



An Enhancer-Based Analysis Revealed a New Function of Androgen Receptor in Tumor Cell Immune Evasion

Yuan Wang, Jiajia Li, Jingjing Li, Peipei Li, Li Wang and Lijun Di*

Cancer Center, Faculty of Health Sciences, University of Macau, Macau, China

OPEN ACCESS

Edited by:

Yi Zhao,
Chinese Academy of Sciences (CAS),
China

Reviewed by:

Zheng Xia,
Oregon Health and Science
University, United States
Zhiyun Guo,
Southwest Jiaotong University, China

*Correspondence:

Lijun Di
lijundi@um.edu.mo

Specialty section:

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

Received: 18 August 2020

Accepted: 02 November 2020

Published: 02 December 2020

Citation:

Wang Y, Li J, Li J, Li P, Wang L
and Di L (2020) An Enhancer-Based
Analysis Revealed a New Function
of Androgen Receptor in Tumor Cell
Immune Evasion.
Front. Genet. 11:595550.
doi: 10.3389/fgene.2020.595550

Cancer is characterized by dysregulation at multiple levels, such as gene transcription. Enhancers are well-studied transcription regulators that can enhance target transcripts through DNA loop formation mediated by chromosome folding. The gain or loss of the interaction between an enhancer and its target gene has a critical effect on gene expression. In this study, we analyzed GRO-seq data to identify active enhancers from seven common cancer cell lines and studied the function of these enhancers across multiple cancer types. By constructing an “enhancer effect score” (EES), we found a significant correlation between EES and tumor-infiltrating lymphocytes (TILs) in prostate cancer. Further analysis revealed that androgen receptor (AR) plays an important role in regulating the immune checkpoint gene PVR via its enhancer. These results suggest that AR contributes to prostate cancer aggressiveness by promoting cancer cell immune evasion.

Keywords: enhancer, GRO-seq, prostate cancer, androgen receptor, immune evading

INTRODUCTION

Cancer is one of the leading causes of death. Compared with normal cells, cancer cells possess some common features, such as genome instability and unlimited replicative potential, which are known as cancer hallmarks (Hanahan and Weinberg, 2000, 2011). The emergence of these cancer hallmarks is driven by mutations accumulated during the process, in which normal cells transform tissue to states from hyperplasia to dysplasia until neoplasia is established. In addition, large-scale genome-wide studies (GWASs) of cancer patients have indicated that dysregulation of gene expression significantly contributes to cancer development (Sarkar et al., 2013).

In human cells, the protein-coding genes account for only 2% of the genome, and a rough estimation predicts that no more than 40,000 genes, at maximum, are encoded. In addition, recent studies revealed that a significant proportion of the remaining 98% of genomic DNA, which is non-coding, participates in the regulation of gene expression. Some of the relatively conserved DNA sequences serve as either cis or trans regulatory elements but do not encode proteins. Enhancers are well-characterized regulatory elements and contribute mainly to positive gene expression by interacting with gene promoters and increasing the efficiency of RNA polymerase II recruitment and engagement (Koch et al., 2011). Thus, the importance of enhancers has been demonstrated in many important biological processes, such as cancer development, cell identity formation (Whyte et al., 2013), and cell fate determination (Adam et al., 2015). During cancer development, enhancers play important roles in driving oncogene transcription and may determine some cell-specific traits

(Gonen et al., 2018). Enhancers are hotspots enriched with binding motifs for transcriptional factors (TFs) that confer enhancer–gene (EG) specificity. A prevalent model depicts both gene-specific TFs and some common TFs, such as mediator complexes, as participants in the establishment of cross talk between enhancers and their target genes (Soutourina, 2019). Any variation in the enhancer locus that breaks the binding affinity for TFs results in abnormal gene expression, which can be correlated with cancer development (Sur and Taipale, 2016).

Several high-throughput next-generation sequencing (NGS)-based technologies have been developed to characterize enhancers based on common characteristics, such as ChIP-seq analysis of H3K4me1 (Rada-Iglesias et al., 2011), H3K27ac (Creyghton et al., 2010), and P300 (Visel et al., 2009; May et al., 2011); DNase-seq (Bernstein et al., 2010); FAIRE-seq (Giresi et al., 2007); and ATAC-seq (Corces et al., 2018). More recently, enhancer locus-specific transcription activity was captured and revealed the positions of active enhancers. These transcripts are termed “eRNAs.” Most eRNAs are not stable and lack a poly-A tail, which make them difficult to capture by traditional RNA-sequencing methods. Thus, a series of novel technologies, such as GRO-seq (Melgar et al., 2011), PRO-seq (Kwak et al., 2013), and NET-seq (Mayer et al., 2015), were designed to capture and characterize eRNAs and enhancers. GRO-seq is the most widely applied technology to capture eRNAs because it is based on the capture of newly synthesized RNA after transcriptional activity is restarted in an isolated nucleus, which is also known as “nuclear run-on.” Transcripts mapping the intergenic region were recognized as potential eRNAs identified in enhancer regions (Liu et al., 2017; Nagari et al., 2017; Franco et al., 2018).

In this paper, we systematically characterize the active enhancers in several cancer cells based on eRNA transcription as characterized by GRO-seq. The significant EG pairs were further characterized based on the analysis of the activity of these enhancers and the transcription activity of genes in pan-cancer samples. This large-scale analysis of coordinated enhancer and gene expression allowed us to identify an enhancer-dominated cancer hallmark, “evading immune detection,” in prostate cancer. Our data indicate that the characterization of enhancers based on eRNA is a robust way to investigate the important functions of enhancers in cancer.

MATERIALS AND METHODS

Data Collection

GRO-seq data for cell lines of A549, MCF7, HCT116, HeLa, HepG2, LNCap, and SKOV3 were collected from Gene Expression Omnibus (GEO) database. H3K4me1, H3K27ac ChIP-seq, and DNase-seq broad peak files for these cell lines were collected from both GEO and ENCODE. As the H3K4me1 ChIP-seq data are not available, SKOV3 is skipped for enhancer marker enrichment analysis. And DNase-seq for SKOV3 ovarian cancer was from ENCODE ENCSR712PYJ, a dataset from a 30-year-old female adult ovary. The detail of data accession can be found in **Supplementary Table 1**. ATAC-seq data for

relative cancer types of these cell lines were collected from the work of M. Ryan (Corces et al., 2018), which detected 23 cancers from The Cancer Genome Atlas (TCGA). The peak files were used as opening chromatin regions and compared with GRO-seq defined active transcript regions. And the bigwig files were used to calculate the ATAC signal on GRO-seq defined enhancers. CrossMap was used to transfer the hg38 to hg19, and bigWigAverageOverBed from UCSC (University of California, Santa Cruz) tools was used to count the reads on enhancers. The established enhancers were from four different data resources: (1) FANTOM5, (2) dbSUPER, (3) VISTA, and (4) Ensembl. All the data were collected in or transformed to hg19 reference with UCSC Lift Over tools. Tumor immune cell infiltration data were download from The Cancer Immunome Atlas (TCIA)¹ and Tumor Immune Estimation Resource (TIMER)² database. TCIA used normalized enrichment score (NES) to represent 28 immune cell infiltration, while TIMER was based on gene expression, which related with tumor purity to estimate B cell, CD4+ T cell, CD8+ T cell, neutrophil, macrophage, and dendritic cell infiltration. Androgen receptor (AR) ChIP-seq in prostate cancer and normal samples was downloaded from GSE56288. The processed .bw files were used to generate **Supplementary Figure 17** with Integrative Genomics Viewer (IGV). ChIA-PET data of MCF7 were download from ENCODE ENCSR000CAA dataset. Genes that belong to each cancer hallmark were download from Gene Ontology, which were reported to be related with cancer hallmarks (Plaisier et al., 2012). Data for validating AR effect on enhancer PVR regulation are from GSE117193 and GSE120720.

Data Processing

For GRO-seq, data were mapped to hg19 reference with SOAP-2.21 using the parameter `-r 0 -v 3 -n 10 -l 32`. The mapped data were used to call active transcribed regions with R package groHMM. The parameter turning step was executed to give the optimal parameters. The details of the cancer type, relative cell line, parameter for groHMM, and data performance are listed in **Supplementary Table 1** and **Figure 1**. For ChIP-seq, data were mapped to hg19 with bowtie2 with a default parameter, and the mapped reads with quality no less than 30 were filtered.

Enhancer Definition

The pipeline groHMM was performed to call active transcripts for each cell type, average RPKM ≥ 1 was used as cutoff for cell types with replicates, and RPKM ≥ 2 was used for those without replicates to filter out active transcripts with low confidence. In order to avoid the interference from coding genes, the transcripts were filtered with bedtools suite intersect `-v` function to include the ones 5 kb away from any refseq genes 5'/3'-UTR regions. The refseq genes were downloaded from UCSC table with reference hg19. And then, the filtered transcripts were combined to obtain new transcript boundary across all cell types by merging the transcripts within 500 bp (Danko et al., 2015), and the merged transcripts were taken

¹<https://tcia.at/home>

²<https://cistrome.shinyapps.io/timer/>

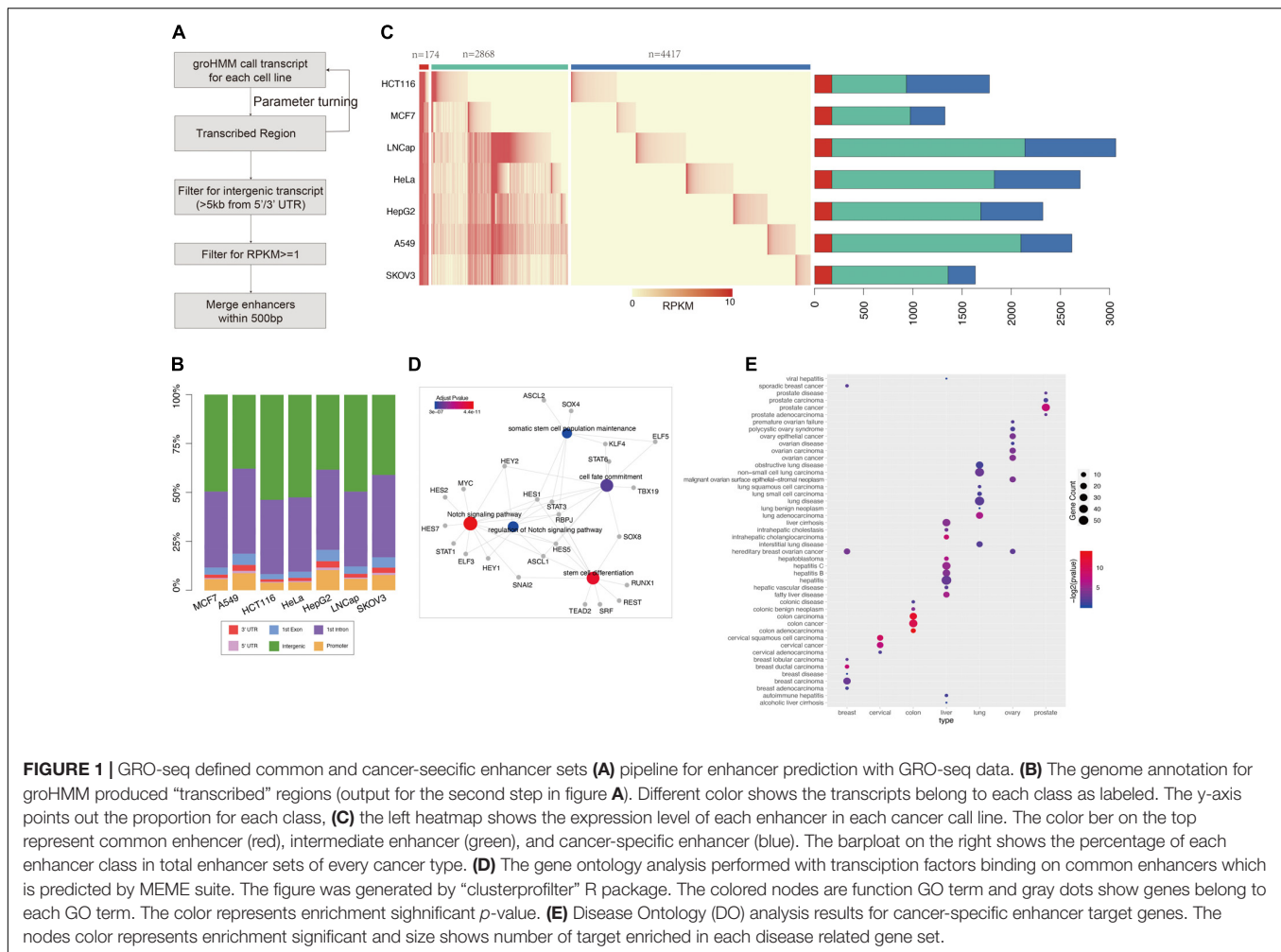


FIGURE 1 | GRO-seq defined common and cancer-specific enhancer sets **(A)** pipeline for enhancer prediction with GRO-seq data. **(B)** The genome annotation for groHMM produced “transcribed” regions (output for the second step in figure **A**). Different color shows the transcripts belong to each class as labeled. The y-axis points out the proportion for each class, **(C)** the left heatmap shows the expression level of each enhancer in each cancer call line. The color bar on the top represent common enhancer (red), intermediate enhancer (green), and cancer-specific enhancer (blue). The barplot on the right shows the percentage of each enhancer class in total enhancer sets of every cancer type. **(D)** The gene ontology analysis performed with transcription factors binding on common enhancers which is predicted by MEME suite. The figure was generated by “clusterprofiler” R package. The colored nodes are function GO term and gray dots show genes belong to each GO term. The color represents enrichment significant p -value. **(E)** Disease Ontology (DO) analysis results for cancer-specific enhancer target genes. The nodes color represents enrichment significant and size shows number of target enriched in each disease related gene set.

as putative enhancers for the following analysis. Functional enhancers were defined as enhancers that show significant Pearson correlation with target genes in any cancer types (adjusted p -value < 0.05).

Enhancer Target Assignment

The genes within 500 kb were considered as enhancer targets at first. In case of missing long-distance pairs, enhancer-target pairs defined by GeneHancer database were also included. Then Pearson correlation was calculated with TCGA data to confirm the regulation relationship between each pair. Three types of data were used to estimate enhancer activity in each TCGA sample: (1) for TCGA ATAC-seq data, the enhancer activity was measured by average ATAC-seq signal on enhancer regions. (2) For TCGA RNA-seq data, as Han Liang et al. have published the enhancer expression profile for TCGA samples (Chen et al., 2018), we simply take their established expression value as our enhancer activity if their defined enhancer is situated within 500 bp of our enhancers. (3) For TCGA methylation data, the average methylation beta value from methylation 450 K array was used to represent our enhancer activity. The Pearson correlation based on these three datasets were calculated, and false discovery

rate (FDR) < 0.05 accompanied with no less than 0.3 absolute correlation value was used as cutoff to filter confident results.

Enhancer Effect Score

The correlation of EG pairs reflects the regulation ability of enhancer to target genes. To quantify the enhancer regulation of a set of genes, such as a set of genes that belong to the same pathway, the average of correlations weighted by target gene expression was used. As limited by dataset, not all the genes were included in our EG pairs. Therefore, hypergeometric distribution test was performed to calculate the enrichment of target genes in specific gene set, and the p -value was used to adjust the average correlation. Finally, this score presents the enhancer regulation to specific pathways and we named it as enhancer effect score (EES), with the formula as follows:

$$EES_s = (-\log_2 P\text{-value}_{\text{enrichment}}) * \frac{\sum_i^n RPKM_i * |Cor_{\max}|_i}{\sum_i^n RPKM_i}$$

$P\text{-value}_{\text{enrichment}}$ is calculated by hypergeometric distribution; n is the total number of genes that belong to a specific gene set; $RPKM_i$ is the gene i expression in sample s ; and Cor_{\max} is the Pearson correlation coefficient between gene i and its enhancer,

when gene *i* is correlated with more than one enhancer, and the max value was used.

Motif Enrichment Analysis

MEME-ChIP was used to detect the binding TF enriched on 174 common active enhancers. We extend 50 bp on both sides of each enhancers and sliced enhancers into 100-bp-long pieces to make sure that every part of the enhancer is under motif searching area. TOMTOM was selected to match the motif to known TFs. Gene Ontology analysis was performed on these enriched TFs with R package “clusterProfiler.”

Modulator Searching

The samples were sorted according to certain gene expression. The first and last quarters were selected as low and high expressed groups, respectively. The correlation was recalculated within these two groups; the difference between two correlation value (Diffcor) was used to filter modulators. The genes with Diffcor of no less than 0.5 were considered related with enhancer regulation to PVR expression.

Statistical Analysis

Differentially expressed genes were defined as fold change larger than 1.5 and *t*-test FDR < 0.05. For those cancer types that have no normal data or less than five normal samples, the average expression from all TCGA normal samples was used to compare with that of tumors. The H3K27ac and H3K4me1 enrichment on enhancer regions was performed with deepTools software. The overlap of opening chromatin regions defined by GRO-seq, ATAC-seq, and DNase-seq was done by “findOverlapsOfPeaks” functions in “ChIPpeakAnno” package. The conservation score was measured with phastCons score obtained by “phastCons100way.UCSC.hg19” package. And the average phastCons score was used to calculate enhancers and their 2.5-kb flanking region with 50 bp as slicing window. And for the genome background, we randomly select the same number of regions with functional enhancers, also with the same length as average of these enhancers, and we use the same method to calculate the phastCons distribution around these random regions.

RESULTS

GRO-Seq Revealed Common and Specific Enhancers Across Cell Lines

To study data on active enhancer behavior in cancer, GRO-seq data for seven common cancers (LUAD, BRCA, CESC, COAD, LIHC, PRAD, and OV) were collected from the GEO database. The GroHMM method was performed on these GRO-seq data (Figure 1A), and the active transcribed regions were identified. These regions were considered open chromatin regions, and potential enhancers were further extracted from them. On average, 86% (ranging at 79% for the HepG2 samples and 91.7% for the HCT116 samples) of these active transcribed regions were located in the non-coding region (Figure 1B).

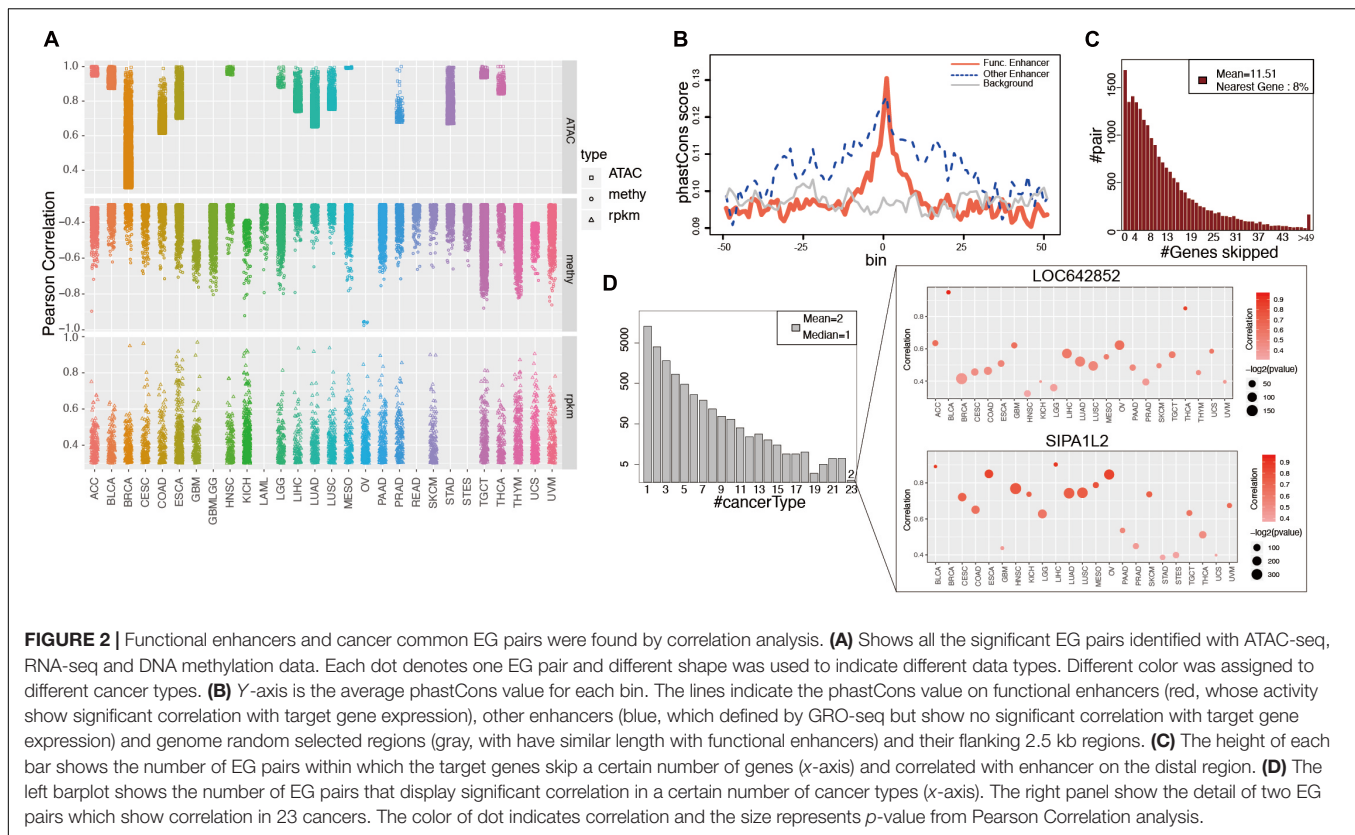
All of these transcribed regions were compared with results from analyses of DNase-seq and ATAC-seq, which are two widely used methods for detecting opening chromatin regions (Roadmap Epigenomics Consortium, Kundaje et al., 2015; Corces et al., 2018). GRO-seq can cover more than 87% peaks obtained from ATAC-seq, and approximately 92% peaks were obtained from DNase-seq. In addition, GRO-seq defined 435,838 active transcribed regions that were covered by either of the two other methods (Supplementary Figure 2). These results show that GRO-seq is more sensitive in predicting regions with an open chromatin status.

Following the pipeline presented in Figure 1A, GRO-seq-defined active transcribed regions situated outside 5 kb of the coding gene were obtained by filtering. The number of GRO-seq-defined “transcribed” regions is related to sequencing depth (Pearson correlation: 0.83) (Figure 1 and Supplementary Figure 3); for example, HeLa cells contain three-fold more active transcripts than HepG2 cells. Thus, to obtain more reliable enhancer sets, a cutoff of RPKM ≥ 1 was used (see section “Materials and Methods”) to filter the enhancer sets. Finally, 7,459 putative enhancers were identified. More than one-half of these enhancers (4,417/7,459) show activity in only one cancer type, and we named them “cancer-specific enhancers.” However, 174 enhancers showed consistent eRNA expression across all the cancer types assessed, and we defined these enhancers as “common enhancers” (Figure 1C). These 174 common enhancers were bound by TFs related to common cancer features, such as “Notch signaling pathway” and “stem cell differentiation” (Figure 1D). “Notch signaling” is pivotal for cell fate determination, and its dysregulation was reported to be involved in oncogenesis (Allenspach et al., 2002). Additionally, cancer stem cells are known to contribute to cancer initiation and progression (Yu et al., 2012; Venkatesh et al., 2018). The correlation analysis of the common enhancers and the expression of their downstream target genes support the hypotheses that the common enhancers contribute to the activation of these pathways and tumor progression. The cell type-specific enhancers were also enriched with corresponding cancer genes according to a disease ontology analysis (Figure 1E).

GRO-seq-defined enhancers showed considerable enrichment with H3K27ac and H3K4me1 (Supplementary Figure 4). Another four enhancer resources (Ensembl, dbSUPER, FANTOM5, and VISTA) were also accessed to further characterize the GRO-seq-defined enhancer sets, and approximately 41% of the GRO-seq-defined enhancers were identified in at least one of these databases (Supplementary Figure 5). A comparison of enhancer markers among these enhancer sets was also performed, and GRO-seq-defined enhancers were similarly enriched with these markers, suggesting that GRO-seq is a reliable approach to explore the function of enhancers in cancer (Supplementary Figure 6).

Enhancer-Based Analysis Revealed Cancer Risk Regulatory Events

To identify the regulatory targets of the identified enhancers, enhancers and genes within 500 kb were used to comprise EG



pairs (Javierre et al., 2016; Corces et al., 2018), and together with the EG pairs from the GeneHancer database, more than 84,000 pairs were collected. Although the enhancers were mainly retrieved from the GRO-seq data based on seven types of cancer cells representative of seven major types of cancer, a significant fraction of these enhancers are active enhancers in more than two types of cancer (from 52.5% in the HCT116 cells to 83% in the SKOV3 cells), suggesting that different types of cancer share enhancers to a certain extent. Hence, in addition to the seven common cancers, more cancer types were considered in the following study of the regulatory relationships between enhancers and genes. To confirm the regulatory relationship, the Pearson correlation between enhancer activity and gene expression was calculated for each established EG pair, within which the enhancer activity was determined by ATAC-seq, DNA methylation, or RNA-seq according to TCGA data (see section Materials and Methods). In total, 20,930 EG pairs comprising 4,815 enhancers and 10,132 genes showed a significant correlation between enhancer activity and gene expression (Figure 2A). The correlation results from the ATAC-seq, DNA methylation, and RNA-seq data were combined, and 75% of the EG pairs were supported by the ATAC-seq (39 and 7% of the findings were supported by the methylation and RNA-seq data, respectively), and 386 EG pairs were supported by all three types of data (Supplementary Figure 7A).

The enhancers that showed a significant correlation with target gene expression were considered “functional enhancers.” Consistent with established knowledge, functional enhancers are

more conserved than flanking regions and random genomic backgrounds (Figure 2B). Consistent with a previous report (Javierre et al., 2016), when only EG pairs from GeneHancer were taken into account, 96% of the targets were within 500 kb of enhancers (Supplementary Figure 7B). More than one-half of target gene expression is regulated by at least two enhancers, whereas on average, one enhancer regulates four different targets (Supplementary Figures 7C,D). Intermediate DNA loop generation is a known mechanism that enables the direct interaction between an enhancer and a promoter (Su et al., 1991). ChIA-PET has been applied to study the long-distance interaction between an enhancer and a promoter mediated by TF or Pol II. We analyzed the Pol II ChIA-PET data for MCF-7 cells and observed that the EG pairs with DNA loops showed a higher correlation than those without loops (Supplementary Figure 7E). Unexpectedly, among all the target genes we analyzed, only 8% of the enhancer target genes were located in neighboring regions (Figure 2C). Previous reports showing that the nearest genes are more likely to be enhancer targets (Heintzman et al., 2009; Visel et al., 2009; Xu et al., 2012); however, our findings suggest that enhancers tend to skip the nearest genes, which is consistent with the result by Sanyal et al. (2012).

In our study, among all 20,930 EG pairs, 70% showed a significant correlation in at least two cancer types. Two EG pairs containing the target genes SIPA1L2 (signal-induced proliferation-associated 1 like 2) and LOC642852 showed a significant correlation in 23 cancers (Figure 2D). SIPA1L2

is related to GTPase activator activity and the Ras signaling pathway. The dysregulation of SIPA1L2 is also known as a risk factor for cancer. Although a non-coding gene, LOC642852 is widely expressed in different cancers, and some studies have reported that LOC642852 is related to colorectal cancer and pancreatic cancer (Giulietti et al., 2018; Zhang et al., 2019).

Enhancers Contribute to Cancer Signaling Pathways

To further characterize enhancer functions in cancer, data on 174 genes in 10 common oncogenic signaling pathways were collected (Sanchez-Vega et al., 2018). More than one-half of the genes in each pathway were significantly correlated with enhancer activity, which indicates the important function of enhancers in regulating the activity of these pathways (**Figure 3** and **Supplementary Figure 8**). For example, MYC is a key oncogene in the MYC signaling pathway, and the dysregulation of MYC is related to the development of many cancers (Boxer and Dang, 2001; Eilers and Eisenman, 2008; Gabay et al., 2014). Three different enhancers of MYC were identified (**Figure 4**). MYC is highly expressed in many cancer types, but the abnormal activation of MYC is related to different enhancers (Lancho and Herranz, 2018). In COAD and STAD, MYC expression is related to the enhancer located approximately 160 kb upstream of the MYC promoter (E1), while for HNSC and BLCA, the expression of MYC is correlated with the enhancer located near the promoter downstream (E2). Interestingly, MYC expression is significantly correlated with both of these enhancers in STAD. To study how these two enhancers contribute to MYC expression in STAD, both one-way and multivariate analyses of variance were performed, and the results showed that E1 contributes to MYC more significantly (**Supplementary Figure 9**). Surprisingly, low expression of both E1 and E2 results in the recovery of MYC expression in STAD, suggesting that other unknown regulatory mechanisms might be involved. The third enhancer (E3) of MYC is located far downstream from the gene, and some other genes are in the intermediate region. However, the activity of this enhancer only shows a strong correlation with MYC expression in LUAD, BRCA, CESC, and ESCA. Interestingly, the single-nucleotide polymorphism (SNP) “rs11780156,” which has been reported as a risk SNP in GWAS in breast cancer was found in the MYC enhancer E3 (Shi et al., 2016). Moreover, several other SNPs (rs6999335, rs6999897, rs11783807, rs56152647, rs6992491, and rs67397162), in linkage disequilibrium with SNP “rs11780156,” are all located in this enhancer according to HaploReg (Ward and Kellis, 2012), suggesting that E3 plays an important role in increasing the risk of BRCA.

Enhancers Regulate Cancer Hallmarks

We also examined the enrichment of enhancer target genes in cancer hallmarks (Hanahan and Weinberg, 2011). As expected, the target genes were significantly enriched in these 10 hallmarks (**Figure 5A**). For example, the hallmark of tissue invasion and metastasis is enriched in nearly all tumor types, suggesting that the genes that promote tumor invasion and metastasis are most likely the regulatory targets of active enhancers. Some other

hallmarks are enriched only in a few types of cancer, suggesting that while some common traits are shared by most types of cancer, the heterogeneity of cancer hallmarks can be used to distinguish different types of cancer. Although the enrichment patterns are different among cancer types, approximately 50% of genes belonging to each cancer hallmark were found to be regulated by enhancers (**Figure 5B**), suggesting that enhancers play important roles in characterizing cancer traits.

Tumor-Infiltrating Lymphocytes Are Regulated by Enhancers

A significant number of genes identified in our EG list were involved in one of the important cancer hallmarks, the “evading immune detection” (EID) pathway in PRAD, COAD, and GBMLGG cancers, among which PRAD shows the most significant enrichment (**Figure 5A**). These three cancer types generally have higher tumor purity than the other types (**Supplementary Figure 10**). To evaluate the effect of an enhancer on the tumor immune response, an EES was developed. EES is a weighted average Pearson correlation calculated for all EG pairs within an EID pathway. We found that, in PRAD, the EES was significantly correlated with tissue infiltration of many immune cells, such as natural killer cells, activated CD8+ T cells, dendritic cells, T helper cells, and immature B cells (**Figure 6A** and **Supplementary Figure 11**). However, the activity of the EID pathway, which was measured by the average gene expression, was not related to tumor-infiltrating lymphocytes (TILs) (**Supplementary Figures 12A,B**). Since each enhancer may regulate more than one gene, depending on the local chromatin organization, this result suggests that enhancers can be better representatives of local genomic activity than any single gene.

By examining every EG pair within the EID pathway, we found that PVR and its enhancer showed the most significant correlation (**Supplementary Figure 13A**). PVR, also known as CD155, is an immune checkpoint that can suppress the immune response through its interaction with the T-cell receptor TIGIT (Inozume et al., 2015). PVR is highly expressed in many cancers, including prostate cancer (**Supplementary Figure 13B**), and its high expression is associated with poor clinical outcomes for prostate cancer patients (**Supplementary Figure 13C**). The enhancer of PVR identified in this study was situated approximately 40 kb upstream of the promoter, and the correlation between the enhancer activity and gene expression was 0.78; PVR is actively transcribed in many cancers (**Figure 6B**), and the putative enhancer of PVR was found to be positively correlated with PVR gene expression across all examined cancer types (**Supplementary Figure 14**). The activation of this enhancer is accompanied by PVR expressed in many cancers, such as ESCA, HNSC, and LUSC (**Figure 6B**). In other cancer types, such as KIRP, MESO, and CHOL, PVR expression also seems to be related to a putative intergenic enhancer in an intron of IGSF23.

The regulation of the target genes by enhancers is influenced not only by enhancer activity but also by factors mediating the interactions between the enhancer and the promoter. Since

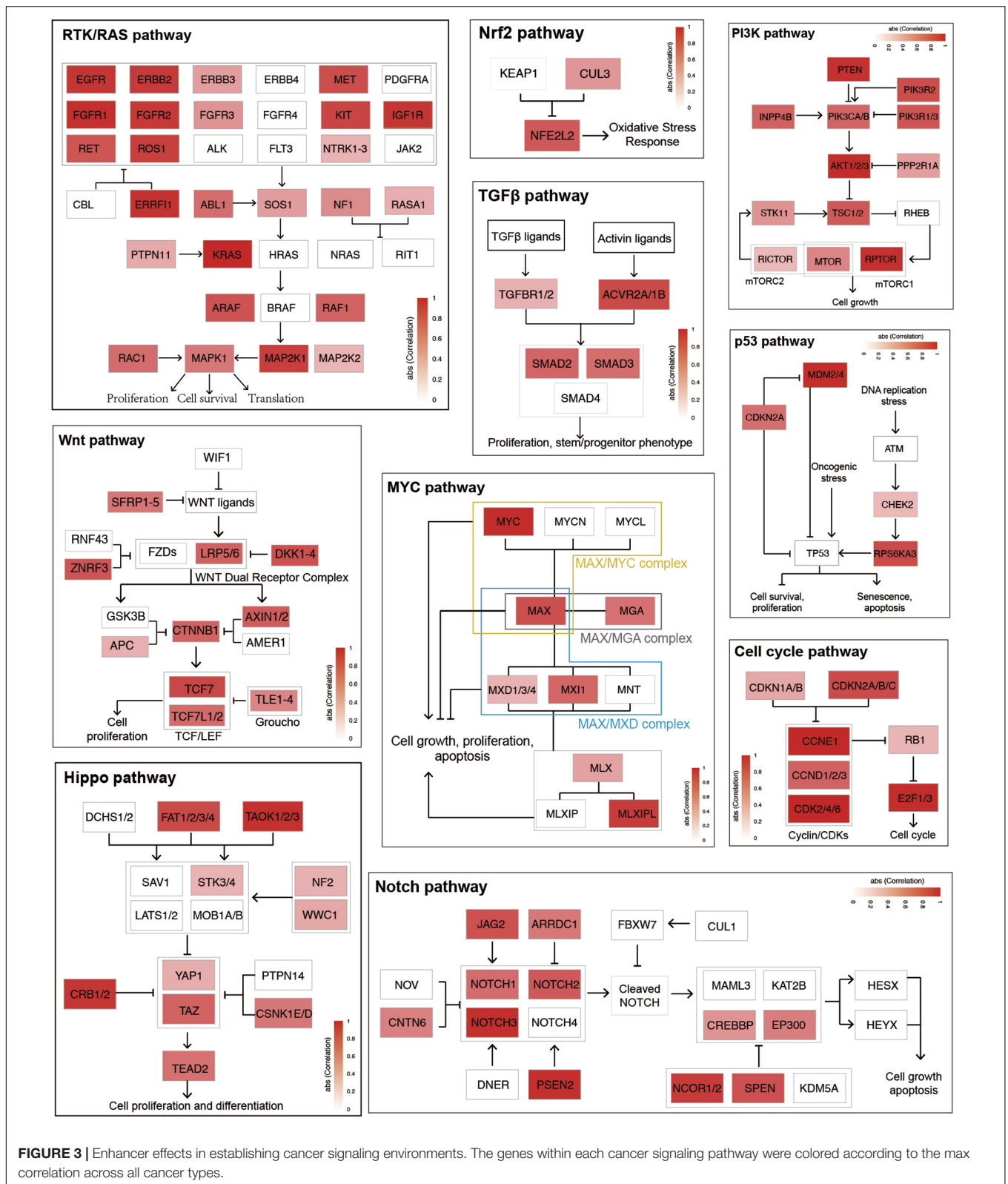
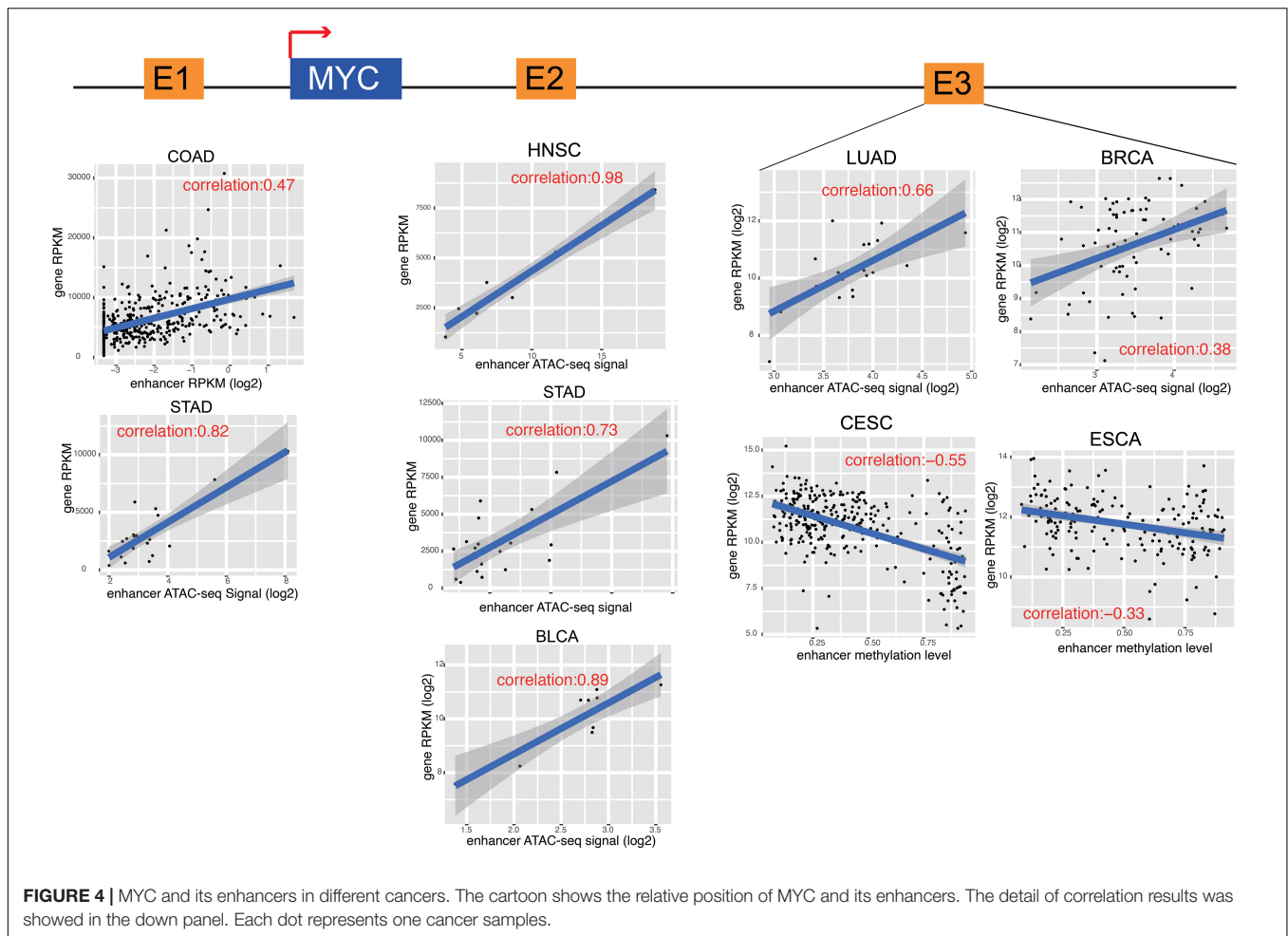


FIGURE 3 | Enhancer effects in establishing cancer signaling environments. The genes within each cancer signaling pathway were colored according to the max correlation across all cancer types.

MED1 is an essential factor for enhancer regulation of the target gene (Sabari et al., 2018; Zamudio et al., 2019), the influence of MED1 on the correlation between the enhancer and

PVR expression has been validated (Supplementary Figure 15). To globally explore the factors that potentially mediate the interaction between the PVR enhancer and promoter, a search for



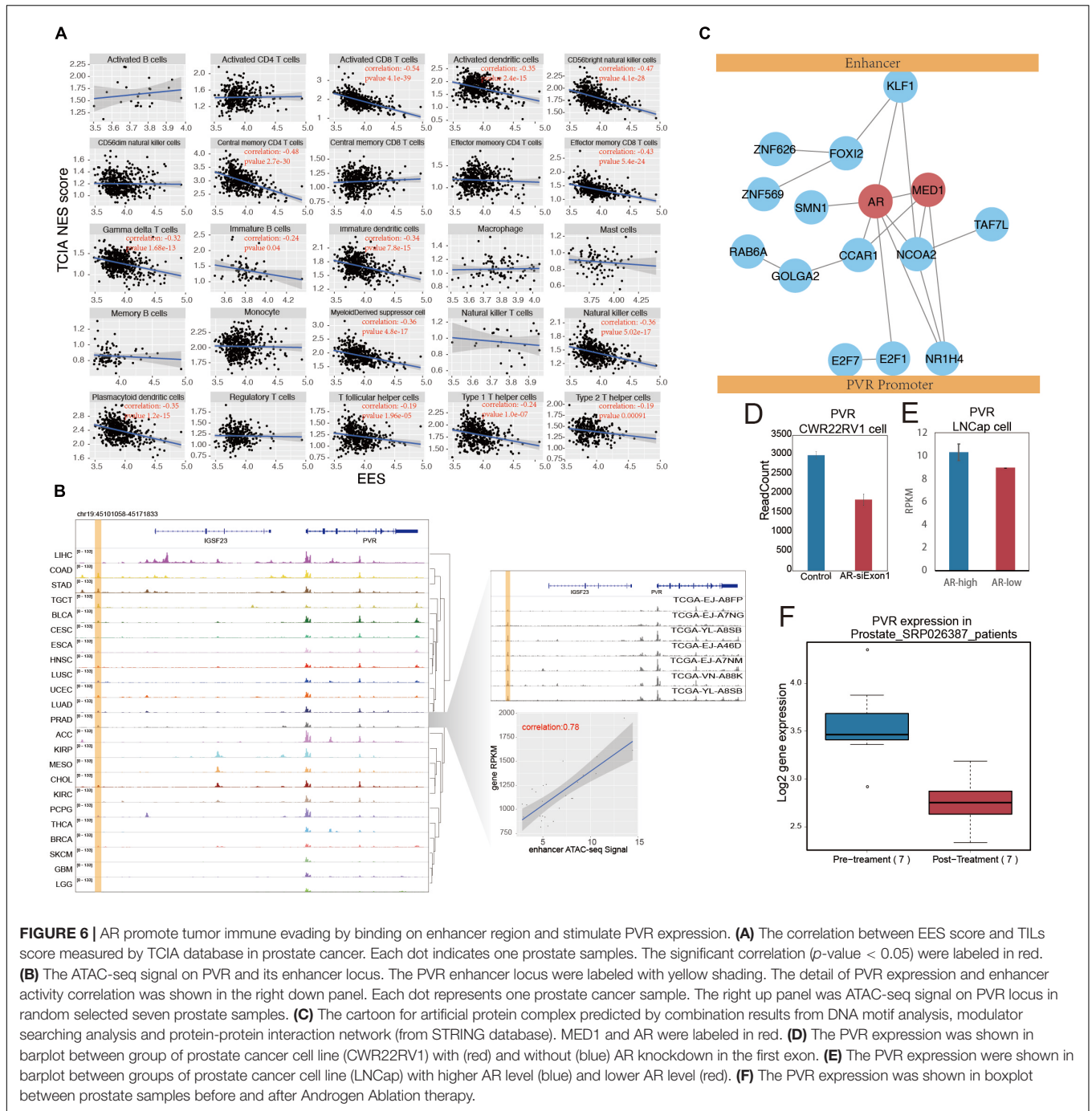
modulators was performed. Highly expressed genes contributing to the high correlation of the PVR enhancer and promoter were considered potential modulators, and 1,215 PVR modulators were identified by this method (**Supplementary Figure 16**). The top 20 modulators were selected for further characterization, and among these candidate, proteins directly binding to the PVR enhancer and promoter were identified by motif screening. The 20 modulators found in this manner were further filtered according to the STRING database. The proteins showing no connection to other modulators were removed, and the remaining modulators were used to form a protein–protein interaction network to predict the protein complex(es) mediating the interaction between the PVR enhancer and its promoter (**Figure 6C**). Within this predicted protein complex, in addition to MED1, AR was found to be an important hub node.

In support of our proposition, a reanalysis of AR binding with an AR ChIP-seq dataset containing both normal and tumor samples of prostate cancer revealed that AR binds only to the PVR enhancer region in the prostate tumor samples but not in the normal prostate tissue (Pomerantz et al., 2015). This result suggests that highly expressed PVR in tumors may correlate with the binding of AR to enhancers (**Supplementary Figure 17**). Moreover, the knockdown of AR in CWR22Rv1 prostate cancer

cell lines led to decreased PVR expression (Hepburn et al., 2019; **Figure 6D**). PVR expression is also significantly higher in prostate cancer cells with high AR levels (Lee et al., 2019; **Figure 6E**). These results suggest that AR mediates PVR expression in prostate cancer. AR binds to androgen and then enters the nucleus to regulate gene transcription. The function of AR is affected when androgen levels are modulated. We generated a clinical dataset (Rajan et al., 2014) based on prostate cancer patients who had accepted androgen ablation therapy, and we found a significant decrease in PVR expression in patients after treatment (**Figure 6F**). Together, these results indicate that AR plays potential roles in regulating PVR expression by mediating its promoter interaction with a distal enhancer.

DISCUSSION

Enhancers are critical regulatory elements of gene expression, and the characterization of the dysregulation of enhancers is important in cancer research (Kron et al., 2014; Yuan et al., 2017). In this study, we analyzed GRO-seq data of cancer cell lines to globally identify enhancers based on eRNA expression. The importance of these enhancers was further investigated in



Nishiwada et al., 2015). Our finding of a PVR enhancer is the first report indicating the potential role of PVR in prostate cancer. To our surprise, we were only able to observe that the EES of the EID pathway correlated with the outcome of patients. The correlation of the average expression of an EID pathway gene and the outcome for patients was not significant. However, we found that the PVR enhancer is in a very gene-rich region, suggesting that the PVR enhancer might also play roles in regulating the expression of other genes in addition to its dominant role as a PVR gene enhancer. In fact, regulating the expression of several

genes by a single enhancer is not unusual. Thus, the EES of EID genes might be a better predictor of patient outcome than average gene expression.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

The authors are grateful for all the support from lab members. LD helped to improve the project and manuscripts. JiaL, PL, and JinL provided technical support for analysis and manuscript assessment. All authors contributed to the article and approved the submitted version.

FUNDING

This work is supported by the Science and Technology Development Fund (FDCT) of Macao SAR to LD

REFERENCES

- Adam, R. C., Yang, H., Rockowitz, S., Larsen, S. B., Nikolova, M., Oristian, D. S., et al. (2015). Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. *Nature* 521, 366–370. doi: 10.1038/nature14289
- Allenspach, E. J., Maillard, I., Aster, J. C., and Pear, W. S. (2002). Notch signaling in cancer. *Cancer Biol. Ther.* 1, 466–476.
- Bernstein, B. E., Stamatoyannopoulos, J. A., Costello, J. F., Ren, B., Milosavljevic, A., Meissner, A., et al. (2010). The NIH roadmap epigenomics mapping consortium. *Nat. Biotechnol.* 28, 1045–1048. doi: 10.1038/nbt1010-1045
- Bevelacqua, V., Bevelacqua, Y., Candido, S., Skarmoutsou, E., Amoroso, A., Guarneri, C., et al. (2012). Nectin like-5 overexpression correlates with the malignant phenotype in cutaneous melanoma. *Oncotarget* 3, 882–892. doi: 10.18632/oncotarget.594
- Bowers, J. R., Readler, J. M., Sharma, P., and Excoffon, K. (2017). poliovirus receptor: more than a simple viral receptor. *Virus Res.* 242, 1–6. doi: 10.1016/j.virusres.2017.09.001
- Boxer, L. M., and Dang, C. V. (2001). Translocations involving c-myc and c-myc function. *Oncogene* 20, 5595–5610. doi: 10.1038/sj.onc.1204595
- Chen, H., Li, C., Peng, X., Zhou, Z., Weinstein, J. N., Cancer Genome Atlas Research Network, et al. (2018). A pan-cancer analysis of enhancer expression in nearly 9000 patient samples. *Cell* 173, 386.e312–399.e312.
- Corces, M. R., Granja, J. M., Shams, S., Louie, B. H., Seoane, J. A., Zhou, W., et al. (2018). The chromatin accessibility landscape of primary human cancers. *Science* 362:eaav1898.
- Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21931–21936. doi: 10.1073/pnas.1016071107
- Danko, C. G., Hyland, S. L., Core, L. J., Martins, A. L., Waters, C. T., Lee, H. W., et al. (2015). Identification of active transcriptional regulatory elements from GRO-seq data. *Nat. Methods* 12, 433–438. doi: 10.1038/nmeth.3329
- Eilers, M., and Eisenman, R. N. (2008). Myc's broad reach. *Genes Dev.* 22, 2755–2766. doi: 10.1101/gad.1712408
- Franco, H. L., Nagari, A., Malladi, V. S., Li, W., Xi, Y., Richardson, D., et al. (2018). Enhancer transcription reveals subtype-specific gene expression programs controlling breast cancer pathogenesis. *Genome Res.* 28, 159–170. doi: 10.1101/gr.226019.117
- Gabay, M., Li, Y., and Felsher, D. W. (2014). MYC activation is a hallmark of cancer initiation and maintenance. *Cold Spring Harb. Perspect. Med.* 4:a014241. doi: 10.1101/cshperspect.a014241
- Giresi, P. G., Kim, J., Mcdaniell, R. M., Iyer, V. R., and Lieb, J. D. (2007). FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res.* 17, 877–885. doi: 10.1101/gr.5533506
- Giulietti, M., Righetti, A., Principato, G., and Piva, F. (2018). LncRNA co-expression network analysis reveals novel biomarkers for pancreatic cancer. *Carcinogenesis* 39, 1016–1025. doi: 10.1093/carcin/bgy069
- (FDCT/0014/2018/A1 and FDCT/0117/2018/A3), the Multi-Year Research Grant from the University of Macau to LD (MYRG2018-00158-FHS), and the National Natural Science Foundation of China (NSFC 81772980) to LD. This work is also supported by the Multi-Year Research Grant from the University of Macau to LW (MYRG2016-00251-FHS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.595550/full#supplementary-material>

- Liu, Y., Chen, S., Wang, S., Soares, F., Fischer, M., Meng, F., et al. (2017). Transcriptional landscape of the human cell cycle. *Proc. Natl. Acad. Sci. U.S.A.* 114, 3473–3478.
- May, D., Blow, M. J., Kaplan, T., McCulley, D. J., Jensen, B. C., Akiyama, J. A., et al. (2011). Large-scale discovery of enhancers from human heart tissue. *Nat. Genet.* 44, 89–93. doi: 10.1038/ng.1006
- Mayer, A., Di Iulio, J., Maleri, S., Eser, U., Vierstra, J., Reynolds, A., et al. (2015). Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell* 161, 541–554. doi: 10.1016/j.cell.2015.03.010
- Melgar, M. F., Collins, F. S., and Sethupathy, P. (2011). Discovery of active enhancers through bidirectional expression of short transcripts. *Genome Biol.* 12:R113.
- Nagari, A., Murakami, S., Malladi, V. S., and Kraus, W. L. (2017). Computational Approaches for Mining GRO-Seq Data to Identify and Characterize Active Enhancers. *Methods Mol. Biol.* 1468, 121–138. doi: 10.1007/978-1-4939-4035-6_10
- Nakai, R., Maniwa, Y., Tanaka, Y., Nishio, W., Yoshimura, M., Okita, Y., et al. (2010). Overexpression of Necl-5 correlates with unfavorable prognosis in patients with lung adenocarcinoma. *Cancer Sci.* 101, 1326–1330. doi: 10.1111/j.1349-7006.2010.01530.x
- Nishiwada, S., Sho, M., Yasuda, S., Shimada, K., Yamato, I., Akahori, T., et al. (2015). Clinical significance of CD155 expression in human pancreatic cancer. *Anticancer. Res.* 35, 2287–2297.
- Plaisier, C. L., Pan, M., and Baliga, N. S. (2012). A miRNA-regulatory network explains how dysregulated miRNAs perturb oncogenic processes across diverse cancers. *Genome Res.* 22, 2302–2314. doi: 10.1101/gr.133991.111
- Pomerantz, M. M., Li, F., Takeda, D. Y., Lenci, R., Chonkar, A., Chabot, M., et al. (2015). The androgen receptor citrome is extensively reprogrammed in human prostate tumorigenesis. *Nat. Genet.* 47, 1346–1351. doi: 10.1038/ng.3419
- Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. A., Flynn, R. A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470, 279–283. doi: 10.1038/nature09692
- Rajan, P., Sudbery, I. M., Villasevil, M. E., Mui, E., Fleming, J., Davis, M., et al. (2014). Next-generation sequencing of advanced prostate cancer treated with androgen-deprivation therapy. *Eur. Urol.* 66, 32–39. doi: 10.1016/j.eururo.2013.08.011
- Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317–330.
- Sabari, B. R., Dall'agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 361:eaar3958. doi: 10.1126/science.aar3958
- Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W. K., Luna, A., La, K. C., et al. (2018). Oncogenic signaling pathways in the cancer genome atlas. *Cell* 173, 321.e310–337.e310.
- Sanyal, A., Lajoie, B. R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature* 489, 109–113. doi: 10.1038/nature11279
- Sarkar, S., Horn, G., Moulton, K., Oza, A., Byler, S., Kokolus, S., et al. (2013). Cancer development, progression, and therapy: an epigenetic overview. *Int. J. Mol. Sci.* 14, 21087–21113. doi: 10.3390/ijms141021087
- Shi, J., Zhang, Y., Zheng, W., Michailidou, K., Ghoussaini, M., Bolla, M. K., et al. (2016). Fine-scale mapping of 8q24 locus identifies multiple independent risk variants for breast cancer. *Int. J. Cancer* 139, 1303–1317.
- Soutourina, J. (2019). Mammalian mediator as a functional link between enhancers and promoters. *Cell* 178, 1036–1038. doi: 10.1016/j.cell.2019.07.040
- Su, W., Jackson, S., Tjian, R., and Echols, H. (1991). DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. *Genes Dev.* 5, 820–826. doi: 10.1101/gad.5.5.820
- Sur, I., and Taipale, J. (2016). The role of enhancers in cancer. *Nat. Rev. Cancer* 16, 483–493. doi: 10.1038/nrc.2016.62
- Venkatesh, V., Nataraj, R., Thangaraj, G. S., Karthikeyan, M., Gnanasekaran, A., Kagineeli, S. B., et al. (2018). Targeting Notch signalling pathway of cancer stem cells. *Stem Cell Investig.* 5:5. doi: 10.21037/sci.2018.02.02
- Visel, A., Blow, M. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854–858. doi: 10.1038/nature07730
- Ward, L. D., and Kellis, M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 40, D930–D934.
- Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., et al. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307–319. doi: 10.1016/j.cell.2013.03.035
- Xu, J., Shao, Z., Glass, K., Bauer, D. E., Pinello, L., Van Handel, B., et al. (2012). Combinatorial assembly of developmental stage-specific enhancers controls gene expression programs during human erythropoiesis. *Dev. Cell* 23, 796–811. doi: 10.1016/j.devcel.2012.09.003
- Yu, Z., Pestell, T. G., Lisanti, M. P., and Pestell, R. G. (2012). Cancer stem cells. *Int. J. Biochem. Cell Biol.* 44, 2144–2151.
- Yuan, J., Jiang, Y. Y., Mayakonda, A., Huang, M., Ding, L. W., Lin, H., et al. (2017). Super-Enhancers Promote Transcriptional Dysregulation in Nasopharyngeal Carcinoma. *Cancer Res.* 77, 6614–6626. doi: 10.1158/0008-5472.can-17-1143
- Zamudio, A. V., Dall'agnese, A., Henninger, J. E., Manteiga, J. C., Afeyan, L. K., Hannett, N. M., et al. (2019). Mediator condensates localize signaling factors to key cell identity genes. *Mol. Cell.* 76, 753.e6–766.e6.
- Zhang, L., Chen, S., Wang, B., Su, Y., Li, S., Liu, G., et al. (2019). An eight-long noncoding RNA expression signature for colorectal cancer patients' prognosis. *J. Cell. Biochem.* 120, 5636–5643. doi: 10.1002/jcb.27847

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Wang, Li, Li, Wang and Di. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.