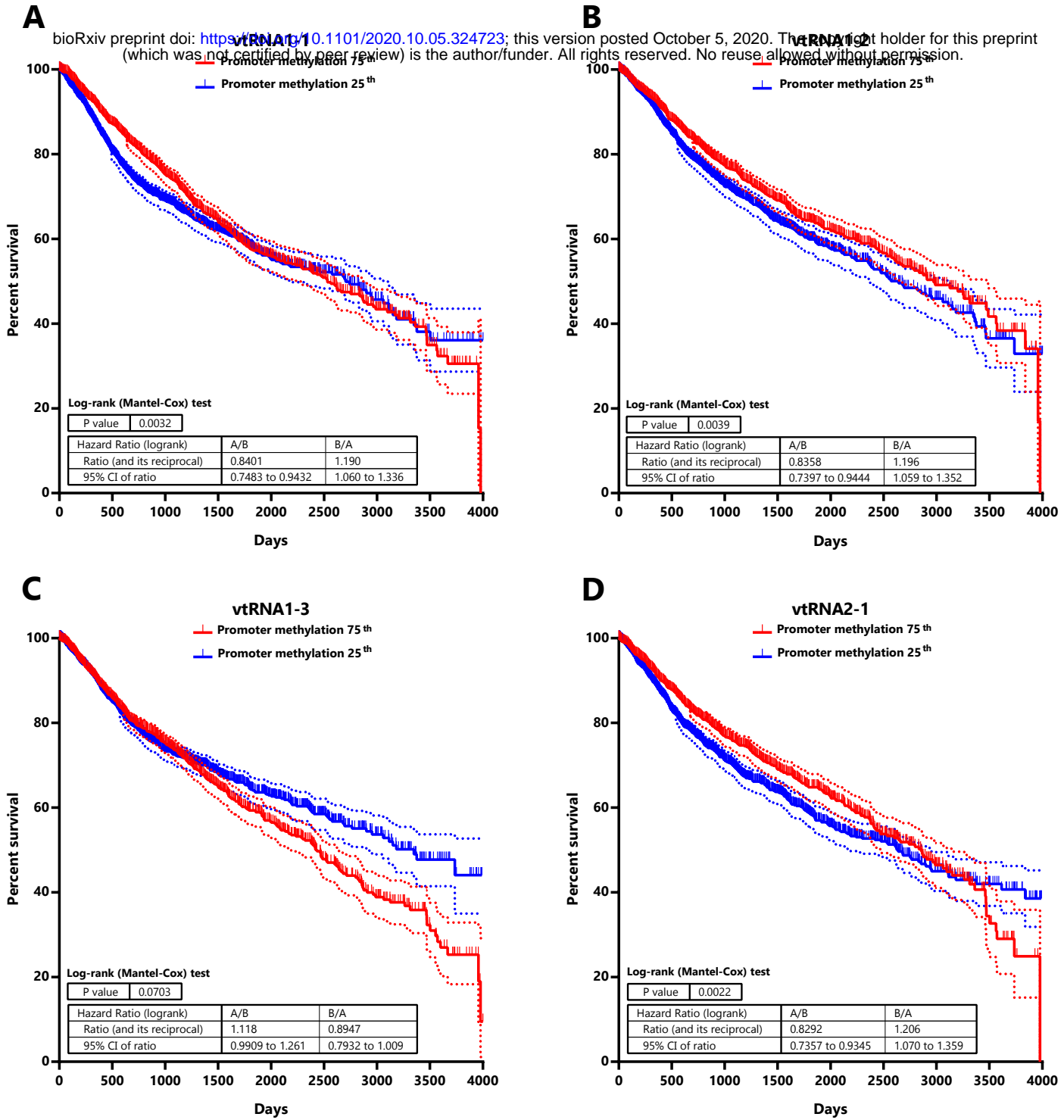


FIGURE 8

bioRxiv preprint doi: <https://doi.org/10.1101/2020.10.05.324725>; this version posted October 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

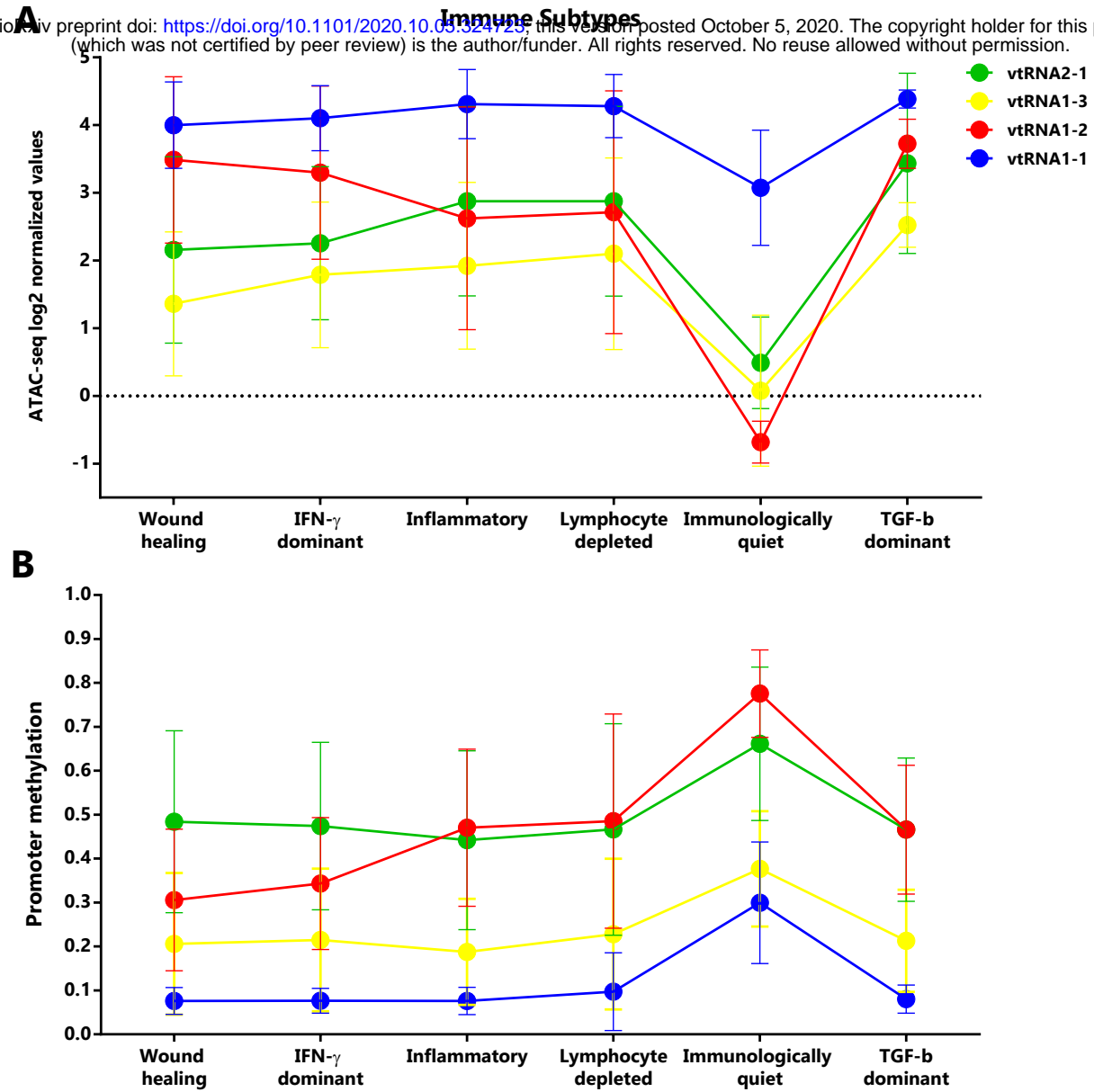


Table 1

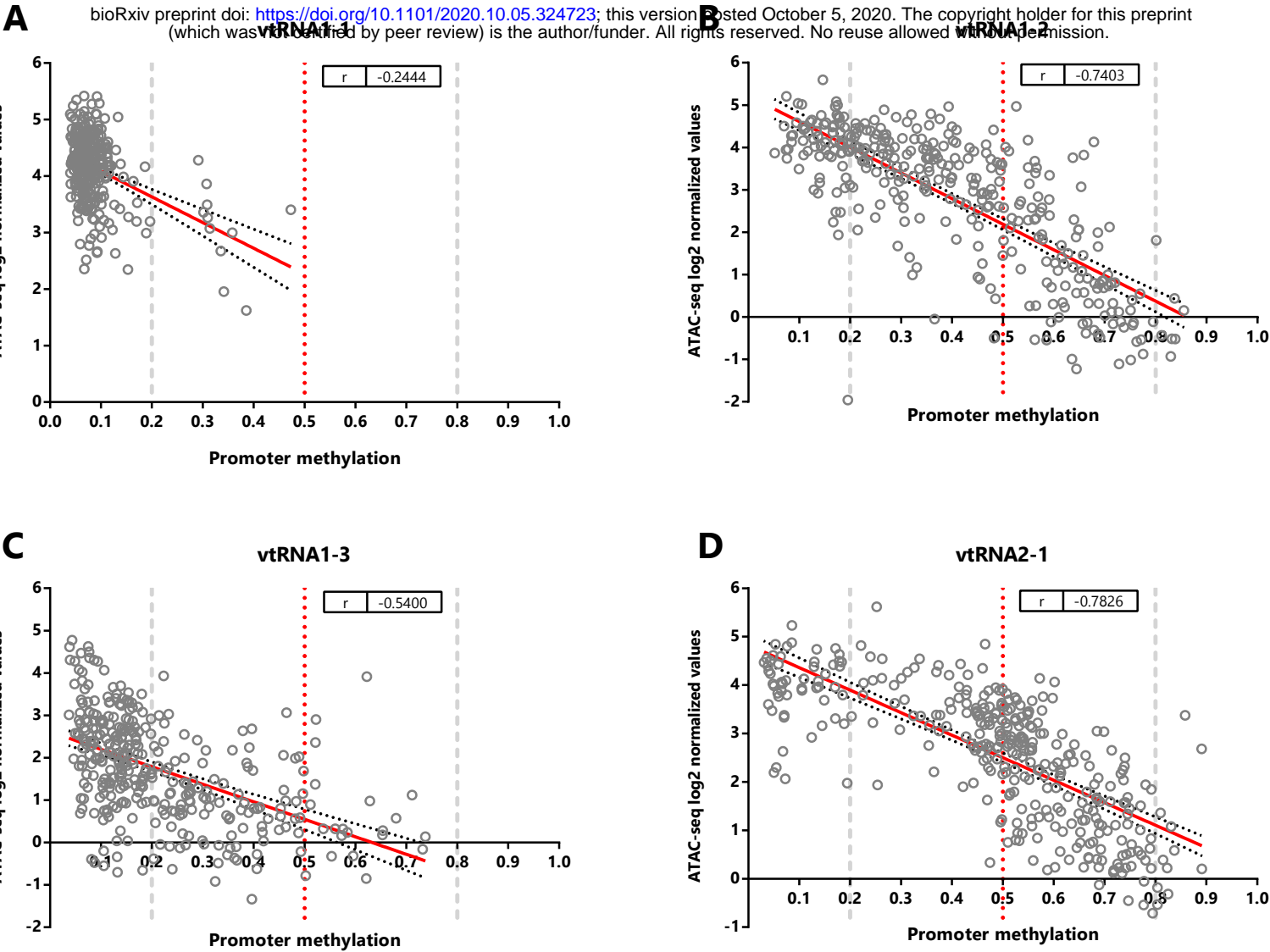
Average \pm SD			Spearman correlation (ATAC-seq vs DNA methylation)	
bioRxiv preprint doi: https://doi.org/10.1101/2020.10.05.31723 ; this version posted October 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.				
vtRNA	ATAC-seq	DNA methylation	<i>Rs</i>	<i>p-value</i>
vtRNA1-1	4.1 \pm 0.6	0.09 \pm 0.07	-0.24	< 0.0001
vtRNA1-2	2.9 \pm 1.6	0.42 \pm 0.20	-0.74	< 0.0001
vtRNA1-3	1.7 \pm 1.2	0.22 \pm 0.15	-0.54	< 0.0001
vtRNA2-1	2.5 \pm 1.4	0.47 \pm 0.21	-0.78	< 0.0001

Supplementary Figure 1

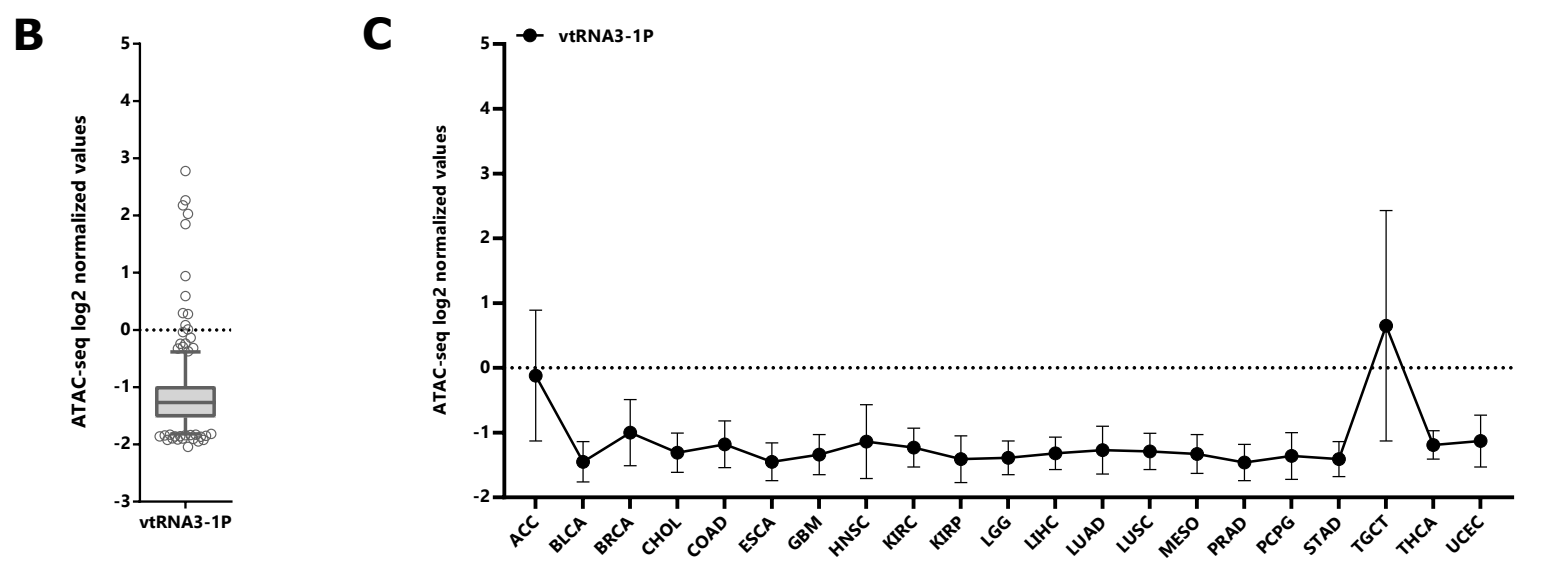
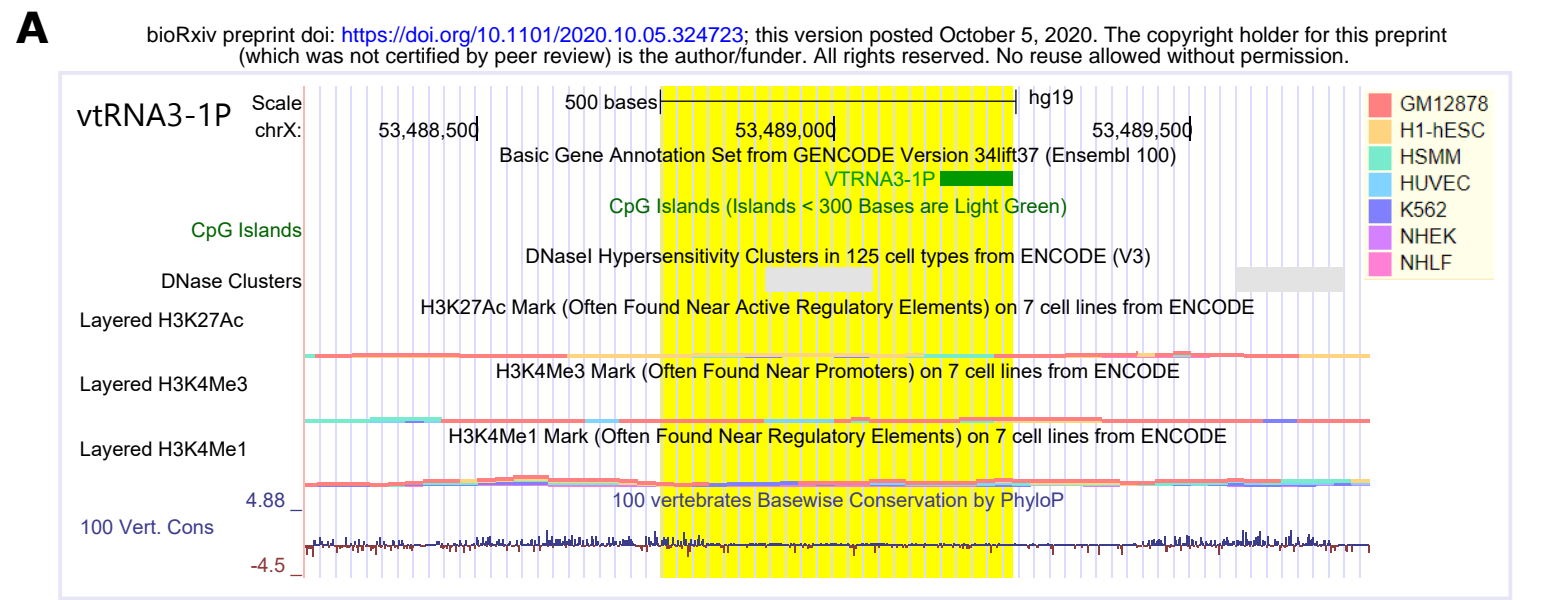
TISSUES	N° Samples			
	Primary Tumor (ATAC-seq)	Primary Tumor (DNA methylation)	Primary Tumor (ATAC-seq & DNA methylation)	Adjacent Normal (DNA methylation)
ACC	9	80	9	0
BLCA	9	415	9	21
BRCA	73	789	42	96
CESC	2	307	2	3
CHOL	5	36	2	9
COAD	37	307	22	38
DLBC	0	48	0	0
ESCA	18	185	18	16
GBM	9	140	9	2
HNSC	8	528	8	50
KICH	0	66	0	0
KIRC	16	322	16	160
KIRP	33	275	33	45
LGG	10	516	10	0
LIHC	17	377	17	50
LUAD	22	469	22	32
LUSC	16	370	16	42
MESO	7	87	7	0
OV	0	10	0	0
PAAD	0	184	0	10
PCPG	9	179	9	3
PRAD	26	502	26	50
READ	0	98	0	7
SARC	0	261	0	4
SKCM	4	104	4	2
STAD	21	395	21	2
TGCT	8	150	8	0
THCA	14	507	14	56
THYM	0	124	0	2
UCEC	12	435	5	46
UCS	0	57	0	0
UVM	0	80	0	0
TOTAL N°	385	8403	329	746

B	Percentages				
	TISSUES	Primary Tumor (ATAC-seq)	Primary Tumor (DNA methylation)	Primary Tumor (ATAC-seq & DNA methylation)	Adjacent Normal (DNA methylation)
	ACC	2	1	3	0
	BLCA	2	5	3	3
	BRCA	19	9	13	13
	CESC	1	4	1	0
	CHOL	1	0	1	1
	COAD	10	4	7	5
	DLBC	0	1	0	0
	ESCA	5	2	5	2
	GBM	2	2	3	0
	HNSC	2	6	2	7
	KICH	0	1	0	0
	KIRC	4	4	5	21
	KIRP	9	3	10	6
	LGG	3	6	3	0
	LIHC	4	4	5	7
	LUAD	6	6	7	4
	LUSC	4	4	5	6
	MESO	2	1	2	0
	OV	0	0	0	0
	PAAD	0	2	0	1
	PCPG	2	2	3	0
	PRAD	7	6	8	7
	READ	0	1	0	1
	SARC	0	3	0	1
	SKCM	1	1	1	0
	STAD	5	5	6	0
	TGCT	2	2	2	0
	THCA	4	6	4	8
	THYM	0	1	0	0
	UCEC	3	5	2	6
	UCS	0	1	0	0
	UVM	0	1	0	0
	TOTAL N°	100	100	100	100

Supplementary Figure 2

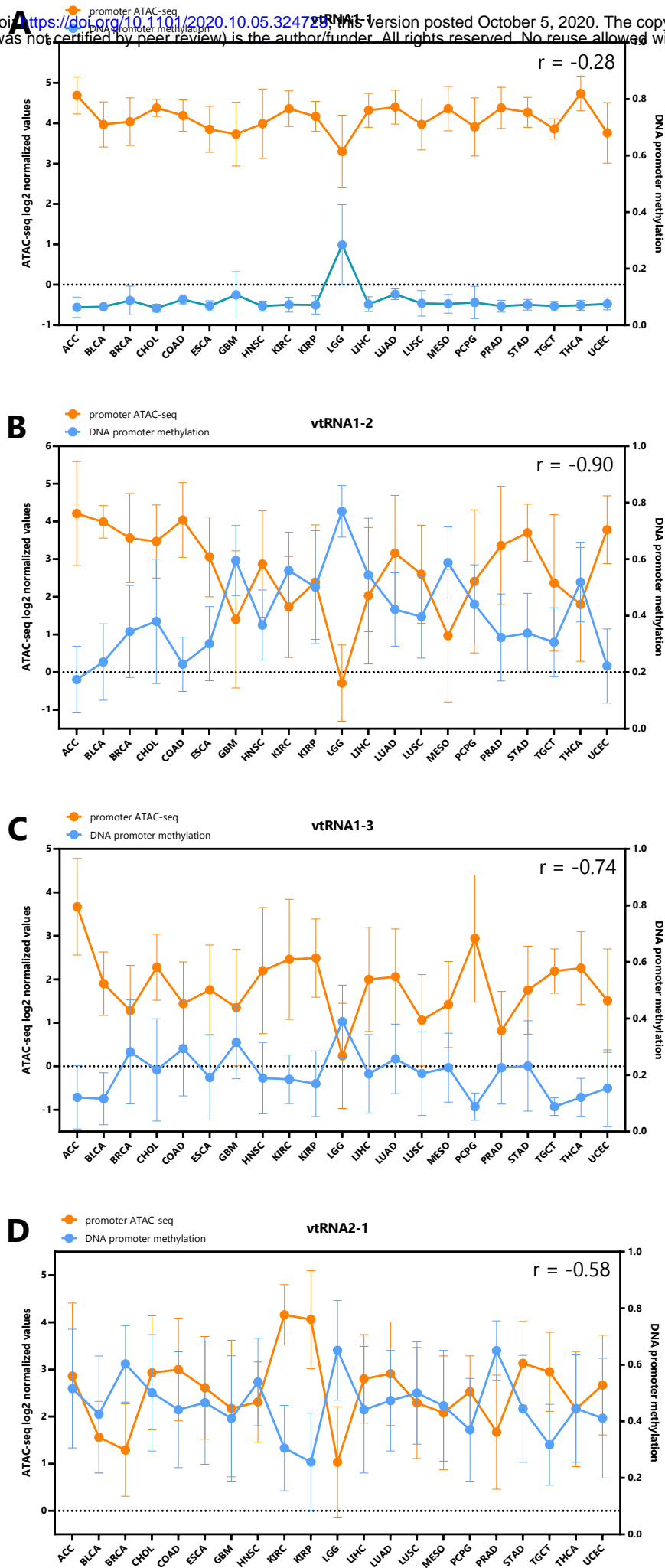


Supplementary Figure 3



Supplementary Figure 4

bioRxiv preprint doi: <https://doi.org/10.1101/2020.10.05.324728>; this version posted October 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



Supplementary Figure 5

bioRxiv preprint doi: <https://doi.org/10.1101/2020.10.05.324723>; this version posted October 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

