



Article Regulation Network of Colorectal-Cancer-Specific Enhancers in the Progression of Colorectal Cancer

Bohan Chen, Yiping Ma, Jinfang Bi, Wenbin Wang, Anshun He, Guangsong Su, Zhongfang Zhao, Jiandang Shi and Lei Zhang *

State Key Laboratory of Medicinal Chemical Biology and College of Life Sciences, Nankai University, 94 Weijin Road, Tianjin 300071, China; 1120200511@mail.nankai.edu.cn (B.C.); 1120200510@mail.nankai.edu.cn (Y.M.); 1120180423@mail.nankai.edu.cn (J.B.); 2120181086@mail.nankai.edu.cn (W.W.); 1120160377@mail.nankai.edu.cn (A.H.); 819085@nankai.edu.cn (G.S.); zhfzhao@nankai.edu.cn (Z.Z.); shijd@nankai.edu.cn (J.S.) * Correspondence: joyleizhang@nankai.edu.cn; Tel./Fax: +86-(22)-23503617

Abstract: Enhancers regulate multiple genes via higher-order chromatin structures, and they further affect cancer progression. Epigenetic changes in cancer cells activate several cancer-specific enhancers that are silenced in normal cells. These cancer-specific enhancers are potential therapeutic targets of cancer. However, the functions and regulation networks of colorectal-cancer-specific enhancers are still unknown. In this study, we profile colorectal-cancer-specific enhancers and reveal their regulation network through the analysis of HiChIP data that were derived from a colorectal cancer cell line and Hi-C and RNA-seq data that were derived from tissue samples by in silico analysis and in vitro experiments. Enhancer–promoter loops in colorectal cancer cells containing colorectal-cancer-specific enhancers are involved in more than 50% of the topological associated domains (TADs) changed in colorectal cancer cells compared to normal colon cells. In addition, colorectal-cancer-specific enhancers interact with 152 genes that are significantly and highly expressed in colorectal cancer cells. These colorectal-cancer-specific enhancer target genes include *ITGB4*, *RECQL4*, *MSLN*, and *GDF15*. We propose that the regulation network of colorectal-cancer-specific enhancers plays an important role in the progression of colorectal cancer.

Keywords: colorectal-cancer-specific enhancer; long-range interaction; Hi-C; HiChIP; TAD

1. Introduction

Colorectal cancer is the world's fourth most deadly cancer [1]. Previous studies on colorectal cancer have indicated that the sequential accumulation of genetic mutations and chromosomal instability are the most important reasons for the occurrence and development of colorectal cancer [2]. A typical genomic event in the initiation of colorectal cancer is *APC* mutation, followed by *RAS* activation or the function loss of *TP53* [1]. Additionally, the *MAPK*, *WNT*, *PI3K*, and TGF- β signaling pathways have been reported to play important roles in the initiation and progression of colorectal cancer [2,3]. Chromosomal changes including deletion, amplification, and translocation induce alterations in gene expression and function that may further initiate colorectal cancer. Studies on higher-order chromatin structures have revealed that long-range gene regulation networks are important in modulating biological processes in cells [4–6]. Further studies on gene regulation networks may help to unveil the occurrence and progression mechanisms of colorectal cancer.

Gene regulatory elements such as enhancers or silencers can regulate multiple target genes through higher-order chromatin structures [7,8] and functions in gene regulation networks in cancer cells (Figure 1). Mutations in gene regulatory elements and changes in higher-order chromatin structures are related to gene transcription and cancer progression [8–10]. A previous study revealed reorganizations of chromatin structures in colorectal cancers in comparison to normal colons, and the topological changes of chromatins were



Citation: Chen, B.; Ma, Y.; Bi, J.; Wang, W.; He, A.; Su, G.; Zhao, Z.; Shi, J.; Zhang, L. Regulation Network of Colorectal-Cancer-Specific Enhancers in the Progression of Colorectal Cancer. *Int. J. Mol. Sci.* 2021, *22*, 8337. https://doi.org/ 10.3390/ijms22158337

Academic Editor: Andrea Angius

Received: 28 June 2021 Accepted: 1 August 2021 Published: 3 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). found to restrain the malignant progression of colorectal cancers [11]. In addition, changes of higher-order chromatin structures in colorectal cancer cells are accompanied by epigenetic changes such as alterations in DNA methylation [11,12]. These epigenetic changes are associated with changes in the binding of transcription factors and the activation of some gene regulatory elements such as enhancers. Some enhancers are activated in cancer cells but not in normal cells [13]. These "cancer-specific enhancers" may regulate several target genes through higher-order chromatin structures and play roles in cancer progression. These cancer-specific enhancers are potential therapeutic targets of cancer. However, how colorectal-cancer-specific enhancers function in cancer initiation and progression is still not very clear.



Figure 1. Schematic diagram of regulation networks of enhancers/silencers through higher-order chromatin structure in cancer cells. Enhancers/silencers regulate target genes through higher-order chromatin structure and affect the expressions of proteins in signaling pathways and further function in cancer progression.

In this paper, we show the profiles of colorectal-cancer-specific enhancers and their gene regulation networks through the mining of databases and the combination analysis of Hi-ChIP data that were derived from a colorectal cancer cell line and Hi-C and RNA-seq data that were derived from tissue samples. The results of this study suggest that colorectal-cancer-specific enhancers can regulate multiple target genes including cancer-driver genes, as well as that genes regulated by these colorectal-cancer-specific enhancers are associated with the occurrence and development of colorectal cancer.

2. Results

2.1. H3K27ac Profiles Define Colorectal-Cancer-Specific Enhancers

H3K27ac is a well-characterized histone modification marker of active enhancers [14]. To identify colorectal-cancer-specific enhancers, we analyzed published H3K27ac CHIP-seq datasets from seven colorectal cancer datasets and ten normal tissue datasets from Roadmap Epigenomics [15], ENCODE [16], and a previously published study [17]. Unsupervised hierarchical clustering was performed on H3K27ac enrichment loci, identifying 10 clusters from the 59 samples (7 colorectal cancer samples, 10 normal colon samples, 8 brain samples, 5 ESC samples, 4 IPS samples, 13 blood samples, 4 heart samples, 2 melanocyte samples, 4 muscle samples, and 2 fibroblast samples from Roadmap Epigenomics [15], ENCODE [16], and a previously published study [17]) (Figure 2A) and showing that the distribution pattern of enhancers is highly tissue-specific, which is consistent with previous studies [13].

We identified 21,132 normal colon-tissue-specific enhancers and 11,463 colorectal-cancerspecific enhancers, and the *t*-distributed stochastic neighbor embedding (*t*-SNE) analysis indicated that colorectal cancer enhancer profiles are distinct from normal colon tissue enhancers (Figure 2B). In addition, we identified 763 super enhancers among the colorectalcancer-specific enhancers (Figure 2C). Some of these super enhancers are associated with oncogenes, such as *EPHA2*, *LIF*, *ID1*, *SMAD7*, *BMP4*, *FOXA1*, and *NOTCH1*. Super enhancers are characterized as large clusters of enhancers in close proximity, and they are required for the maintenance of cell identity [14,18].



Figure 2. H3K27ac profiles define colorectal-cancer-specific enhancers. (**A**) The unsupervised hierarchical clustering of the 11,463 enhancer loci detected in colorectal-cancer samples (n = 7) compared to that of normal tissue samples (n = 10). (**B**) *t*-distributed stochastic neighbor embedding (*t*-SNE) analysis of normal colon-tissue-specific enhancers (Colon) and colorectal-cancer-specific enhancers. *t*-SNE_1 $p = 2.03 \times 10^{-15}$, *t*-SNE_2 $p = 8.54 \times 10^{-12}$. (**C**) Inflection plot showing the identified super enhancers among the cancer-specific enhancers.

2.2. HiChIP Identifies Chromatin Interactions Containing Colorectal-Cancer-Specific Enhancers

The positions of colorectal-cancer-specific enhancers are shown in Figure 3A; these colorectal-cancer-specific enhancers are evenly distributed on each chromosome (Figure 3A). To identify the regulation network of colorectal-cancer-specific enhancers, HiChIP for H3K27ac was performed in the HCT116 colorectal cancer cell line. The HiChIP data indicated interaction loops between enhancers/super enhancers and other chromatin fragments (in Figure 3B,C, interaction loops in chromosome 1 are shown as examples). Additionally, colorectal-cancer-specific enhancers/super enhancers interact with several chromatin fragments, including some long-range chromatin regions (in Figure 3D,E, the interaction loops in chromosome 1 are shown as examples). Most of the interaction loops

that contain colorectal-cancer-specific enhancers are less than 100 kb in length and include more short-range interactions than the random interaction loops from the HiChIP data (Figure 3F,G). In addition, the interaction loops that were measured to be more than 100 kb in length—which have lower proportions (Figure 3F,G), but longer interaction loops (>100 kb)—that contain colorectal-cancer-specific enhancers did not have a significantly lower proportion (Figure 3F,G), which suggests that more long-range interactions are associated with colorectal-cancer-specific enhancers.



Figure 3. HiChIP identified chromatin interactions containing colorectal-cancer-specific enhancers. (**A**) Distributions of colorectal-cancer-specific enhancers of the human genome (hg19). (**B**,**C**) Circos plot showing interactions in Chr1, as indicated by curves extending from enhancers and super enhancers in colorectal cancer cells. Each curve in the Circos plot indicates one interaction loop that contains enhancers or super enhancers. (**D**,**E**) Circos plot showing interactions in Chr1, as indicated by curves extending from colorectal-cancer-specific enhancers and super enhancers in colorectal cancer cells. Each curve in the Circos plot showing interactions in Chr1, as indicated by curves extending from colorectal-cancer-specific enhancers and super enhancers in colorectal-cancer cells. Each curve in the Circos plot indicates one interaction loop that contains colorectal-cancer-specific enhancers or super enhancers. (**F**,**G**) An analysis of the length distribution of random interaction loops from HiChIP data (red), the interaction loops that contain colorectal-cancer-specific enhancers (green) and colorectal-cancer-specific enhancers (blue, *p* < 2.2 × 10⁻¹⁶ compared with the other two groups).

2.3. Change in TAD Boundaries in Colorectal Cancer Cells Compared to Normal Colon Cells

A comparison to normal colon cells and an analysis of Hi-C data in colorectal cancer cells and normal colon cells [11] showed that about 65% of the boundaries of TADs changed in colorectal cancer cells (Figure 4A–F). The boundary-changed TADs were divided into six categories—expand (Figure 4A), narrow (Figure 4B), fusion (Figure 4C), split (Figure 4D), appear (Figure 4E), and disappear (Figure 4F)—that represent six kinds of change patterns of TAD boundaries in colorectal cancer cells compared to normal colon cells. Previous studies have indicated that gene transcription is correlated with insulation at TAD boundaries [6,11]. Expand, fusion, and appearing TADs in colorectal cancer cells are associated with more interactions between enhancers and genes, whereas split, narrow, and disappearing TADs are associated with less interactions between some enhancers and genes in colorectal cancer cells. The sixty-five percent of changed TADs in colorectal cancer cells in comparison to normal colon cells suggests global alterations of gene regulations in colorectal cancer cells. We noticed that more than 50% (51–86%, except for the disappear group, which is 24%) of these changed TADs contain colorectal-cancer-specific enhancers in the alterations of the gene regulation network.



Figure 4. Changes of TAD boundaries in colorectal cancer cells compared to normal colon cells. (**A**) Expanded TADs in colorectal cancer cells compared to normal colon cells. (**B**) Narrowed TADs in colorectal cancer cells compared to normal colon cells. (**C**) Fusion TADs in colorectal cancer cells, where two or more TADs in normal colon cells are fused to one TAD in colorectal cancer cells. (**D**) Split TADs in colorectal cancer cells, where new insulation boundaries are formed in one TAD in colorectal cancer cells. (**E**) Appeared TADs in colorectal cancer cells that do not exist in normal colon cells. (**F**) Disappeared TADs, which are TADs in normal colon cells that disappear in colorectal cancer cells. The pie charts (**A**–**F**) show the percentages of specific categories of boundary-changed TADs (category-TAD) in the total changed TADs (C-TAD) in colorectal cancer cells (the first pie chart in each panel) and the percentages of specific categories of boundary-changed TADs containing colorectal-specific enhancers (SpE-category-TAD) in the total changed TADs of this category (the second pie chart in each panel).

2.4. Transcriptome Change in Colorectal Cancer Cells Is Associated with Colorectal-Specific Enhancers

To determine whether the transcription change of genes in colorectal cancer cells is associated with colorectal-specific enhancers, we analyzed the transcriptome data of colon tissues (n = 349) and colorectal cancer tissues (n = 275) from TCGA. We found 4714 genes to be significantly upregulated (fold change > 1.5; q < 0.01) and 4924 genes to be significantly downregulated (fold change > 1.5; q < 0.01) in colorectal cancer tissue (Figure 5A). The combination analysis of RNA-seq and HiChIP data identified 152 genes that are significantly upregulated (fold change > 1.5; q < 0.01) in colorectal cancer tissue and have comparatively high interaction signals (interaction counts \geq 5) with colorectal-cancerspecific enhancers (Figure 5B). These genes are potential target genes that are directly regulated by colorectal-cancer-specific enhancers. About 79% of the interactions between colorectal-cancer-specific enhancers and target genes are long-range (≥20 kb) interactions (Figure 5C). A pathway and process enrichment analysis by Metascape [19] showed that potential target genes of colorectal-cancer-specific enhancers are enriched in the RNA catabolic process, the $\alpha 6\beta 1\alpha 6\beta 4$ integrin pathway, and the regulation of cyclin-dependent protein serine/threonine kinase activity (Figure 5D). These pathways or biological processes are associated with colorectal cancer progression [20–24].



Figure 5. Transcriptome change in colorectal cancer cells is associated with colorectal-specific enhancers. (**A**) Volcano plot depicting gene expression changes in normal colon tissue (n = 349) and colorectal cancer tissue (n = 275). (**B**) Comparison of fold changes in gene expression based on RNA-seq and HiChIP signals. The plot shows 735 target genes. Red dots indicate target genes that were found to significantly change (fold change > 1.5; q < 0.01) in colorectal cancer tissue compared to those of normal colon tissue and have significant HiChIP signals (counts \geq 5). (**C**) Percentage of long-range (>20 kb) interaction loops between colorectal-cancer-specific enhancers and target genes. (**D**) Functions of genes as analyzed by Metascape.

ITGB4 encodes integrin subunit β 4 and is involved in the $\alpha 6\beta 1 \alpha 6\beta 4$ integrin pathway. HiChIP data showed that *ITGB4* is regulated by a colorectal-cancer-specific enhancer (Figure 6A), and RNA-seq data showed that *ITGB4* is upregulated (fold change = 2.31) in colorectal cancer tissue, suggesting the regulation of *ITGB4* by colorectal-cancer-specific enhancers. Mutations and aberrant expression of cancer-driver genes affect key cellular functions and induce cancer occurrence. *RECQL4* is a cancer-driver gene [25] and potential target gene that is regulated by colorectal-cancer-specific enhancers (Figure 6B). In addition, *RECQL4* is one of the most often mutated genes in colorectal cancer [26], and it is

upregulated (fold change = 4.10) in colorectal cancer tissues. The colorectal-cancer-specific enhancer is about 15 kb away from the *RECQL4* promoter, and HiChIP data showed interactions between colorectal-cancer-specific enhancers and *RECQL4* (Figure 6B). In addition, *MSLN* and *GDF15* were found to be significantly upregulated (fold change = 12.24 for *MSLN* and fold change = 14.76 for *GDF15*) in colorectal cancer tissues, as well as to interact with colorectal-cancer-specific enhancers in colorectal cancer cells (Figure 6C, D). Mesothelin (MSLN) is reportedly associated with the prognosis of colorectal cancer [27], and growth differentiation factor 15 (GDF15) is reportedly associated with the recurrence and 5-fluorouracil resistance of colorectal cancer [28–30]. ChIP and HiChIP data showed that several colorectal-specific enhancers interact with *MSLN* and *GDF15*, and some of these enhancers interact with *MSLN* or *GDF15* through long-range chromatin interactions (>20 kb) (Figure 6C, D).



Figure 6. Target genes of the colorectal-cancer-specific enhancers. (**A**) H3K27ac enrichment for normal colon cells (N; blue) and colorectal cancer cells (C; red), as well as HiChIP interaction loops (red curves) between the *ITGB4* target gene and colorectal-cancer-specific enhancers. Colorectal-cancer-specific enhancers are shaded in red. Blue curves indicate other interaction loops in this area. (**B**) H3K27ac enrichment for normal colon cells (N; blue) and colorectal cancer cells (C; red), as well as HiChIP interaction loops (red curves) between the *RECQL4* target gene and colorectal-cancer-specific enhancers. The colorectal-cancer-specific enhancers are shaded in red. Blue curves indicate other interaction loops in this area. (**C**) H3K27ac enrichment for normal colon cells (N; blue) and colorectal cancer cells (C; red), as well as HiChIP interaction loops in this area. (**C**) H3K27ac enrichment for normal colon cells (N; blue) and colorectal cancer cells (C; red), as well as HiChIP interaction loops (red curves) between the *MSLN* target gene and colorectal-cancer-specific enhancers. Colorectal-cancer-specific enhancers are shaded in red. Blue curves indicate other interaction loops in this area. (**D**) H3K27ac enrichment for normal colon cells (N; blue) and colorectal cancer cells (C; red), as well as HiChIP interaction loops (red curves) between the *MSLN* target gene and colorectal-cancer-specific enhancers. **(D)** H3K27ac enrichment for normal colon cells (N; blue) and colorectal cancer cells (C; red), as well as HiChIP interaction loops (red curves) between the *GDF15* target gene and colorectal-cancer-specific enhancers. Colorectal-cancer-specific enhancers are shaded in red. Blue curves indicate other interaction colors in this area.

3. Discussion

This study shows that colorectal-specific enhancers interact with several potential target genes that are upregulated in colorectal cancer cells compared with normal colon cells, which is consistent with a previous study that show changes in the epigenome at enhancers are associated with transcriptional program to promote colon carcinogenesis [17]. Our study further unveils the regulation network of the colorectal-cancer-specific enhancers in the view of higher-order chromatin structures. HiChIP data indicated that higher-order chromatin structures induce interactions between colorectal-cancer-specific enhancers and target genes. Most altered TADs in colorectal cancer genome contain colorectal-specific enhancers. However, it is unclear whether changes in chromatin topological structures or the activation of colorectal-specific enhancers comprise the primary event in the initiation of colorectal cancer. A previous study showed that chromosomal topological changes repress stemness and invasion programs; these changes may restrain the malignant progression of colorectal cancer, and tumor-associated epigenomic changes are primarily oncogenic [11]. HiChIP data showed that some target genes of colorectal-cancer-specific enhancers are cancer-driver genes, including *RECQL4* (shown in Figure 6B), and an analysis of Hi-C data indicated that most boundary-changed TADs contain colorectal-specific enhancers (Figure 4). It is possible that the activation of colorectal-specific enhancers occurs before the alteration of higher-order chromatin structures during the occurrence and progression of colorectal cancer. Further studies are needed to verify this hypothesis.

With the development of chromatin conformation capture techniques [31–34], the analysis of higher-order chromatin structures in cancer cells has revealed complicated long-range gene networks of enhancers and indicated important functions of these enhancers or gene networks in cancer progression [8–10,13]. A recent study reported that mutations in enhancer loci lead to activity changes in the enhancer, affect the expression of the long-range target genes of this enhancer, and further improve the progression of prostate cancer [8]. A previous study showed that ependymoma enhancer profiles are distinct from other tissues and most super enhancers are tumor-specific and enriched with cancer-associated genes [13], suggesting important functions of these tumor-specific enhancers in the initiation of tumors. We profiled the colorectal-cancer-specific enhancers and presented the regulation networks of these enhancers based on HiChIP data, and RNA-seq data further unveiled the initiation and progression mechanisms of colorectal cancer.

The combination analysis of HiChIP data and RNA-seq data indicated several potential target genes of colorectal-specific enhancers. ITGB4 is one of these potential target genes. ITGB4 is one of the integrin (ITG) molecules that is associated with cell migration, proliferation, and cancer development [35–37]. Aberrant ITGB4 expression has been reported in several cancers, including colorectal cancers [21,38]. RNA-seq analysis has also shown the significant upregulation of *ITGB4* in colorectal cancer tissue. A recent study indicated that ITGB4 might be a prognosis marker for the individual therapy of colon cancer [21]. Data in our study suggested that the upregulation of ITGB4 in colorectal cancer cells may due to the regulation of colorectal-cancer-specific enhancers. We also noticed that GALK1, which is beside ITGB4, may be one of the potential target genes of the colorectalcancer-specific enhancer that is shown in Figure 6A. HiChIP data indicated interactions between the colorectal-cancer-specific enhancer and GALK1, and RNA-seq analysis showed upregulation (fold change = 1.97; q = 7.76×10^{-49}) of *GALK1* in colorectal cancer tissue compared to that in normal colon tissue. However, the interaction signal was not shown to be robust in the HiChIP data (count = 3). This may have been due to the long-range interactions between the colorectal-cancer-specific enhancer and GALK1 (~45 kb). Our study suggests that GALK1 may be regulated by colorectal-cancer-specific enhancers and be associated with the progression of colorectal cancer. However, few studies have reported the function of *GALK1* in colorectal cancer until now.

RECQL4 is an example of a cancer-driver gene that is regulated by colorectal-cancerspecific enhancers. RECQL4 was reported to be a DNA helicase and to function in DNA replication, repair, and recombination [39], and it plays an important role in maintaining genomic stability [40]. Several studies have indicated that RECQL4 is associated with cancer progression and may be a diagnostic marker for cancer [40–44], e.g., the overexpression of RECQL4 is associated with a poor prognosis of gastric cancer [42] and predicts poor prognosis in hepatocellular carcinoma [43]. An investigation of the mutation spectrum of cancer-associated genes in patients with early onset colorectal cancer showed that *RECQL4* is one of the most often mutated genes [26]. Additionally, the present analysis of RNA-seq data from 275 colon cancer tissues and 349 normal colon tissues showed that RECQL4 is upregulated in colorectal cancer cells. Based on the literature and the data analysis in the present study, RECQL4 may be an important target gene of colorectalcancer-specific enhancers and involved in the progression of colorectal cancer. HiChIP data also indicated that MLSN and GDF15 interact with colorectal-specific enhancers through long-range chromatin interactions. MSLN is a cell surface protein that is reportedly highly expressed in several types of malignant tumors, including colorectal cancer [27]. Highly expressed MSLN induces an increase in cell proliferation in colorectal cancer, as well as worse prognosis [27]. In addition, MSLN was reported as one of the top cell therapy targets for tumors [45]. The upregulation of GDF15 was reported correlate with an increased risk of recurrence and a decreased overall survival of colorectal cancer [28]. Our study suggests that the target genes of colorectal-cancer-specific enhancers may be potential targets for treating colorectal cancer. Methods such as the use of zinc-finger proteins and dCas9 to modify chromatin structures or manipulate specific enhancer-promoter interactions have been reported [46–49], indicating the potential to treat cancer by focusing on 3D chromatin structures. Our study suggests that the chromatin interactions of colorectal-cancer-specific enhancers and their target genes can be potential new targets to treat colorectal cancer. The analyses in this study were mainly based on database mining and in vitro HiChIP data from one colorectal cancer cell line, so more in vitro/vivo experiments on larger scales of human tissue samples and mouse models are needed to further validate the regulations of specific target genes by the colorectal-cancer-specific enhancers, as well as further unveil the whole regulation network of colorectal-cancer-specific enhancers. Though further studies are needed to pinpoint whether there are key driver genes in the colorectal-cancer-specific enhancer regulation network or the progression of colorectal cancer is a combined result of these target genes, our study suggests important functions of colorectal-cancer-specific enhancers in colorectal cancer progression and that these enhancers may be potential therapeutic targets of colorectal cancer.

4. Materials and Methods

4.1. Cell Culture

The colorectal cancer cell line HCT116 was obtained from the ATCC. HCT116 cells were incubated at 37 °C with 5% CO₂ and cultured in Dulbecco's Modified Eagle Medium (GIBCO) supplemented with 10% fetal bovine serum (FBS) (BI) and 1% penicillin-streptomy cin (GIBCO).

4.2. HiChIP

HiChIP was performed as previously described with modifications [33]. Briefly, cells were crosslinked with 1% formaldehyde and lysed. Then, the chromatin was digested using *Mbo*I (NEB), and the restricted ends were ligated by T4 ligase (NEB). Pelleted nuclei were dissolved in a nuclear lysis buffer (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1% SDS, and protease inhibitors), and they were sonicated and diluted in a ChIP Dilution Buffer (0.01% SDS, 1.1% Triton X–100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 7.5, and 167 mM NaCl). Then, immunoprecipitation was performed overnight at 4 °C by incubating H3K27ac antibodies (Abcam, Cambridge, UK) precoated on protein A-coated magnetic beads (Thermo Fisher Scientific, Waltham, MA, USA). Immunocomplexes were washed three times each with a low-salt buffer (0.1% SDS, 1% Triton X–100, 2 mM EDTA, 20 mM Tris-HCl, pH 7.5, and 500 mM NaCl), and an LiCl buffer (10 mM Tris-HCl, pH 7.5, cont 500 mM NaCl), and an LiCl buffer (10 mM Tris-HCl, pH 7.5, cont 500 mM NaCl), and an LiCl buffer (10 mM Tris-HCl, pH 7.5, cont 500 mM NaCl).

pH 7.5, 250 mM LiCl, 1% NP-40, 1% Na-Doc, and 1 mM EDTA). Beads were resuspended in a DNA elution buffer (50 mM NaHCO₃ and 1% SDS). After elution, ChIP samples were incubated with 10 mg/mL of proteinase K 4 h at 55 °C. Then, DNA was purified using AMPure XP Beads (Beckman, Brea, CA, USA). Streptavidin C1 beads were used to capture biotinylated DNA. QIAseq FX DNA Library Kits were used to generate the sequencing library, and HiChIP libraries were size-selected to 300–700 bp using AMPure XP beads (Beckman) and subjected to 2 × 50-bp paired-end sequencing on HiSeq XTen (Illumina, San Diego, CA, USA).

4.3. ChIP-Seq Data Processing and Super Enhancer Analysis

The mapping of ChIP-seq data (Table 1) to hg19 was performed using bowtie2 [50]. H3K27ac peak calling was performed using MACS1.4 with default parameters [51]. Peak calling was separately performed for each sample. Peaks that could not be identified in each colorectal cancer sample and peaks that appeared within the region surrounding ± 2.5 kb of transcriptional start sites were excluded from any further analysis. Afterwards, the H3K27ac peaks of the 7 individual samples (7 published colorectal cancer H3K27ac CHIP-seq datasets from ENCODE [16] and a published study [17]) were merged into a single set of peaks. Super enhancers were identified using the rank ordering of super enhancers (ROSE) algorithm [14]. For colorectal-cancer-specific enhancers, we removed all peak regions that contained any overlap with a peak detected in each normal colon region.

| Tab | le | 1. | ChIP-seq | l data | sets. |
|-----|----|----|----------|--------|-------|
|-----|----|----|----------|--------|-------|

| Sample Class | NO. | Sample | Data Source |
|-------------------|-----|------------------------|-------------|
| | 1 | E075-H3K27ac.broadPeak | Roadmap |
| | 2 | E101-H3K27ac.broadPeak | Roadmap |
| | 3 | E102-H3K27ac.broadPeak | Roadmap |
| | 4 | E106-H3K27ac.broadPeak | Roadmap |
| | 5 | ENCFF116PPS | ENCODE |
| Normal Colon | 6 | ENCFF154YYF | ENCODE |
| | 7 | ENCFF193RGE | ENCODE |
| | 8 | ENCFF345LME | ENCODE |
| | 9 | ENCFF389IUW | ENCODE |
| | 10 | ENCFF779JEH | ENCODE |
| | 1 | V400_H3K27ac | GSE36204 |
| | 2 | V429_H3K27ac | GSE36204 |
| | 3 | V503_H3K27ac | GSE36204 |
| Colorectal Cancer | 4 | V9M_H3K27ac | GSE36204 |
| | 5 | ENCFF436VPF | ENCODE |
| | 6 | ENCFF570UFJ | ENCODE |
| | 7 | ENCFF720SPY | ENCODE |

4.4. Unsupervised Hierarchical Clustering Analysis

A matrix of the normalized H3K27ac density was generated using HOMER [52]. Variant enhancer loci (VELs) were defined as enhancers that exhibited the greatest median absolute deviation (MAD) across all samples used for clustering. After the unsupervised hierarchical clustering between colorectal cancer and Roadmap Epigenomics samples data, 11,463 VELs were retained. These enhancers were used for unsupervised hierarchical clustering a Pearson correlation as a distance metric.

4.5. t-SNE Analysis

For the clustering of H3K27ac ChIP-seq data from the colorectal cancer and normal colon cohorts together, we generated a normalized H3K27ac density matrix for both cohorts. The distance between samples was calculated by using "1- Spearman correlation coefficient" as the distance measure. The resulting distance matrix was used to perform the *t*-SNE analysis (Rtsne package).

4.6. HiChIP Data Processing

Valid fragment pairs and interaction matrixes were generated using Hi-C-Pro [53]. Then, interaction loop calling was performed using the hichipper [54], and the resulting interaction loops in Figure 3B–E are shown as a Circos plot [55]. For colorectal-cancer-specific enhancer loops detection, we selected loops that contained anchors that overlapped with cancer-specific enhancers.

4.7. Hi-C Data Analysis

Two HiC data sets in GSE133928, BRD3179 and BRD3179N [11], representing colorectal cancer tissue and normal colon tissue, respectively, were analyzed to detect TADs in colorectal cancer cells and normal colon cells. Expanded TADs or narrowed TADs refers to the boundaries of TADs in colorectal cancer cells that are expanded or narrowed, respectively, in comparison to normal colon cells. Fusion TADs refers to two or more TADs in normal colon cells fused to one TAD in colorectal cancer cells. In comparison to normal colon cells, new insulation boundaries were formed in one TAD in colorectal cancer cells; this is referred to as a split TAD. Appeared TADs refers to TADs in colorectal cancer cells that do not exist in normal colon cells. Disappeared TADs refers to TADs in normal colon cells that disappear in colorectal cancer cells.

5. Conclusions

In summary, our study shows that colorectal-cancer-specific enhancers regulate several target genes including cancer-driver genes and may play important roles in the occurrence and progression of colorectal cancer.

Author Contributions: B.C. and L.Z., conception, design, drafting, and revision of the article; B.C. and L.Z., acquisition of, analysis, and interpretation of data; B.C., Y.M., A.H., J.B., W.W., and G.S., acquisition of data; B.C., analysis of all high-throughput sequencing data; B.C. and W.W., acquisition of data and production of all bioinformatics figures; J.S., Z.Z., and L.Z., revision of the article. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Fundamental Research Funds for the Central Universities, Nankai University (63201087) and the National Natural Science Foundation of China (81772687).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: HiChIP data were deposited in the NCBI Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/, accessed on 5 May 2021) under the accession number GSE173699. The following published datasets were used in our analysis: the GSE36204 dataset was used for H3K27ac ChIP-seq analyses [17]. Other ChIP-seq data were obtained from Roadmap Epigenomics [15] (http://www.roadmapepigenomics.org/, accessed on 24 December 2020) and ENCODE (https://www.encodeproject.org/, accessed on 24 December 2020) [16]. The GSE133928 dataset was used for Hi-C analysis (Hi-C data in colorectal cancer cells and normal colon cells) [11], and RNA-seq data were obtained from the TCGA and GEPIA database (http://gepia.cancer-pku.cn/, accessed on 10 January 2021) [56].

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

References

- 1. Dekker, E.; Tanis, P.J.; Vleugels, J.L.A.; Kasi, P.M.; Wallace, M.B. Colorectal cancer. *Lancet* 2019, 394, 1467–1480. [CrossRef]
- Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012, 487, 330–337. [CrossRef]
- 3. Fearon, E.R. Molecular genetics of colorectal cancer. Ann. Rev. Pathol. 2011, 6, 479–507. [CrossRef]
- 4. Wei, Z.; Gao, F.; Kim, S.; Yang, H.; Lyu, J.; An, W.; Wang, K.; Lu, W. Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. *Cell Stem Cell* **2013**, *13*, 36–47. [CrossRef]
- Su, G.; Guo, D.; Chen, J.; Liu, M.; Zheng, J.; Wang, W.; Zhao, X.; Yin, Q.; Zhang, L.; Zhao, Z.; et al. A distal enhancer maintaining Hoxa1 expression orchestrates retinoic acid-induced early ESCs differentiation. *Nucleic Acids Res.* 2019, 47, 6737–6752. [CrossRef] [PubMed]

- Bonev, B.; Mendelson Cohen, N.; Szabo, Q.; Fritsch, L.; Papadopoulos, G.L.; Lubling, Y.; Xu, X.; Lv, X.; Hugnot, J.P.; Tanay, A.; et al. Multiscale 3D genome rewiring during mouse neural development. *Cell* 2017, *171*, 557–572.e24. [CrossRef]
- Zhang, L.; He, A.; Chen, B.; Bi, J.; Chen, J.; Guo, D.; Qian, Y.; Wang, W.; Shi, T.; Zhao, Z.; et al. A HOTAIR regulatory element modulates glioma cell sensitivity to temozolomide through long-range regulation of multiple target genes. *Genome Res.* 2020, 30, 155–163. [CrossRef]
- 8. Qian, Y.; Zhang, L.; Cai, M.; Li, H.; Xu, H.; Yang, H.; Zhao, Z.; Rhie, S.K.; Farnham, P.J.; Shi, J.; et al. The prostate cancer risk variant rs55958994 regulates multiple gene expression through extreme long-range chromatin interaction to control tumor progression. *Sci. Adv.* **2019**, *5*, eaaw6710. [CrossRef]
- Kloetgen, A.; Thandapani, P.; Ntziachristos, P.; Ghebrechristos, Y.; Nomikou, S.; Lazaris, C.; Chen, X.; Hu, H.; Bakogianni, S.; Wang, J.; et al. Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. *Nat. Genet.* 2020, *52*, 388–400. [CrossRef] [PubMed]
- 10. Rhie, S.K.; Perez, A.A.; Lay, F.D.; Schreiner, S.; Shi, J.; Polin, J.; Farnham, P.J. A high-resolution 3D epigenomic map reveals insights into the creation of the prostate cancer transcriptome. *Nat. Commun.* **2019**, *10*, 4154. [CrossRef] [PubMed]
- 11. Johnstone, S.E.; Reyes, A.; Qi, Y.; Adriaens, C.; Hegazi, E.; Pelka, K.; Chen, J.H.; Zou, L.S.; Drier, Y.; Hecht, V.; et al. Large-scale topological changes restrain malignant progression in colorectal cancer. *Cell* **2020**, *182*, 1474–1489.e23. [CrossRef] [PubMed]
- 12. Berman, B.P.; Weisenberger, D.J.; Aman, J.F.; Hinoue, T.; Ramjan, Z.; Liu, Y.; Noushmehr, H.; Lange, C.P.; van Dijk, C.M.; Tollenaar, R.A.; et al. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat. Genet.* **2011**, *44*, 40–46. [CrossRef]
- 13. Mack, S.C.; Pajtler, K.W.; Chavez, L.; Okonechnikov, K.; Bertrand, K.C.; Wang, X.; Erkek, S.; Federation, A.; Song, A.; Lee, C. Therapeutic targeting of ependymoma as informed by oncogenic enhancer profiling. *Nature* **2018**, *553*, 101–105. [CrossRef]
- 14. Whyte, W.A.; Orlando, D.A.; Hnisz, D.; Abraham, B.J.; Lin, C.Y.; Kagey, M.H.; Rahl, P.B.; Lee, T.I.; Young, R.A. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **2013**, *153*, 307–319. [CrossRef]
- 15. Roadmap Epigenomics, C.; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; et al. Integrative analysis of 111 reference human epigenomes. *Nature* **2015**, *518*, 317–330.
- 16. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. Nature 2012, 489, 57–74. [CrossRef]
- Akhtar-Zaidi, B.; Cowper-Sal-Lari, R.; Corradin, O.; Saiakhova, A.; Bartels, C.F.; Balasubramanian, D.; Myeroff, L.; Lutterbaugh, J.; Jarrar, A.; Kalady, M.F.; et al. Epigenomic enhancer profiling defines a signature of colon cancer. *Science* 2012, 336, 736–739. [CrossRef]
- 18. Thandapani, P. Super-enhancers in cancer. Pharmacol. Ther. 2019, 199, 129–138. [CrossRef]
- 19. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A.H.; Tanaseichuk, O.; Benner, C.; Chanda, S.K. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **2019**, *10*, 1523. [CrossRef] [PubMed]
- Mercurio, A.M.; Bachelder, R.E.; Rabinovitz, I.; O'Connor, K.L.; Tani, T.; Shaw, L.M. The metastatic odyssey: The integrin connection. Surg. Oncol. Clin. N. Am. 2001, 10, 313–328. [CrossRef]
- Li, M.; Jiang, X.; Wang, G.; Zhai, C.; Liu, Y.; Li, H.; Zhang, Y.; Yu, W.; Zhao, Z. ITGB4 is a novel prognostic factor in colon cancer. J. Cancer 2019, 10, 5223–5233. [CrossRef]
- 22. Goodall, G.J.; Wickramasinghe, V.O. RNA in cancer. Nat. Rev. Cancer 2021, 21, 22–36. [CrossRef] [PubMed]
- Shi, X.N.; Li, H.; Yao, H.; Liu, X.; Li, L.; Leung, K.S.; Kung, H.F.; Lin, M.C. Adapalene inhibits the activity of cyclin-dependent kinase 2 in colorectal carcinoma. *Mol. Med. Rep.* 2015, 12, 6501–6508. [CrossRef]
- 24. Cam, W.R.; Masaki, T.; Shiratori, T.Y.; Kato, N.; Okamoto, M.; Yamaji, Y.; Igarashi, K.; Sano, T.; Omata, M. Activation of cyclin E-dependent kinase activity in colorectal cancer. *Digest. Dis. Sci.* 2001, *46*, 2187–2198. [CrossRef]
- 25. Martinez-Jimenez, F.; Muinos, F.; Sentis, I.; Deu-Pons, J.; Reyes-Salazar, I.; Arnedo-Pac, C.; Mularoni, L.; Pich, O.; Bonet, J.; Kranas, H.; et al. A compendium of mutational cancer driver genes. *Nat. Rev. Cancer* **2020**, *20*, 555–572. [CrossRef]
- 26. Zhunussova, G.; Afonin, G.; Abdikerim, S.; Jumanov, A.; Perfilyeva, A.; Kaidarova, D.; Djansugurova, L. Mutation spectrum of cancer-associated genes in patients with early onset of colorectal cancer. *Front. Oncol.* **2019**, *9*, 673. [CrossRef] [PubMed]
- Inoue, S.; Tsunoda, T.; Riku, M.; Ito, H.; Inoko, A.; Murakami, H.; Ebi, M.; Ogasawara, N.; Pastan, I.; Kasugai, K.; et al. Diffuse mesothelin expression leads to worse prognosis through enhanced cellular proliferation in colorectal cancer. *Oncol. Lett.* 2020, 19, 1741–1750. [CrossRef] [PubMed]
- Wallin, U.; Glimelius, B.; Jirstrom, K.; Darmanis, S.; Nong, R.Y.; Ponten, F.; Johansson, C.; Pahlman, L.; Birgisson, H. Growth differentiation factor 15: A prognostic marker for recurrence in colorectal cancer. *Br. J. Cancer* 2011, *104*, 1619–1627. [CrossRef] [PubMed]
- 29. Wang, B.; Ma, N.; Zheng, X.; Li, X.; Ma, X.; Hu, J.; Cao, B. GDF15 repression contributes to 5-fluorouracil resistance in human colon cancer by regulating epithelial-mesenchymal transition and apoptosis. *BioMed Res. Int.* **2020**, 2020, 2826010. [CrossRef]
- 30. Zheng, X.; Ma, N.; Wang, X.; Hu, J.; Ma, X.; Wang, J.; Cao, B. Exosomes derived from 5-fluorouracil-resistant colon cancer cells are enriched in GDF15 and can promote angiogenesis. *J. Cancer* 2020, *11*, 7116–7126. [CrossRef]
- 31. Dekker, J.; Rippe, K.; Dekker, M.; Kleckner, N. Capturing chromosome conformation. Science 2002, 295, 1306–1311. [CrossRef]
- 32. Belaghzal, H.; Dekker, J.; Gibcus, J.H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* 2017, *123*, 56–65. [CrossRef]
- 33. Mumbach, M.R.; Rubin, A.J.; Flynn, R.A.; Dai, C.; Khavari, P.A.; Greenleaf, W.J.; Chang, H.Y. HiChIP: Efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **2016**, *13*, 919–922. [CrossRef] [PubMed]

- Lieberman-Aiden, E.; van Berkum, N.L.; Williams, L.; Imakaev, M.; Ragoczy, T.; Telling, A.; Amit, I.; Lajoie, B.R.; Sabo, P.J.; Dorschner, M.O.; et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 2009, 326, 289–293. [CrossRef] [PubMed]
- 35. Moreno-Layseca, P.; Streuli, C.H. Signalling pathways linking integrins with cell cycle progression. *Matrix Biol. J. Int. Soc. Matrix Biol.* **2014**, *34*, 144–153. [CrossRef]
- Wilkinson, E.J.; Woodworth, A.M.; Parker, M.; Phillips, J.L.; Malley, R.C.; Dickinson, J.L.; Holloway, A.F. Epigenetic regulation of the ITGB4 gene in prostate cancer. *Exp. Cell Res.* 2020, 392, 112055. [CrossRef] [PubMed]
- 37. Stewart, R.L.; O'Connor, K.L. Clinical significance of the integrin alpha6beta4 in human malignancies. *Lab. Investig. J. Tech. Methods Pathol.* **2015**, *95*, 976–986. [CrossRef]
- 38. Ni, H.; Dydensborg, A.B.; Herring, F.E.; Basora, N.; Gagne, D.; Vachon, P.H.; Beaulieu, J.F. Upregulation of a functional form of the beta4 integrin subunit in colorectal cancers correlates with c-Myc expression. *Oncogene* **2005**, *24*, 6820–6829. [CrossRef]
- 39. Croteau, D.L.; Popuri, V.; Opresko, P.L.; Bohr, V.A. Human RecQ helicases in DNA repair, recombination, and replication. *Ann. Rev. Biochem.* **2014**, *83*, 519–552. [CrossRef] [PubMed]
- Mo, D.; Zhao, Y.; Balajee, A.S. Human RecQL4 helicase plays multifaceted roles in the genomic stability of normal and cancer cells. *Cancer Lett.* 2018, 413, 1–10. [CrossRef]
- Guo, L.; Li, Y.; Zhao, C.; Peng, J.; Song, K.; Chen, L.; Zhang, P.; Ma, H.; Yuan, C.; Yan, S.; et al. RECQL4, negatively regulated by miR-10a-5p, facilitates cell proliferation and invasion via MAFB in ovarian cancer. *Front. Oncol.* 2020, 10, 524128. [CrossRef]
- 42. Chen, H.; Yuan, K.; Wang, X.; Wang, H.; Wu, Q.; Wu, X.; Peng, J. Overexpression of RECQL4 is associated with poor prognosis in patients with gastric cancer. *Oncol. Lett.* **2018**, *16*, 5419–5425. [CrossRef] [PubMed]
- Li, J.; Jin, J.; Liao, M.; Dang, W.; Chen, X.; Wu, Y.; Liao, W. Upregulation of RECQL4 expression predicts poor prognosis in hepatocellular carcinoma. *Oncol. Lett.* 2018, 15, 4248–4254. [CrossRef]
- 44. Arora, A.; Agarwal, D.; Abdel-Fatah, T.M.; Lu, H.; Croteau, D.L.; Moseley, P.; Aleskandarany, M.A.; Green, A.R.; Ball, G.; Rakha, E.A.; et al. RECQL4 helicase has oncogenic potential in sporadic breast cancers. *J. Pathol.* **2016**, *238*, 495–501. [CrossRef] [PubMed]
- 45. Upadhaya, S.; Yu, J.X.; Shah, M.; Correa, D.; Partridge, T.; Campbell, J. The clinical pipeline for cancer cell therapies. *Nat. Rev. Drug Discov.* **2021**, *20*, 503–504. [CrossRef] [PubMed]
- 46. Deng, W.; Lee, J.; Wang, H.; Miller, J.; Reik, A.; Gregory, P.D.; Dean, A.; Blobel, G.A. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **2012**, *149*, 1233–1244. [CrossRef]
- 47. Deng, W.; Rupon, J.W.; Krivega, I.; Breda, L.; Motta, I.; Jahn, K.S.; Reik, A.; Gregory, P.D.; Rivella, S.; Dean, A.; et al. Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **2014**, *158*, 849–860. [CrossRef] [PubMed]
- 48. Morgan, S.L.; Mariano, N.C.; Bermudez, A.; Arruda, N.L.; Wu, F.; Luo, Y.; Shankar, G.; Jia, L.; Chen, H.; Hu, J.F.; et al. Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. *Nat. Commun.* **2017**, *8*, 15993. [CrossRef]
- Feng, Y.; Pauklin, S. Revisiting 3D chromatin architecture in cancer development and progression. *Nucleic Acids Res.* 2020, 48, 10632–10647. [CrossRef]
- 50. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. Nat. Methods 2012, 9, 357–359. [CrossRef]
- Feng, J.; Liu, T.; Qin, B.; Zhang, Y.; Liu, X.S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* 2012, 7, 1728–1740. [CrossRef] [PubMed]
- Heinz, S.; Benner, C.; Spann, N.; Bertolino, E.; Lin, Y.C.; Laslo, P.; Cheng, J.X.; Murre, C.; Singh, H.; Glass, C.K. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 2010, *38*, 576–589. [CrossRef]
- 53. Servant, N.; Varoquaux, N.; Lajoie, B.R.; Viara, E.; Chen, C.J.; Vert, J.P.; Heard, E.; Dekker, J.; Barillot, E. HiC-Pro: An optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **2015**, *16*, 259. [CrossRef] [PubMed]
- 54. Lareau, C.A.; Aryee, M.J. Hichipper: A preprocessing pipeline for calling DNA loops from HiChIP data. *Nat. Methods* **2018**, *15*, 155–156. [CrossRef] [PubMed]
- 55. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An information aesthetic for comparative genomics. *Genome Res.* 2009, 19, 1639–1645. [CrossRef] [PubMed]
- 56. Tang, Z.; Kang, B.; Li, C.; Chen, T.; Zhang, Z. GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **2019**, *47*, W556–W560. [CrossRef] [PubMed]