NAR Cancer, 2024, **6**, zcae020
https://doi.org/10.1093/narcan/zcae020
Advance access publication date: 8 May 2024
Cancer Gene Regulation, Chromatin, and Epigenetics

Integrated enhancer regulatory network by enhancer–promoter looping in gastric cancer

Tianhui Zhu[1], Atsushi Okabe[1,2], Genki Usui[1,3], Ryoji Fujiki[1], Daichi Komiyama[1], Kie Kyon Huang[4], Motoaki Seki[1], Masaki Fukuyo[1], Hiroyuki Abe[3], Meng Ning[1], Tomoka Okada[1], Mizuki Minami[1], Makoto Matsumoto[1], Qin Fan[1], Bahityar Rahmutulla[1], Takayuki Hoshii[1], Patrick Tan[4,5,6], Teppei Morikawa[7], Tetsuo Ushiku[3] and Atsushi Kaneda[1,2,*]

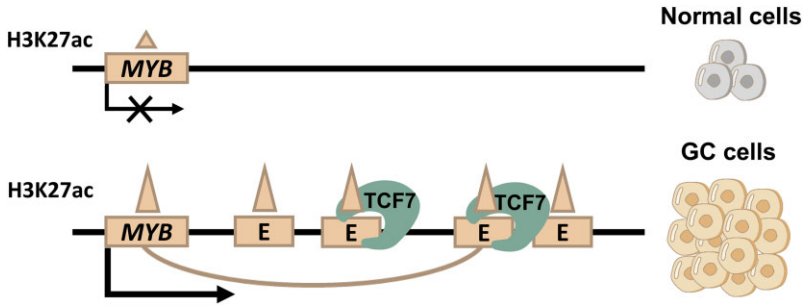
[1]Department of Molecular Oncology, Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan
[2]Health and Disease Omics Center, Chiba University, Chiba 260-8670, Japan
[3]Department of Pathology, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan
[4]Program in Cancer and Stem Cell Biology, Duke–NUS Medical School, Singapore 169857, Singapore
[5]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore 138632, Singapore
[6]Cancer Science Institute of Singapore, National University of Singapore, Singapore 117599, Singapore
[7]Department of Diagnostic Pathology, NTT Medical Center Tokyo, Tokyo 141-8625, Japan

[*]To whom correspondence should be addressed. Tel: +81 043 226 2039; Fax: +81 043 226 2039; Email: kaneda@chiba-u.jp

Abstract

Enhancer *cis*-regulatory elements play critical roles in gene regulation at many stages of cell growth. Enhancers in cancer cells also regulate the transcription of oncogenes. In this study, we performed a comprehensive analysis of long-range chromatin interactions, histone modifications, chromatin accessibility and expression in two gastric cancer (GC) cell lines compared to normal gastric epithelial cells. We found that GC-specific enhancers marked by histone modifications can activate a population of genes, including some oncogenes, by interacting with their proximal promoters. In addition, motif analysis of enhancer–promoter interacting enhancers showed that GC-specific transcription factors are enriched. Among them, we found that *MYB* is crucial for GC cell growth and activated by the enhancer with an enhancer–promoter loop and *TCF7* upregulation. Clinical GC samples showed epigenetic activation of enhancers at the *MYB* locus and significant upregulation of *TCF7* and *MYB*, regardless of molecular GC subtype and clinicopathological factors. Single-cell RNA sequencing of gastric mucosa with intestinal metaplasia showed high expression of *TCF7* and *MYB* in intestinal stem cells. When we inactivated the loop-forming enhancer at the *MYB* locus using CRISPR interference (dCas9-KRAB), GC cell growth was significantly inhibited. In conclusion, we identified *MYB* as an oncogene activated by a loop-forming enhancer and contributing to GC cell growth.

Graphical abstract



Introduction

cis-Regulatory elements (CREs) occupy a significant portion of the human genome and contain immeasurable and distinctive information about cell types through the regulation of gene transcriptomes (1). Aberrant gene transcription often occurs in cancer, leading to tumorigenesis by upregulating oncogenes or silencing tumor suppressor genes and their associated pathways (2). In CREs, enhancers and promoters promote the transcriptional machinery for gene expression (1). Histone modifications, such as H3K4me1 and H3K4me3, regularly occur in the enhancer and promoter regions, and H3K27ac indicates the active status of the enhancer and promoter. These epigenetic codes provide a path for discovering the connections between CREs and transcriptomes (3). One of the mechanisms by which genes are activated by enhancers and promoters is *de novo* enhancer–promoter (E–P) loop

Received: December 12, 2023. **Revised:** April 7, 2024. **Editorial Decision:** April 22, 2024. **Accepted:** April 25, 2024

formation (4). Several vital factors, including transcription factors (TFs) and critical transcriptional enzymes, are involved in this long-range interaction during different gene transcription processes.

In recent years, the structural aberrations of chromatin in cancer have been explored through the development of chromosome conformation capture techniques. Colorectal cancer cells revealed different open (A) and closed (B) chromatin compartments compared to normal cells using Hi-C, HiChIP and image data. These alterations are accompanied by changes in DNA methylation and histone modification and are associated with gene repression and oncogenic pathways, such as epithelial-to-mesenchymal transition (EMT), invasion and Wnt signaling (5). For a smaller unit, topologically associated domain boundary alterations regulated by CTCF on the chromatin structure in T-cell acute lymphoblastic leukemia can lead to E–P looping and higher *MYC* expression (6). A study of the 3D genome alterations in acute myeloid leukemia confirmed the activation role of enhancer hijacking in gene transcription (7).

Gastric cancer (GC) was the fifth leading cause of cancer in 2020 and causes many deaths every year (8,9). Mutations in genes such as *PIK3CA*, *TP53*, *KRAS*, *ARID1A* and *CDH1* are enriched in GC patients (10). These characteristics may provide information on the GC subtypes and indicate the precise treatment of patients with the disease (8,11,12). Promoter regions were somatically altered in GC clinical samples. Cancer-associated promoters exhibited enrichment of the *SUZ12* and *EZH2* (two subunits of the polycomb complex PRC2) TF binding sites (13). The TFs *KLF5*, *GATA4* and *GATA6* have been found to occupy promoter regions in a genome-wide manner. They participate in GC development by cooperatively regulating GC cell proliferation, colony formation and oncogene *MYC* expression (14). Genome-wide enhancer alterations also occur in GC. Activated and super-enhancers are associated with elevated gene expression and long-range chromatin interactions. Super-enhancers are involved in cancer development, such as signal transduction, cell death and cell proliferation (15). Genome-wide enhancer alterations regulated by *TEAD1* were enriched in biological adhesion, locomotion and cell migration (16). All of the above studies indicate the importance of enhancer and promoter dysregulation and E–P loop structure aberrations in oncogene activation, tumorigenic pathways and cancer development (4,17–19). However, aberrant E–P interactions in GC have not yet been fully investigated.

To clarify the role of the E–P loop in gene regulation and its contribution to gastric tumorigenesis, we utilized chromatin immunoprecipitation (ChIP) followed by sequencing (ChIP-seq), HiChIP, assay for transposase-accessible chromatin using sequencing (ATAC-seq) and RNA sequencing (RNA-seq), along with a comprehensive, integrated genome-wide analysis of histone modifications, chromatin interactions, chromatin states and gene expression in GC and normal cells. We identified a set of enhancer-interacting genes in GC, extracted several GC-specific TFs that regulate gene transcription via E–P loop formation and found an oncogene, *MYB*, associated with GC cell growth. The E–P loop activation mechanism was validated by the CRISPR interference (CRISPRi; dCas9-KRAB) system in GC cell lines, where repressive epigenetic changes in the *MYB* target enhancer-induced H3K27ac level decreased in the target region and *MYB* downregulation inhibited GC cellular growth.

Materials and methods

Cell culture

The human GC cell line SNU719 cells (Korean Cell Line Bank) were cultured in RPMI 1640 medium (Fujifilm, #189-02025) supplemented with 10% fetal bovine serum (FBS; Corning, Ref# 35-079-CV) and penicillin/streptomycin (Sigma–Aldrich, #P4333). YCC10 cells (Yonsei Cancer Center) were cultured in minimum essential media (Eagle's minimum essential medium; Fujifilm, #051-07615) supplemented with 10% FBS, penicillin/streptomycin and 1% non-essential amino acid solution. GES1 (Beijing Institute for Cancer Research) is a normal fetal gastric epithelial cell line immortalized with SV40 cultured in RPMI 1640 supplemented with 10% FBS and penicillin/streptomycin. HEK293T cells (ATCC) were cultured in Dulbecco's modified Eagle medium (Fujifilm, #044-29765) supplemented with 10% FBS and penicillin/streptomycin.

ChIP for histone modification

Two million cells were cross-linked with 1% formaldehyde and quenched with 0.2 M glycine. The fixed cells were sonicated by Picoruptor (Diagenode) in 0.25% sodium dodecyl sulfate (SDS) sonication buffer supplemented with 1% protease inhibitor cocktail (Roche, Ref# 05056489001). Twenty microliters of Protein A Sepharose beads (Cytiva, #17528001) was incubated with 2 μl histone modification antibody (H3K4me3, Active Motif, #39159; H3K4me1, CST, #5326S; H3K27ac, Active Motif, #91193) for 6 h at 4°C. The sonicated chromatin fragments were mixed with antibody-bonded Protein A Sepharose beads overnight at 4°C. Beads with immunoprecipitated DNA were washed three times with ChIP dilution buffer, low-salt wash buffer, high-salt wash buffer and TE buffer, eluted with 100 μl elution buffer (Tris–HCl, pH 8.0). De-cross-linking of DNA was performed at 65°C with 3 μl Proteinase K incubation overnight. After removing the beads, DNA fragments were purified by QIAquick PCR Purification Kit (#28106).

ChIP for TCF7

Ten million cells were cross-linked with 1% formaldehyde and quenched with 0.2 M glycine. The fixed cells were sonicated by Picoruptor (Diagenode) in 0.25% SDS sonication buffer supplemented with 1% protease inhibitor cocktail (Roche, Ref# 05056489001). Twenty microliters of Protein A Dynabeads (Invitrogen, #10002D) was incubated with 10 μl TCF7 (T-cell factor 7) antibody (CST, #2203) for 6 h at 4°C. The sonicated chromatin fragments were mixed with antibody-bonded beads overnight at 4°C. Beads with immunoprecipitated DNA were washed three times with ChIP dilution buffer, low-salt wash buffer, high-salt wash buffer and TE buffer, eluted with 100 μl elution buffer (Tris–HCl, pH 8.0). De-cross-linking of DNA was performed at 65°C with 3 μl Proteinase K incubation overnight. After removing the beads, DNA fragments were purified by QIAquick PCR Purification Kit (#28106).

Quantitative polymerase chain reaction for ChIP samples

Real-time quantitative polymerase chain reaction (qPCR) was performed using SYBR Green and CFX96 Touch Real-Time PCR (Bio-Rad Laboratories) to quantify the target regions of

the genomic DNA. All the primers are listed in Supplementary Table S1.

ChIP-seq

Libraries were prepared using the KAPA Hyper Prep Kit (KAPA Biosystems) and sequenced by Illumina NextSeq 500 or NovaSeq 6000, as previously described (20). Sequence reads were aligned to the hg38 reference genome by Bowtie2 (version 2.3.3.1) (21). Duplicates were removed by MarkDuplicates (Picard); peak calling was performed by findPeaks (HOMER 4.10). All ChIP-seq data were visualized by IGV (v2.11.9).

E–P interactome analysis by HiChIP

HiChIP for H3K27ac was performed based on the previously published study (22). Approximately 10 million cells were cross-linked with 1% of formaldehyde for 10 min, quenched with 0.2 M glycine and lysed with Hi-C lysis buffer (10 mmol/l Tris–HCl, pH 8.0, 10 mmol/l NaCl, 0.2% NP-40) supplemented with protease inhibitors to obtain cell nuclei. DNA was digested by MboI (NEB, #R0147) and 10× NEB Buffer 2 at 37°C with rotation. The digested overhangs were then filled with dCTP (Invitrogen, #10217016), dGTP (Invitrogen, #18254011), dTTP (Invitrogen, #18255018) and biotin-labeled dATP (Active Motif, #14139) and then ligated with T4 DNA ligase (Roche). After ligation, the nuclei were resuspended in 0.25% SDS sonication buffer (50 mmol/l Tris–HCl, pH 7.5, 10 mmol/l EDTA, 1% SDS) supplemented with protease inhibitors and sonicated by Covaris E220. The sonicated chromatin was then processed with H3K27ac ChIP. The purified and enriched DNA fragments were then captured by the streptavidin T-1 magnetic beads and washed, processed with Nextera XT DNA Library Preparation Kit (#FC-131-1096) preparation and sequenced by Illumina NovaSeq 6000.

As for HiChIP loop calling, the HiChIP sequencing reads were aligned to the hg38 reference genome by the HiC-Pro pipeline (23). Chromatin loops were identified by 10 and 25 kb resolution using the FitHiChIP (24). The loops supported by at least three pairs of valid reads (PET) and associated with a false discovery rate of <0.05 were considered as high-confidence loops for subsequent analyses. HiChIP loops were visualized by WashU genome browser (v54.0.4, https://epigenomegateway.wustl.edu/browser/). To find interacting promoter/enhancer in two distal elements, we used bedtools to intersect obtained overlapped regions between ChIP and HiChIP (25).

Cleavage under targets and tagmentation

Cleavage under targets and tagmentation (CUT&Tag) was conducted as previously described (26). A total of 1×10^5 cells were incubated in the antibody buffer with 10 μl pre-cleaned A-MyOne T1 beads (Thermo Fisher Scientific, #65601) and 1 μl H3K27ac antibody (Active Motif, #91193). After primary antibody binding, cells were incubated with 1 μl secondary antibody (Rabbit Anti-Mouse IgG H&L, Abcam, #ab6709) in Dig-wash buffer for 1 h at room temperature and washed using a Dig-wash buffer five times. Then, adapters were added to the fragments with the pA-Tn5 Kit (EpiCypher, #15-1017), and the mixture was washed twice using a Dig-wash buffer. The beads were resuspended in tagmentation buffer and incubated for 1 h at 37°C. After DNA extraction, library was constructed using KAPA HiFi HotStart ReadyMix (Roche, #07958927001), and sequenced by NovaSeq 6000.

Circularized chromosome conformation capture with sequencing

Circularized chromosome conformation capture with sequencing (4C-seq) was performed as previously described (20,27). Approximately 1×10^6 cells were cross-linked with 1% formaldehyde for 10 min, quenched with 0.2 M glycine, lysed with lysis buffer and digested with primary restriction enzyme DpnII (NEB, R0543). The digested DNA fragments were then subjected to ligation by using T4 DNA ligase (Roche, 10799009001). After digestion with the second restriction enzyme Csp6I (Thermo Fisher Scientific, ER0211) and circularization by T4 DNA ligase, the 4C template was amplified by a PCR reaction and purified with AMPure XP beads. The 4C library was sequenced by NextSeq 500 (Illumina).

Transcriptome analysis by RNA-seq

RNA was extracted using TRIzol reagent (Invitrogen, Ref# 15596018) and Direct-zol RNA MiniPrep Kit (Zymo Research Corporation, #R2052) according to the manufacturer's instructions. RNA-seq libraries were prepared using the TruSeq Stranded mRNA Sample Prep Kit (Illumina) and sequenced using Illumina NextSeq 500 or NovaSeq 6000. Gene expression levels were expressed as fragments per kilobase of exon per million mapped sequence reads (FPKM) by Cufflinks (https://github.com/cole-trapnell-lab/cufflinks) (28).

Clinical cancer patient data

Clinical cancer patient transcriptome data were obtained from The Cancer Genome Atlas (TCGA; https://portal.gdc.cancer.gov/, TCGA_STAD) (10).

Gene Ontology analysis

Enrichment analysis of Gene Ontology (GO) terms was performed by using g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) (29).

Chromatin accessibility analysis by ATAC-seq

ATAC-seq was performed as described previously (30). The cell was lysed by ATAC lysis buffer for 7 min and centrifuged at $500 \times g$ and 4°C for 10 min to obtain cell nuclei. Nuclei isolated from $\sim 5 \times 10^4$ cells were used for transposition reaction with transposase (Illumina Tagment DNA TDE1 Enzyme and Buffer Kits, #20034198). Purified transposed DNA was amplified in 50 μl reactions with KAPA HiFi HotStart ReadyMix (Roche, Ref# 07958927001) using primers with unique barcodes for the library construction. Sequence data analysis was performed similarly to ChIP-seq. Accessible chromatin was called with HOMER findPeaks with 300-bp nucleosome-free region.

Motif analysis

The motif enrichment of open, interacting enhancer was calculated using HOMER findMotifsGenome. The motif sequence was annotated by HOMER annotatePeaks-m.

CRISPRi-mediated epigenetic editing

GC cell lines expressing dCas9-KRAB were established according to (31). Plasmid (pLV-dCas9-KRAB-PGK-HygR, Addgene #83890) was packaged by lentivirus (Promega, FUGENE 6 transfection reagent, Ref# E2692) and introduced into GC cells. Guide RNAs were designed by CRISPick (https://portals.broadinstitute.org/gppx/crispick/public) (32,33) and then cloned in the backbone (pLV-U6-gRNA-UbC-eGFP-P2A-Bsr, Addgene #83925) by Gibson assembly. The sequences of constructed single-guide RNA (sgRNA) plasmids were confirmed by Sanger sequence, and the sgRNA sequences are listed in Supplementary Table S2. After lentivirus packaging, sgRNA was introduced into dCas9-expressed GC cell lines. To obtain a modest 30–50% efficiency, we used a suitable amount of virus for infection. Three days after infection, 8 μg/ml blasticidin was used to select sgRNA-infected cells for 7 days.

Competitive growth assay by FACS

Competitive growth assay was started after 4 days from sgRNA infection (defined by day 0) by measuring fluorescent protein expression using CytoFLEX flow cytometry (Beckman Coulter). The eGFP percentage in each group was measured every 7 days for a total of 28 days. Cell viability was calculated by comparing with the day 0 in each group.

Immunohistochemistry

Clinical GC tissue samples [$n = 49$, including 5 MSI (microsatellite instability) GC, 29 EBV(+) (Epstein–Barr virus-positive) GC and 15 other GC cases] were obtained from patients undergoing gastrectomy at the University of Tokyo Hospital. Normal gastric mucosa samples ($n = 22$) were obtained from healthy individuals undergoing biopsies in the health check-up program at NTT Medical Center Tokyo (34). Formalin-fixed, paraffin-embedded tissue blocks were cut into 4-μm-thick sections. Three serial sections were prepared for each tissue block and used for hematoxylin and eosin staining and immunohistochemistry (IHC) using c-MYB antibody (Abcam, Cat# ab45150). IHC was performed using BenchMark ULTRA automated staining system (Roche Diagnostics, Basel, Switzerland), and visualized with OptiView DAB IHC Detection Kit (760700; Roche) according to the manufacturer's instructions. Two independent experienced pathologists scored the staining as follows: 3 (strong), 2 (moderate), 1 (weak) and 0 (negative). The study design was approved by the Institutional Review Board at NTT Medical Center Tokyo (18-102), Chiba University Hospital (1209) and the University of Tokyo Hospital [G3521-(21)].

siRNA knockdown

Small interfering RNA (siRNA) was purchased from Thermo Fisher Scientific (MYB#1: s9109; #2: s9110; TCF7#1: s13877; #2: s13878; TCF7L2#1: s13880; #2: s13881; TCFL5#1: s21080; #2: s21081), and was transfected into GC cells by RNAiMAX (Thermo Fisher Scientific), according to the manufacturer's instructions. The siRNA effect was confirmed by reverse transcription qPCR (RT-qPCR) after 48 h transfection, and the primers for RT-qPCR are listed in Supplementary Table S1.

CCK-8 growth assay

Cell growth assay was performed from the day of siRNA transfection to day 8 in a 96-well plate. Cell viability was calculated by comparing the OD value (450 nm) detected each day to day 0. The OD value was detected 1.5 h after adding CCK-8 (Cell Counting Kit 8; Dojindo, #CK04).

Colony formation assay

For colony formation assay, 1×10^4 cells with and without CRISPRi editing were seeded into 6-cm plates with agarose. After 28-day incubation, plates were washed with phosphate-buffered saline (PBS) and stained by staining solution (1 mg/ml nitro blue tetrazolium in PBS) for 3 days.

Western blot

Cells were lysed by SDS loading buffer suspended in 2-mercaptoethanol for 5 min to obtain whole cell protein. The protein was quantified by the Bradford assay. The same amount of protein was loaded on the gel (Bio-Rad, 7.5%) and transferred to 0.45 μm polyvinylidene fluoride membrane. The membrane was blocked in 5% skim milk for 1 h, and incubated in the primary antibody (cMYB, Abcam, Cat# ab45150; ACTB, MBL, M177-3) at 4°C overnight. After being washed by TBST, the membrane was incubated in the secondary antibody (anti-mouse IgG, NA931V; anti-rabbit IgG, NA934V), and visualized using Amersham ECL Prime (Cytiva, RPN2232).

Statistical analysis

Error bars represent standard deviation. For statistical comparison, we performed Student's *t*-test using R software. Data with statistical significance ($^*P < 0.05$, $^{**}P < 0.01$) are shown in figures.

Results

Active enhancers displayed a specified histone modification landscape in GC

To identify the enhancer and promoter regions in normal and GC cell lines, we utilized the ChIP-seq of H3K4me1, H3K4me3 and H3K27ac, which are well-known histone modification markers of the enhancer and promoter regions (3). First, we identified the 'active enhancers' with H3K4me1(+), H3K4me3(–) and H3K27ac(+), as well as the 'active promoters' with H3K4me3(+) and H3K27ac(+) in each cell line (Supplementary Figure S1A). Using the heatmap, we compared the enhancer regions of the three cell types and found that they were highly specific in the normal and GC tissues. The regions with normal active enhancers showed strong signals for H3K4me1 and H3K27ac in GES1, but not in the two GC lines. Likewise, the GC active enhancer exhibited higher signals in the GC cells than in the normal cells, especially in H3K27ac. However, the histone modification markers H3K4me3 and H3K27ac were at least partly shared in the three cell lines at the active promoters. Enhanced aberrations and stable promoter patterns were also observed (15). This evidence led us to use different methods to extract the enhancers and promoters for further analysis.

For the enhancers, we extracted specific regions from each cell line. We named 8261 enhancers active in GES1 as 'normal-specific enhancers' and 6306 commonly active

enhancers in two GC lines as 'GC-specific enhancers' (Figure 1A; Supplementary Tables S3 and S4). A heatmap of specific enhancers showed obvious differentiation between the normal and GC cells (Figure 1B). Two representative genes, *CDH2* with a normal-specific enhancer and *EHF* with two GC-specific enhancers, were shown in the H3K4me1, H3K4me3 and H3K27ac ChIP-seq data (Figure 1C). *EHF* is a critical TF involved in GC tumorigenesis as demonstrated in our previous study (35). Regarding most of the promoters shared in the three cell lines, the active promoters in the normal and GC cells may have different patterns of E–P loop formation. To avoid losing some interactions between the common promoters and specific enhancers in the following steps, we extracted all the promoters in each cell line, including overlapping regions, to achieve a total of 12 267 active promoters in the three cell lines (Supplementary Figure S1B; Supplementary Table S5). Thus, we confirmed enhancer aberrations in GC and identified the active enhancers and promoters using ChIP-seq.

Genes with GC-specific active E–P loops showed higher expression in GC

To explore the chromatin interactome, we performed H3K27ac HiChIP, which detects interactions between the H3K27ac histone active markers (22). More than 10 000 'active interactions' were detected in the three cell lines using HiChIP (Supplementary Figure S2). Instead of calling specific loops in the GC cells, we preferred to identify loops formed by the GC enhancer aberrations. We applied the 'specific enhancer' (Figure 1A) and 'active promoter' (Supplementary Figure S1B) extracted by ChIP-seq and further classified those active loops into enhancer–enhancer (E–E), E–P and promoter–promoter (P–P) interactions. Around one-third of the active loops were E–P loops (GES1: 7260; SNU719: 3717; YCC10: 6193) (Figure 2A). We obtained all the enhancers with E–P loops in the GES1 and GC cells, named 'interacted enhancer'. By integrating interactome with histone modification data, the differences between the normal and GC cell lines can be visualized and compared. Here, we showed that the representative gene, *EHF*, gained E–P loops in the GC cells. In the *EHF* transcription starting site (TSS) region, H3K4me3 and H3K27ac signals were detected in the GC cells but not in the normal cells, and H3K4me1 and H3K27ac signals in enhancer regions. These active elements were confirmed to form several loops at the *EHF* locus (Figure 2B). To identify the genes with E–P loops in the GC lines, we annotated their corresponding promoters. In the GC cells, there are 2369 genes with an E–P loop, called the 'E–P interacted gene' (Supplementary Table S6). E–P interacting genes of GC cells were expressed at significantly higher levels in the GC cells compared to those in the normal cells (Figure 2C; Supplementary Table S7). Thus, we analyzed H3K27ac HiChIP using selected regulatory proximal and distal elements and found an association between E–P contact and high gene expression.

The E–P interacted gene *MYB* has higher expression in GC cells and clinical samples

According to previous studies, genes related to E–P loops may be enriched in some tumorigenic pathways (36). We obtained gene expression data of the cell lines by RNA-seq, and those of clinical tissues from the TCGA_STAD, and extracted highly expressed genes in GC cell lines compared with GES1 [average transcripts per million (TPM) of two cell lines, $P < 0.05$, fold change (FC) > 2, $n = 2480$; Supplementary Table S8] and clinical GC tissues ($P < 0.05$, FC > 2, TPM > 5, $n = 3466$; Supplementary Table S9). Among the 2369 GC E–P interacted genes, 188 were highly expressed in both GC cell lines and clinical GC tissues (Figure 3A). Furthermore, GO analysis of these 188 genes revealed several tumorigenesis-related terms, such as cell migration and cell motility (Figure 3B; Supplementary Table S10). In addition, we wanted to determine whether oncogenes or those that function in tumorigenesis are included. We utilized 'gene effect score' in SNU719 from the DepMap public database, a dependency map through CRISPR screening (https://depmap.org/portal/). A lower gene effect score indicates that SNU719 cell growth is more dependent on the gene. Among these 188 genes, the top 20 genes showing critical roles in SNU719 cellular growth (ranging from −2.33 to −0.35) included *CDC7*, *CCND1*, *ATP2A2*, *POLR1B*, etc. (Figure 3C; Supplementary Table S11). These results indicated that both oncogenic pathways and cell growth-related genes are enriched in the E–P interacting genes.

When we compared the FCs of expression for the top 20 genes, *MYB* showed the highest FC (Figure 3D). To confirm *MYB* expression in clinical tissues, we compared the expression in normal and GC tissues by using RNA-seq data from TCGA_STAD, where *MYB* showed significantly higher expression in all molecular subtypes of GC, especially in MSI GC and EBV GC (Figure 3E). Additionally, we performed IHC for clinical GC tissue samples of our cohort ($n = 49$) and normal gastric mucosae from healthy individuals ($n = 22$). Nearly half of the GC cases showed high *MYB* expression (IHC score 2 or 3), and the IHC staining score was consistent with the TCGA_STAD data (Figure 3F; Supplementary Table S12). We further classified TCGA patients of each GC subtype into two groups by age (<70 versus ≥ 70), gender (female versus male), stage (I/II versus III/IV) and TNM factors (I/II versus III/IV for T; 0 versus ≥ 1 for N; 0 versus ≥ 1 for M). There was no significant difference detected in these comparisons, suggesting that *MYB* expression could be upregulated in any subtypes of GC, regardless of these clinicopathological factors (Supplementary Figure S3A). IHC results showed similar distribution of GC cases with high and low IHC scores regardless of these clinicopathological factors, which were consistent with TCGA results (Supplementary Figure S3B).

ChIP and HiChIP data at the *MYB* locus were visualized, along with a significant interaction between the active enhancer and promoter in GC cells (Figure 3G). To confirm whether enhancers of *MYB* are activated in clinical GC tissues as well, we analyzed published NanoChIP-seq data for seven GC patients, including GS, EBV(+) and CIN subtypes (GSE76153 and GSE75898) (13,15). When H3K4me1 signals of normal and GC tissues were compared, GC tissue showed significantly higher tag count (Supplementary Figure S4).

Taken together, we found that E–P interacting genes in GC are associated with oncogenic pathways and that *MYB* was highly upregulated in GC cell lines and clinical GC tissue samples, regardless of GC subtypes or clinicopathological factors.

MYB regulates GC cellular growth

To assess how *MYB* contributes to GC tumorigenesis, we knocked down *MYB* in two GC cell lines by siRNA.

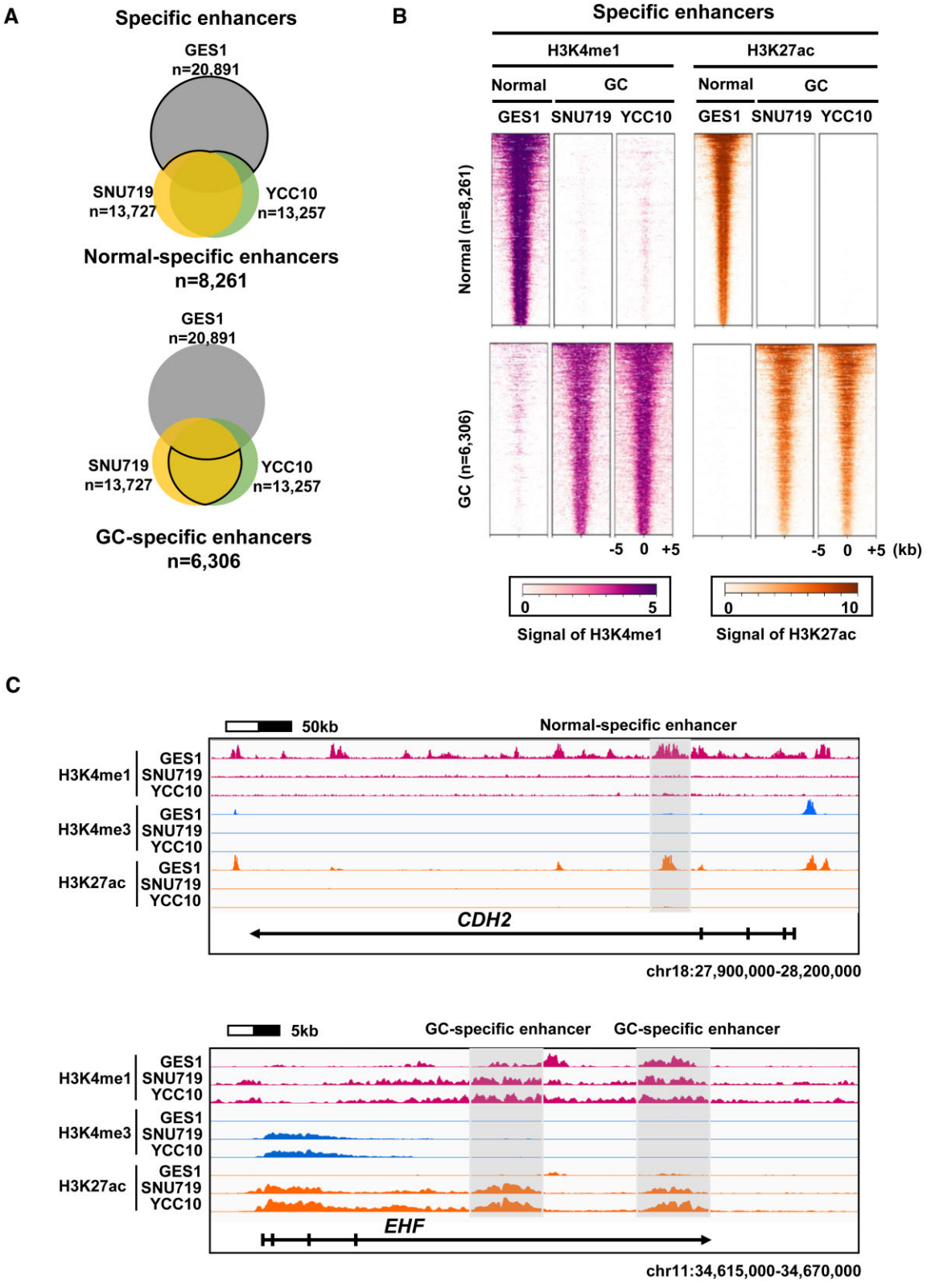


Figure 1. Active enhancers displayed a specified histone modification landscape in GC. (**A**) Black outline area presenting normal/GC-specific enhancer regions. (**B**) Heatmap of normal/GC-specific enhancers: normal-specific enhancer, $n = 8261$; GC-specific enhancer, $n = 6306$. (**C**) Representative ChIP-seq results at the *CDH2* and *EHF* loci. Active promoters and normal/GC-specific enhancer are shown.

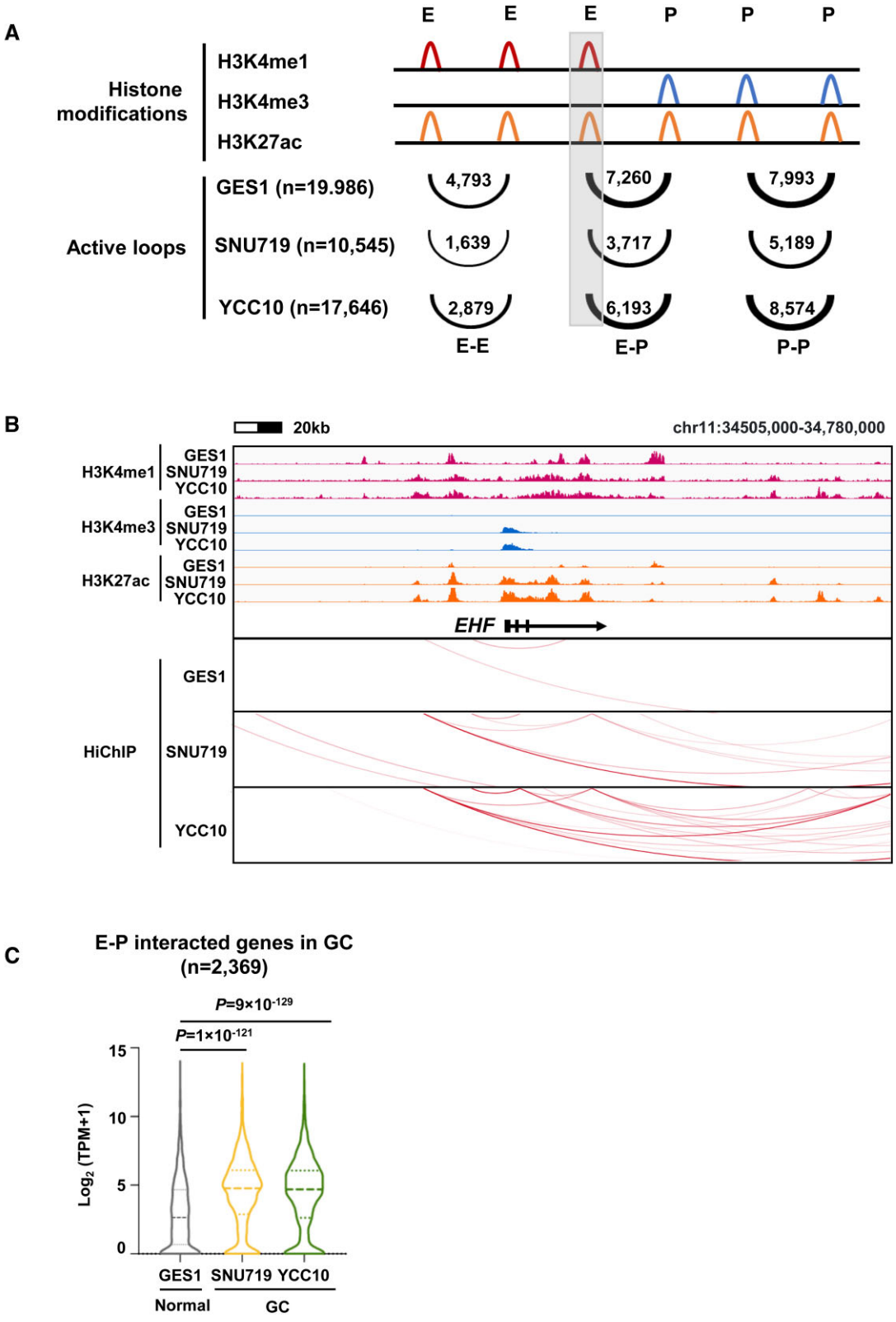


Figure 2. Genes with GC-specific active E–P loops showed higher expression in GC. (**A**) The feature of active loops in normal and GC cells. E, enhancer; P, promoter. Active loops are classified into three types: E–E, E–P and P–P loops. The number of E–P loop interactions identified is as follows: GES1, $n = 7260$; SNU719, $n = 3717$; YCC10, $n = 6193$. (**B**) A representative ChIP-seq and H3K27ac HiChIP result at *EHF* locus. *EHF* was from the list of 'genes with active E–P interaction' of the GC cell lines. (**C**) Expression levels of E–P interacted genes in GC. GC enhancer-interacted genes ($n = 2369$) show higher expression in GC cells compared with GES1.

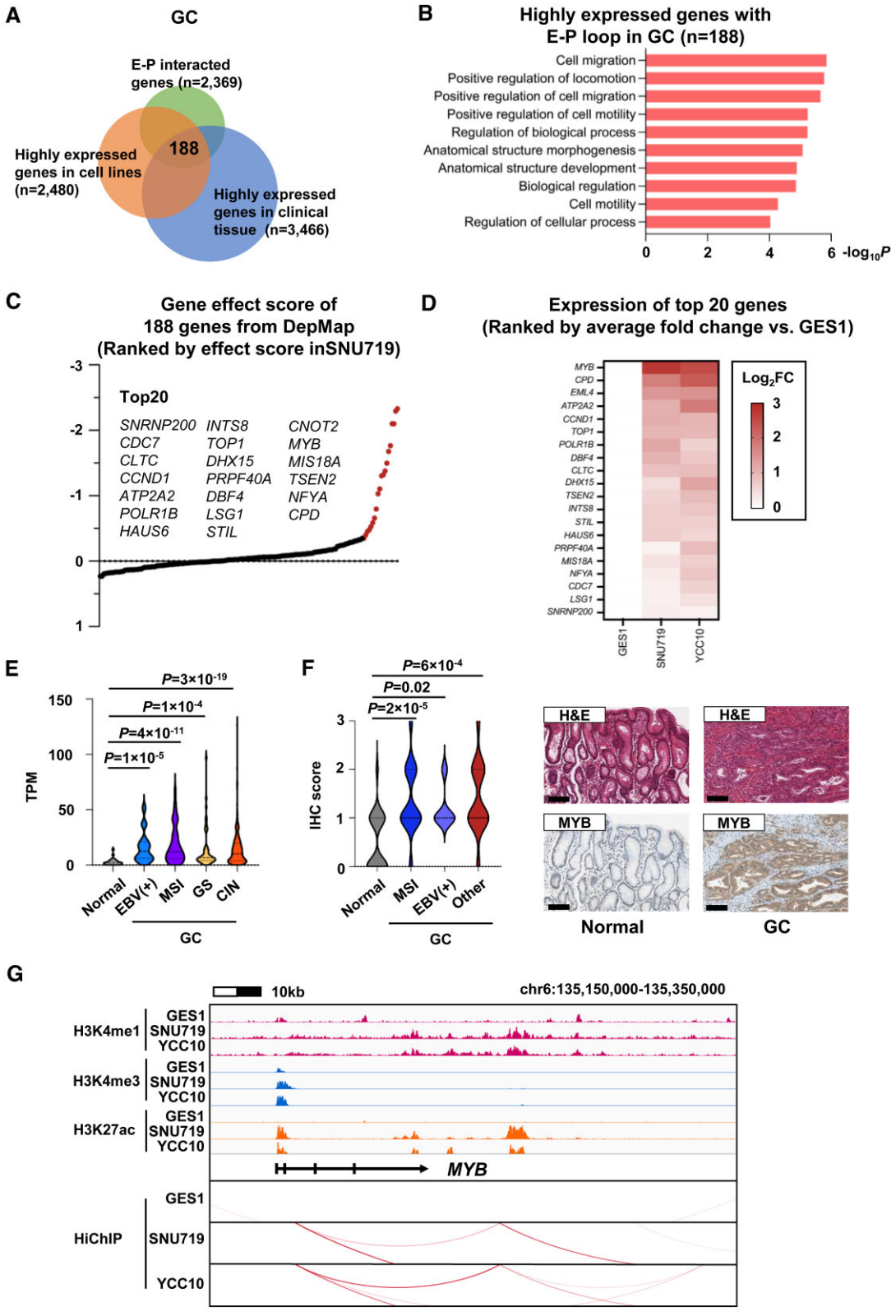


Figure 3. E–P interacting gene *MYB* has higher expression in GC cells and clinical samples. (**A**) One hundred eighty-eight genes were overlapped among E–P interacted genes detected by HiChIP ($n = 2369$), highly expressed genes in GC cell lines detected by RNA-seq ($n = 2480$, $P < 0.05$, FC > 2) and highly expressed genes in GC tissues detected by TCGA data of RNA-seq ($n = 3466$, $P < 0.05$, FC > 2). (**B**) Enrichment of GO terms for the 188 extracted genes by g:Profiler (https://biit.cs.ut.ee/gprofiler/gost). (**C**) DepMap score of the 188 genes analyzed for SNU719 [CRISPR_(DepMap_Public_24Q1)]. Top 20 genes are listed, and marked by red. Gene names are shown by the ascending order of gene effect score and descending order of gene dependency. (**D**) Expression levels of the top 20 genes in GC cell lines compared to GES1. Genes are listed from the highest to the lowest FC. (**E**) *MYB* expression in normal and GC tissues. TCGA_STAD data were analyzed. Normal tissues, $n = 29$; EBV(+) GC, $n = 30$; MSI GC, $n = 73$; CIN (chromosomal instability) GC, $n = 223$; GS (genomically stable) GC, $n = 50$. (**F**) IHC of *MYB* in clinical sample of gastric mucosae ($n = 22$) and GC tissues. EBV(+) GC, $n = 29$; MSI GC, $n = 5$; other GC, $n = 15$. Scale bar: 100 μm. (**G**) Representative ChIP-seq and HiChIP results at *MYB* locus.

After siRNA transfection, messenger RNA (mRNA) expression and protein expression of MYB were both downregulated, as shown by RT-qPCR and western blotting. The results of growth assay in SNU719 were consistent with the public data on DepMap; repression of *MYB* leads to an inhibitory effect on the cell growth of GC. Similar results were obtained in another GC cell line, YCC10 (Figure 4A and B), confirming that *MYB* is a critical gene in GC cell growth. *MYB* is a TF that plays an important role and is overexpressed in many types of cancer (37–39). However, the roles of *MYB* in GC remain unclear. To explore downstream effect of *MYB* in GC cells, we performed RNA-seq by using *MYB* knockdown (KD) cells in both GC cell lines. GO enrichment analysis showed that the commonly downregulated genes (FPKM, $P < 0.05$, FC < 0.8) tended to regulate cell cycle process. The commonly upregulated genes (FPKM, $P < 0.05$, FC > 1.2) showed a structure and cell adhesion pattern, including the well-known cell adhesion genes *CLDN1* and *ITGB* family members (Figure 4C and D; Supplementary Tables S13 and S14). These results indicate the importance of *MYB* in GC and suggest that it may regulate GC cell growth through the cell cycle.

Colocalization and correlation between TFs and the loop-forming enhancers

Similar motifs are usually enriched in a group of CREs, and motif-associated TFs bind to accessible elements to participate in the transcriptome of their downstream genes (40–43). We hypothesized that enhancers with E–P loops in GC might have a similar tendency; there would be one or more motifs accumulated in those interacting enhancers. These motifs may indicate critical factors associated with loop-forming enhancers in GC. In terms of this, we planned to find motifs enriched at the 'interacted enhancer' regions (Figure 2A). We performed ATAC-seq and obtained 300-bp open-chromatin regions at the center of the interacting enhancers. After obtaining the accessible chromatin regions by ATAC (GES1, $n = 29\ 341$; SNU719, $n = 34\ 489$; YCC10, $n = 33\ 365$), we combined the regions with interacting enhancer regions from HiChIP (GES1, $n = 4588$; SNU719, $n = 2643$; YCC10, $n = 3501$; Supplementary Tables S15–S17) and found 3381, 2159 and 3085 accessible enhancers with E–P loops in the normal and GC cells, respectively (Supplementary Figure S5A; Supplementary Tables S18–S20). Chromatin accessibility was similar at the genome-wide level in the normal and GC cells but increased at all active enhancers and E–P loop enhancers (Supplementary Figure S5B). Similar to the ChIP-seq data, the promoter of *CDH2* and its interacting enhancer only showed an open status in the normal GES1 cells, and the ATAC signal specifically occurred in two GC lines at the *EHF* promoter and enhancers (Figure 5A).

To find the TF that is associated with loop-forming enhancers during gene transcription, we performed motif analysis by accessible interacting enhancers. Enhancers with E–P interactions in normal and GC cells exhibited different occupancies. Except for common BATF/ATF3 motifs, TEAD, RUNX and ATF motifs were enriched in the normal cells (blue), whereas TCF, FOX and KLF were the top motifs enriched in the GC cell lines (red) (Figure 5B and C). These motif analysis results indicated that E–P loop development may associate with different TFs in GC cells than in normal cells, as expected. We further checked the expression of all possible TFs predicted from each motif (TCF, FOX or KLF), and compared their expression in the cells and clinical tissues. *TCF7*, *TCF7L2*, *TCFL5*, *KLF5*, *FOXA1* and *FOXC1* were found to be highly expressed at mRNA levels in GC (Figure 5D and E). Collectively, according to the motif analysis of open enhancers with an E–P loop, GC showed different TF occupancy. The TCF, FOX and KLF TF families were the most frequent motifs in GC.

TCF7 is a critical TF that regulates MYB transcription in GC

To validate the motif enrichment results, we investigated the TF involved in *MYB* activation and the transcriptional mechanisms of *MYB*. Interestingly, when we searched for the predicted TCF, FOX and KLF motif sequences at the *MYB* locus, only the TCF motif was found in the enhancers (Figure 6A). Since the predicted results contained several family members of TCF, we referred to the expression in the cell lines and clinical data and selected overexpressed *TCF7*, *TCF7L2* and *TCFL5* to confirm the predicted results. We knocked down these three TFs in the GC cells to check *MYB* expression. RT-qPCR results showed that *TCF7* KD effectively led to lower *MYB* expression, but not the other two (Figure 6B and C; Supplementary Figure S6A–C). These results suggest that TCF7 is a critical regulator in *MYB* transcription processes. Moreover, the GC cells with decreased *TCF7* expression exhibited lower cell growth (Figure 6D), indicating that TCF7 not only regulates gene expression but also plays an important role in GC cell growth.

Next, to confirm that TCF7 binds to the *MYB* enhancer, we performed TCF7 ChIP-seq. Most of the TCF7 binding regions showed open status in GC (Supplementary Figure S6D, by referring to ATAC data), indicating their accessibilities and binding activities in GC. Consistent with the RT-qPCR results, TCF7 binding was enriched in the *MYB* enhancer of the GC cells but not of the normal cells (Figure 6E). Likewise, higher expression of *TCF7* is shown in all the four subtypes of GC regardless of age, gender, stage or TNM factors (Figure 6F; Supplementary Figure S7). To explore the expression of *TCF7* and *MYB* during the clinical GC development, we checked the single-cell RNA-seq of the antrum with intestinal metaplasia (IM) obtained from the patients (44). IM is known as the condition of pre-malignant gastric mucosa, and to be associated with higher GC risk. We assessed gene expression levels in four gastric and four intestinal cell clusters, and found that both *TCF7* and *MYB* were highly expressed in intestinal stem cells (Supplementary Figure S8), suggesting that their upregulation might perhaps occur at an early step of GC tumorigenesis, e.g. IM condition.

To elucidate the function of TCF7 in the regulation of *MYB* enhancers, we analyzed chromatin interactions from *MYB* promoters by 4C sequence and H3K27ac status by CUT&Tag after *TCF7* KD. Interaction frequency from *MYB* promoter did not change after *TCF7* KD, and H3K27ac was not changed either (Supplementary Figure S9). These results suggested that *TCF7* might not have pioneering function in E–P loop formation and H3K27ac induction.

Suppression of interacted enhancer by CRISPRi repressed gene expression and cellular growth

At the *MYB* locus, there are four active enhancers (E1–E4) marked by H3K4me1 and H3K27ac downstream of the TSS. By ATAC-seq and TCF7 ChIP-seq, we observed that three

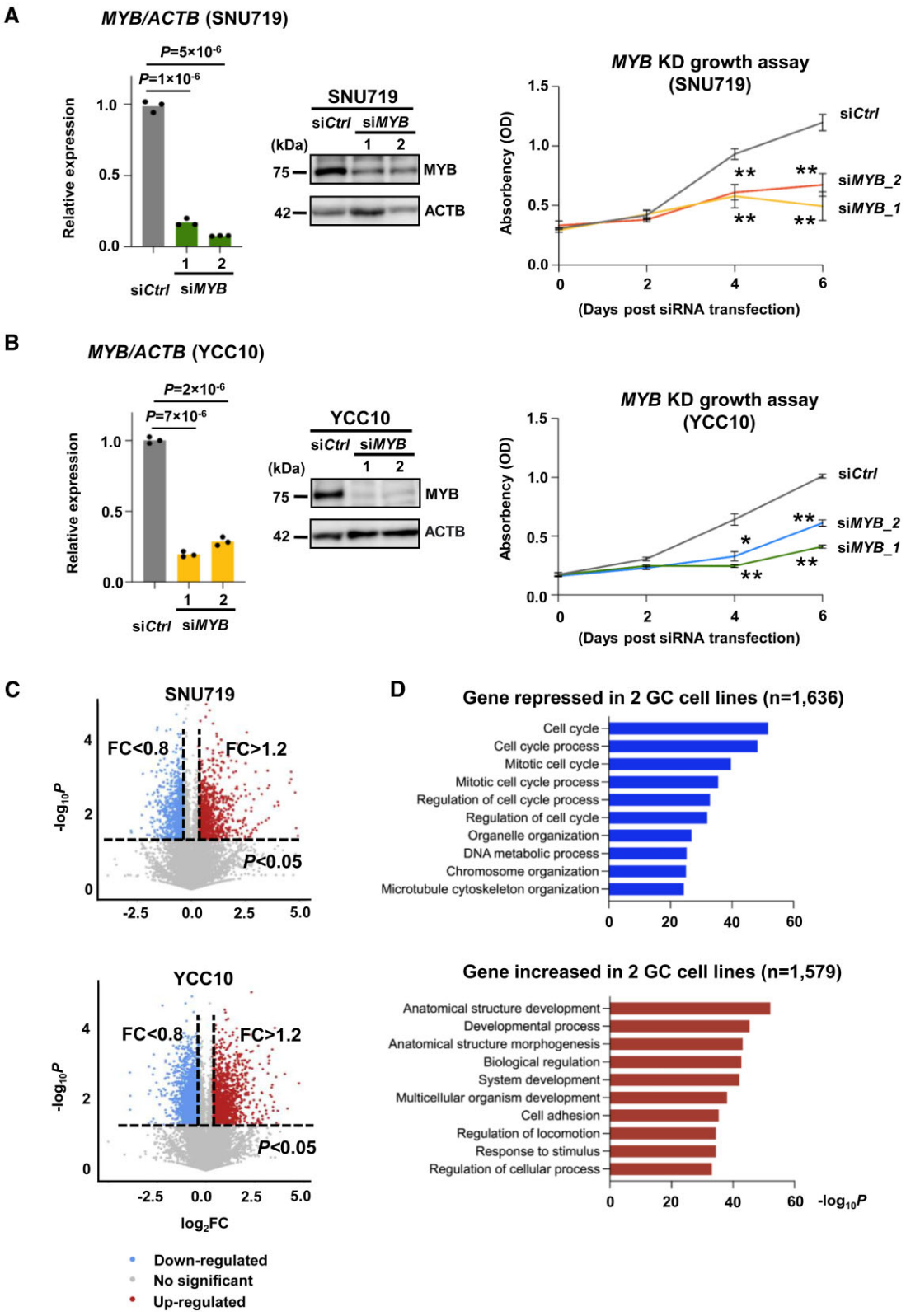


Figure 4. *MYB* regulates GC cellular growth. (**A**, **B**) RT-qPCR and western blotting results reveal that *MYB* was knocked down by two siRNAs at mRNA and protein levels. Cellular growth assay using CCK-8 in both cell lines indicates significant decrease of cellular growth after *MYB* KD. (**C**) RNA-seq after *MYB* KD in GC cells. Downregulated genes ($P < 0.05$, FC < 0.8) and upregulated genes ($P < 0.05$, FC > 1.2) are indicated. (**D**) GO analysis of the downregulated genes and upregulated genes in *MYB* KD in GC cell lines.

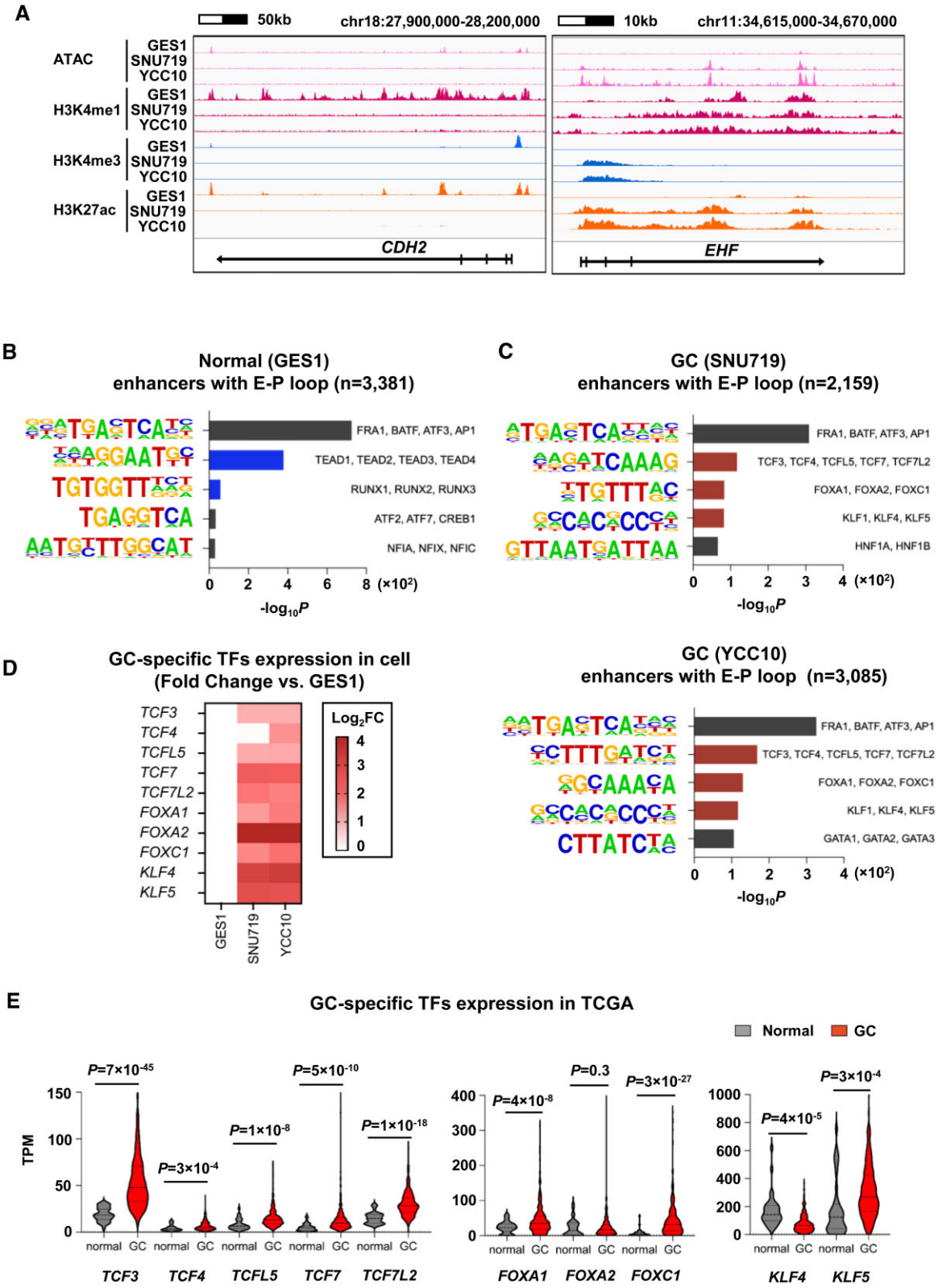


Figure 5. Colocalization of TFs and open chromatin in the loop-forming enhancers. (**A**) Representative views of open-chromatin regions (ATAC-seq) and active histone marks (ChIP-seq). *CDH2* is active in the normal GES1 cells, and *EHF* is active in two GC lines. (**B**, **C**) The top 5 motifs and predicted TFs in normal-specific and GC-specific enhancers. Motifs e.g. TEAD and RUNX (B) and those e.g. TCF, FOX, and KLF (C) are specifically observed in normal-specific and GC-specific enhancers with E–P loops, respectively. (**D**) Expression of GC-specific TFs in three cell lines (RNA-seq). (**E**) Expression of GC-specific TFs in clinical tissues (TCGA_STAD).

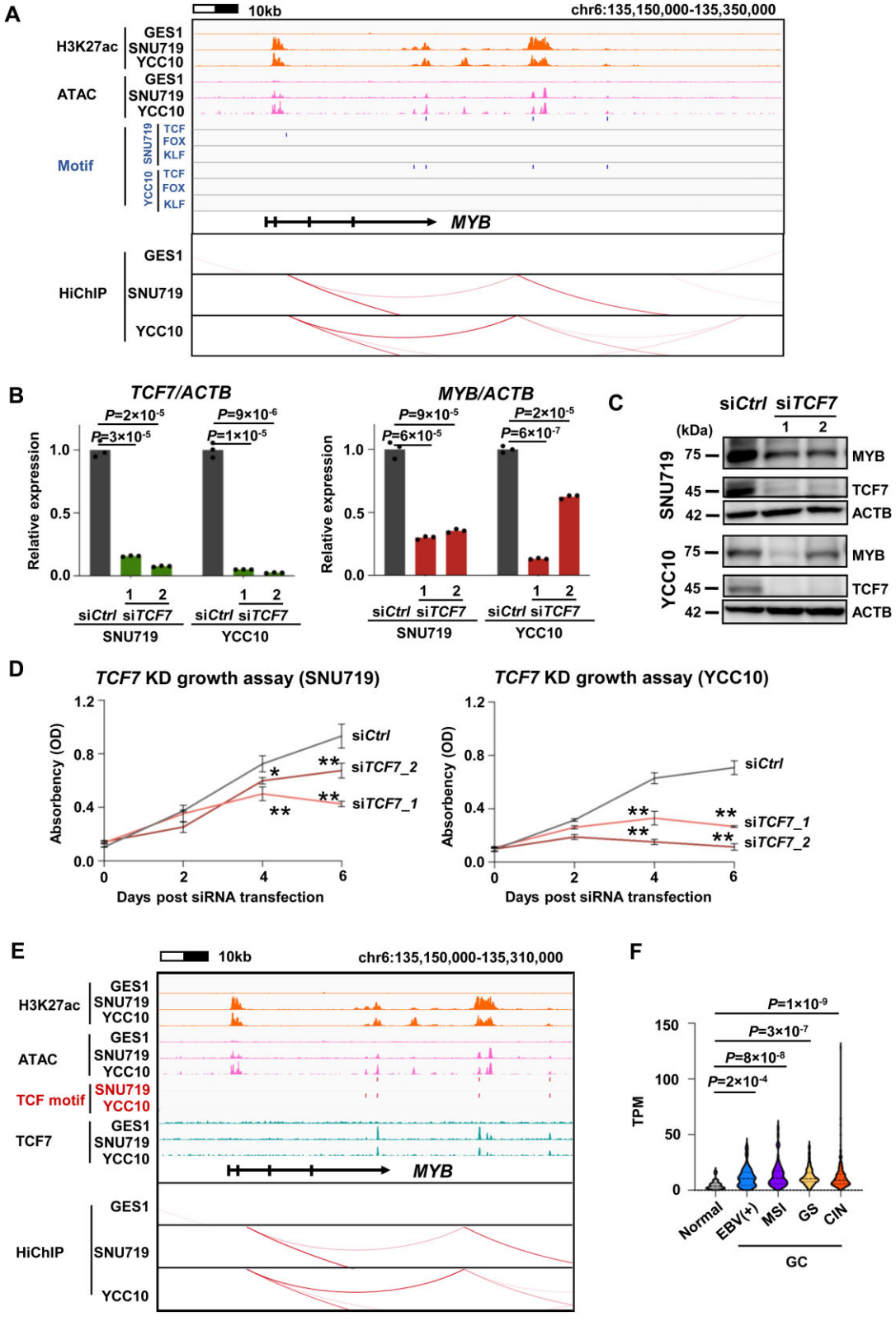


Figure 6. TCF7 is a critical TF that regulates *MYB* transcription in GC. (**A**) Three GC motifs at *MYB* locus. Only TCF motif sequence has been found nearby *MYB*. (**B**) *TCF7* KD in GC cell lines. Expressions of *TCF7* and *MYB* are downregulated after siRNA KD of *TCF7*. (**C**) Western blotting reveals decrease of TCF7 and MYB at protein levels, after *TCF7* KD. (**D**) Cellular growth assay for two GC cell lines after *TCF7* KD. (**E**) TCF7 ChIP-seq signal accumulated in interacting enhancers of *MYB*. (**F**) Expression of *TCF7* in four GC subtypes compared with normal samples (TCGA_STAD).

of the enhancers produced both signals (E1, E3 and E4). The HiChIP loops indicated that only E3 interacted with the *MYB* promoter. If *MYB* is activated by this loop-forming enhancer, and we repress the interacting enhancer, *MYB* expression should also be repressed. As the interacting enhancers were extracted by active histone marks, we introduced the CRISPRi system to validate our hypothesis with the transcriptional activation mechanism. The fusion of dCas9 with the transcription repressor KRAB can act as an epigenetic repressor to induce the heterochromatin mark H3K9me3 and suppress downstream enhancers of *MYB* (31). We introduced the dCas9-KRAB fused protein into our GC cell lines and designed 12 sgRNAs on the target enhancers E1–E4, as well as control sgRNAs between E2 and E3 (Control 1) and at a different chromosome (Control 2) (Figure 7A). As expected, epigenetic editing successfully induced H3K9me3 in the four enhancers downstream of *MYB*, leading to the decrease of H3K27ac signal in corresponding regions. H3K27ac ChIP-qPCR results showed that the sgRNAs against E3 repressed not only E3 H3K27ac but also other regulatory elements of *MYB* (TSS, E1 and E2). However, the other sgRNAs exhibited only an inactivation effect on their targets (Figure 7B). RT-qPCR results after sgRNA treatment showed that *MYB* expression was markedly downregulated when targeting E3, but not markedly when targeting other enhancers, or did not change when using control sgRNAs (Figure 7C).

When we conducted H3K27ac ChIP-seq after sgRNA treatment, H3K27ac signals were shown to be decreased at the *MYB* locus when targeting E3, but not targeting Control 1 region (Supplementary Figure S10A). Decrease of H3K27ac, however, was not observed at neighboring regions, such as *ALDH8A1*, *HBS1L* and *AHI1* loci. The expression of neighboring genes was not decreased either (Supplementary Figure S10B and C).

To elucidate the effect on cellular growth, we performed competitive growth assay after CRISPRi treatment targeting E1–E4. Cell viability significantly decreased on days 14, 21 and 28 when using sgRNAs targeting E3, confirming the significant effect on GC cell growth by inactivating the interacting enhancer, in SNU719 (Figure 7D) and YCC10 (Supplementary Figure S10D and E). Decrease of *MYB* expression after E3 inactivation by protein level was also confirmed by western blotting (Figure 7E; Supplementary Figures S10F and S11A). In addition, we conducted colony formation assay using SNU719 cells after CRISPRi treatment targeting E3 (sg9), and observed decrease of colonies after E3 inactivation (Supplementary Figure S11B and C). These results showed that epigenetic repression of *MYB* enhancer by CRISPRi technique is enough to lead to decrease of gene expression and cellular growth, similarly to siRNA KD of *MYB* in the GC cells (Figure 4A and B).

Furthermore, to elucidate the effect on E–P loop after epigenetic repression of enhancers by CRISPRi, we performed 4C-seq using control sgRNA and sgRNA targeting E3. Interaction frequency from *MYB* promoter was not changed after E3 repression (Supplementary Figure S12A), suggesting that E3 repression may downregulate the activity of chromatin, but not enough to break the loop structure between enhancer and promoter. In addition, we further explored the synergistic effect of E3 repression by CRISPRi and *TCF7* KD on *MYB* expression. The significant combination effect of E3 repression and *TCF7* KD was observed (Supplementary Figure S12B).

Discussion

In this study, we conducted a comprehensive analysis of the long-range interactions in GC. We found that *MYB* was aberrantly activated by such interactions and contributed to gastric tumorigenesis. GC-specific motifs occupy highly accessible distal enhancers. TCF7 is a critical TF that regulates *MYB* activation in GC.

Recently, many studies have been conducted on chromatin remodeling during cell development, differentiation and cancer (42,45–47). These aberrant chromatin changes are typically accompanied by enhancer and transcriptomic plasticity (36,48–51). Chromatin structure is regulated by different factors. CTCF also contributes to the 3D chromatin structure of embryonic stem cells (50) and leukemia by activating the oncogene *MYC* (6,7). In addition to common factors such as CTCF and cohesion, specific TFs or long noncoding RNAs (lncRNAs) also play critical roles in different cell types or stages. YY1 plays an important role in E–P looping in mouse embryonic cells (52); KLF5 has been shown to exhibit the ability to bind lineage-specific regulatory elements and induce gene expression in human epithelial cancer cells by recruiting BRD4 to chromatin through the CBP/EP300 complex (53). CCAT1-L promotes long-range chromatin looping at the *MYC* locus in colorectal cancer tumor development (54). In this study, we found that TCF7 is a critical regulator of GC and is associated with aberrantly formed E–P loops. We demonstrated that TCF7 ChIP-seq signals colocalized with chromatin-accessible regions and interacted with enhancer regions, which specifically occur in GC cells.

In previous studies on GC, cancer-associated super-enhancers have shown tendencies in cancer development, such as signal transduction, cell death and cell proliferation (15). Genome-wide enhancer alterations regulated by TEAD1 were enriched in biological adhesion, locomotion and cell migration (16). In our previous study, we found that ATF3 was enriched in the enhancer aberrations in EBV-positive GC cells (55). None of the reports included an interactome analysis. Another study reported that EHF accumulates at its downstream gene, *FZD5*, thereby promoting the development of EBV-positive GC (35). This study showed chromatin interactions detected by Hi-C but only focused on one gene locus and not genome-wide. A previous study applied H3K27ac HiChIP to EBV-positive GC. They extracted mitogen-activated protein kinase 8 (*MAP3K8*) from interactions and super-enhancers and validated its function in cell proliferation, colony formation and migration. They showed that MAP3K8 is associated with the Notch signaling pathway and EMT in EBV-positive cells (56). However, the genome-wide networks and regulators of the E–P loop have not been studied. In our study, we combined H3K27ac HiChIP with ChIP-seq, RNA-seq and ATAC-seq for an integrated analysis and found TCF7 along with its downstream gene *MYB* as an oncogene that functions in GC cellular growth and might perhaps be a therapeutic target.

MYB is a well-known oncogene involved in leukemia (57). It has been reported to be upregulated in several types of cancers, including GC (37–39,58). It has also been reported to fuse with other partner genes and contribute to cancer development (59,60). *MYB* can activate NF-κB and Wnt pathways, promote cancer growth and metastasis, and further contribute to carcinogenesis (61–63). This confirmed the importance of *MYB* in cancer development. Some studies on GC have shown that *ERBB2* overexpression/amplification

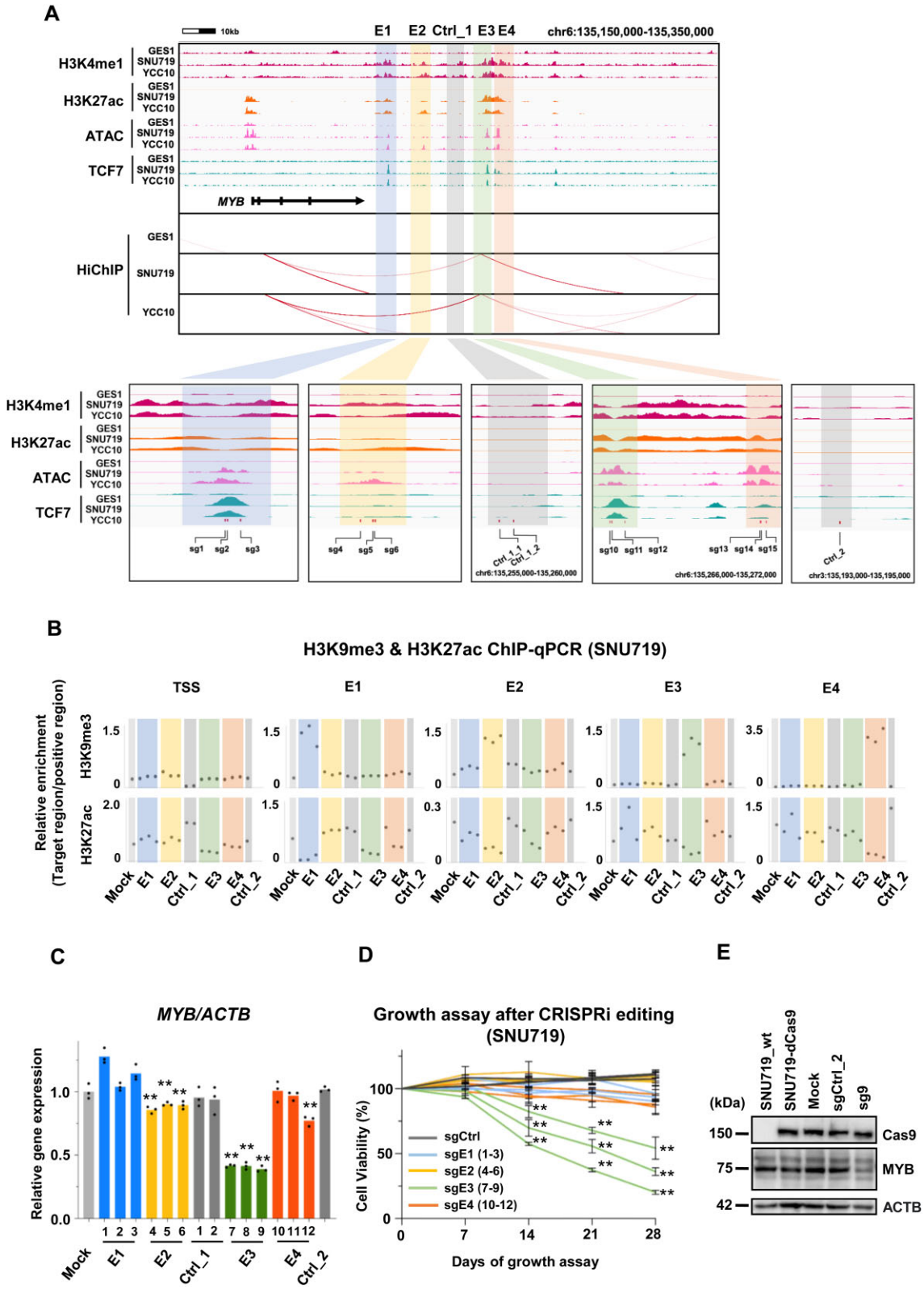


Figure 7. *MYB* interacted enhancer suppression by CRISPRi can repress gene expression and cellular growth. (**A**) Guide RNA targeting four enhancers (E1–E4) around *MYB*. Whole view shows ChIP-seq results for histone modification, ATAC peaks and ChIP-seq for TCF7, along with E–P loop at *MYB* locus. Enlarged view shows the location of sgRNA against enhancers; sgCtrl_1 targeted region between E2 and E3; sgCtrl_2 targeted at another chromosome locus (chr3: 115638881). (**B**) ChIP-qPCR of H3K9me3 and H3K27ac at TSS and E1–E4 regions after CRISPRi treatment in SNU719. Different sgRNA target regions include Mock, sg1–3 for E1, sg4–6 for E2, Ctrl_1_1 and Ctrl_1_2 between E2 and E3, sg7–9 for E3, sg10–12 for E4, and Ctrl_2 (from left to right). Relative enrichment was calculated by comparing the target region enrichment to positive region enrichment. Each data point is the average value of triplicates. (**C**) RT-qPCR result of *MYB* expression after CRISPRi. (**D**) Cell growth detected by fluorescence-activated cell sorting (FACS) after CRISPRi in SNU719. Cell viability was calculated by comparing D7/14/21/28 with D0 GFP%. (**E**) MYB protein expression was downregulated by sgRNA targeting E3.

is associated with poor prognosis. HER2, which is encoded by *ERBB2*, contributes to tumorigenesis by promoting cell proliferation and suppressing cell apoptosis (10,64). *MYC* overexpression/amplification is also found in patients with GC and is related to its presence or metastasis, or may present as an aggressive phenotype of GC (65). In recent years, lncRNAs have been reported to silence tumor suppressor genes in GC, leading to GC development (66). However, little is known about how *MYB* functions in GC. GO analysis showed that GC cell inhibition might be caused by the cell cycle. The colony formation assay indicated its function in tumorigenesis.

In addition, the mechanism of *MYB* transcription has not been analyzed well in GC. A previous study has revealed that ATF4 binds to the *HBS1L-MYB* intergenic enhancer region and regulates *MYB* expression in human erythroid cells (67). In the present study, different TFs exhibited the ability to regulate *MYB* activation. This might indicate that different TFs exist in different cell types during *MYB* transcription and that there might be a more complicated machinery process. TCF7 was named for its function in T-cell development and differentiation (68,69); several studies have found that TCF family members participate in the Wnt/β-catenin pathway in cancer (70,71). Wnt triggers β-catenin to migrate to the nucleus; following this, β-catenin will interact with TCF/LEF and their cofactor to activate Wnt downstream genes (72). TCF proteins also have oncogenic functions in the development of oncogenesis (73,74). A previous study showed that the enhancer looping gene *LDB1* colocalizes with TCF7 signals in liver cells (75). Although no interaction data were shown in this study, the results may indicate a role of TCF7 in chromatin remodeling. In our study, we showed that the TCF7 signal accumulated at the interacting enhancer of *MYB*, especially the enhancers involved in E–P contact formation on the chromatin structure. This finding may explain the critical role of TCFs in tumorigenesis, and that the active E–P loop plays an important role in oncogene activation. Enhancer activation and interaction status, however, were not clearly repressed by *TCF7* KD (Supplementary Figure S9A). This may be because KD by siRNA might be insufficient to show marked changes in H3K27ac levels at the target enhancers and E–P interaction, or TCF7 might not have pioneering function and there might be other TFs having critical functions in *MYB* enhancers.

There might be a few limitations in this study. Although we obtained clinical RNA-seq data from TCGA_STAD and performed IHC to show the upregulation of *MYB* in GC tissues, it is still necessary to conduct experiments to explore the E–P interactions and their regulatory networks in clinical samples. However, owing to the contamination of stromal cells with inflammatory cells, it is difficult to conduct HiChIP and obtain clean loop data from GC clinical tissues. In addition, there might be a group of TFs playing a dominant role in E–P loop formation and activation of downstream oncogenes. Further experiments, e.g. knocking out candidate TFs by CRISPR, should be conducted to elucidate the whole regulatory network and responsible factors to regulate aberrant chromatin interactions.

In summary, our genome-wide integrated analysis revealed that activation of enhancers with aberrantly formed E–P loops contributes to critical oncogene activation, and that *MYB* upregulated by TCF7 binding plays an important role in GC cellular growth.

Data availability

Next-generation sequencing data generated during this study have been deposited in Gene Expression Omnibus (GSE247539), which consist of H3K27ac HiChIP, ChIP-seq, ATAC-seq, RNA-seq, CUT&Tag and 4C-seq. GES1 ATAC-seq data are available at DNA Data Bank of Japan (DRR512981). Previously deposited next-generation sequencing data for RNA-seq and ChIP-seq are also available at GSE164225 and GSE135176, for histone NanoChIP-seq at GSE76153 and GSE75898, and for single-cell RNA-seq at EGAS00001007067. The authors declare that all other data are available within the article or associated supplementary information files, or are available from the author on request. Source data are provided in this paper. ATAC-seq, TCF7 ChIP-seq and CRISPRi H3K27ac ChIP-seq have been uploaded on UCSC Genome Browser (https://genome.ucsc.edu/s/Zhu-Tianhui/10 tracks in MYB locus).

Supplementary data

Supplementary Data are available at NAR Cancer Online.

Acknowledgements

We thank H. Maruyama for technical assistance, and Editage (http://www.editage.jp) for English language editing.

Author contributions: T.Z.: formal analysis, methodology, validation and writing—original draft. A.O.: conceptualization, formal analysis and writing—review & editing. G.U., R.F., D.K. and K.K.H.: formal analysis and methodology. M.S., M.F., H.A., M.N., T.O., M.Mi., M.Ma., Q.F., B.R. and T.H.: methodology. P.T., T.M. and T.U.: writing—review & editing. A.K.: conceptualization and writing—review & editing.

Funding

Japan Society for the Promotion of Science [19H03726 to A.K.]; Japan Agency for Medical Research and Development [19cm0106510h0004 to A.K., 22zf0127008s0301 to A.K., 23jm0210107h0001 to A.O. and A.K.]; Chiba University [Global and Prominent Research grant 2018-Y9 to A.K., IAAR grant to A.K.].

Conflict of interest statement

None declared.

References

1. Fueyo,R., Judd,J., Feschotte,C. and Wysocka,J. (2022) Roles of transposable elements in the regulation of mammalian transcription. *Nat. Rev. Mol. Cell Biol.*, **23**, 481–497.
2. Chakravarty,D. and Solit,D.B. (2021) Clinical cancer genomic profiling. *Nat. Rev. Genet.*, **22**, 483–501.
3. Heintzman,N.D., Stuart,R.K., Hon,G., Fu,Y., Ching,C.W., Hawkins,R.D., Barrera,L.O., Van Calcar,S., Qu,C., Ching,K.A., *et al.* (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.*, **39**, 311–318.
4. Schoenfelder,S. and Fraser,P. (2019) Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.*, **20**, 437–455.

5. 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.* (2020) Large-scale topological changes restrain malignant progression in colorectal cancer. *Cell*, **182**, 1474–1489.
6. Kloetgen,A., Thandapani,P., Ntziachristos,P., Ghebrechristos,Y., Nomikou,S., Lazaris,C., Chen,X., Hu,H., Bakogianni,S., Wang,J., *et al.* (2020) Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. *Nat. Genet.*, **52**, 388–400.
7. Xu,J., Song,F., Lyu,H., Kobayashi,M., Zhang,B., Zhao,Z., Hou,Y., Wang,X., Luan,Y., Jia,B., *et al.* (2022) Subtype-specific 3D genome alteration in acute myeloid leukaemia. *Nature*, **611**, 387–398.
8. Smyth,E.C., Nilsson,M., Grabsch,H.I., van Grieken,N.C. and Lordick,F. (2020) Gastric cancer. *Lancet North Am. Ed.*, **396**, 635–648.
9. Sung,H., Ferlay,J., Siegel,R.L., Laversanne,M., Soerjomataram,I., Jemal,A. and Bray,F. (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.*, **71**, 209–249.
10. Bass,A.J., Thorsson,V., Shmulevich,I., Reynolds,S.M., Miller,M., Bernard,B., Hinoue,T., Laird,P.W., Curtis,C., Shen,H., *et al.* (2014) Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*, **513**, 202–209.
11. Joshi,S.S. and Badgwell,B.D. (2021) Current treatment and recent progress in gastric cancer. *CA Cancer J. Clin.*, **71**, 264–279.
12. Lei,Z.N., Teng,Q.X., Tian,Q., Chen,W., Xie,Y., Wu,K., Zeng,Q., Zeng,L., Pan,Y., Chen,Z.S., *et al.* (2022) Signaling pathways and therapeutic interventions in gastric cancer. *Signal Transduct. Target. Ther.*, **7**, 358.
13. Muratani,M., Deng,N., Ooi,W.F., Lin,S.J., Xing,M., Xu,C., Qamra,A., Tay,S.T., Malik,S., Wu,J., *et al.* (2014) Nanoscale chromatin profiling of gastric adenocarcinoma reveals cancer-associated cryptic promoters and somatically acquired regulatory elements. *Nat. Commun.*, **5**, 4361.
14. Chia,N.Y., Deng,N., Das,K., Huang,D., Hu,L., Zhu,Y., Lim,K.H., Lee,M.H., Wu,J., Sam,X.X., *et al.* (2015) Regulatory crosstalk between lineage-survival oncogenes KLF5, GATA4 and GATA6 cooperatively promotes gastric cancer development. *Gut*, **64**, 707–719.
15. Ooi,W.F., Xing,M., Xu,C., Yao,X., Ramlee,M.K., Lim,M.C., Cao,F., Lim,K., Babu,D., Poon,L.-F., *et al.* (2016) Epigenomic profiling of primary gastric adenocarcinoma reveals super-enhancer heterogeneity. *Nat. Commun.*, **7**, 12983.
16. Ho,S.W.T., Sheng,T., Xing,M., Ooi,W.F., Xu,C., Sundar,R., Huang,K.K., Li,Z., Kumar,V., Ramnarayanan,K., *et al.* (2023) Regulatory enhancer profiling of mesenchymal-type gastric cancer reveals subtype-specific epigenomic landscapes and targetable vulnerabilities. *Gut*, **72**, 226–241.
17. Hua,J.T., Ahmed,M., Guo,H., Zhang,Y., Chen,S., Soares,F., Lu,J., Zhou,S., Wang,M., Li,H., *et al.* (2018) Risk SNP-mediated promoter–enhancer switching drives prostate cancer through lncRNA PCAT19. *Cell*, **174**, 564–575.
18. Okabe,A. and Kaneda,A. (2021) Transcriptional dysregulation by aberrant enhancer activation and rewiring in cancer. *Cancer Sci.*, **112**, 2081–2088.
19. See,Y.X., Chen,K. and Fullwood,M.J. (2022) MYC overexpression leads to increased chromatin interactions at super-enhancers and MYC binding sites. *Genome Res.*, **32**, 629–642.
20. Okabe,A., Huang,K.K., Matsusaka,K., Fukuyo,M., Xing,M., Ong,X., Hoshii,T., Usui,G., Seki,M., Mano,Y., *et al.* (2020) Cross-species chromatin interactions drive transcriptional rewiring in Epstein–Barr virus-positive gastric adenocarcinoma. *Nat. Genet.*, **52**, 919–930.
21. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
22. Mumbach,M.R., Rubin,A.J., Flynn,R.A., Dai,C., Khavari,P.A., Greenleaf,W.J. and Chang,H.Y. (2016) HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods*, **13**, 919–922.
23. Servant,N., Varoquaux,N., Lajoie,B.R., Viara,E., Chen,C.J., Vert,J.P., Heard,E., Dekker,J. and Barillot,E. (2015) HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.*, **16**, 259.
24. Bhattacharyya,S., Chandra,V., Vijayanand,P. and Ay,F. (2019) Identification of significant chromatin contacts from HiChIP data by FitHiChIP. *Nat. Commun.*, **10**, 4221.
25. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
26. Kaya-Okur,H.S., Wu,S.J., Codomo,C.A., Pledger,E.S., Bryson,T.D., Henikoff,J.G., Ahmad,K. and Henikoff,S. (2019) CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.*, **10**, 1930.
27. Krijger,P.H.L., Geeven,G., Bianchi,V., Hilvering,C.R.E. and de Laat,W. (2020) 4C-seq from beginning to end: a detailed protocol for sample preparation and data analysis. *Methods*, **170**, 17–32.
28. Trapnell,C., Williams,B.A., Pertea,G., Mortazavi,A., Kwan,G., Van Baren,M.J., Salzberg,S.L., Wold,B.J. and Pachter,L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.*, **28**, 511–515.
29. Raudvere,U., Kolberg,L., Kuzmin,I., Arak,T., Adler,P., Peterson,H. and Vilo,J. (2019) g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.*, **47**, W191–W198.
30. Buenrostro,J.D., Wu,B., Chang,H.Y. and Greenleaf,W.J. (2015) ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.*, **109**, 21.29.1–21.29.9.
31. Klann,T.S., Black,J.B., Chellappan,M., Safi,A., Song,L., Hilton,I.B., Crawford,G.E., Reddy,T.E. and Gersbach,C.A. (2017) CRISPR–Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat. Biotechnol.*, **35**, 561–568.
32. Sanson,K.R., Hanna,R.E., Hegde,M., Donovan,K.F., Strand,C., Sullender,M.E., Vaimberg,E.W., Goodale,A., Root,D.E., Piccioni,F., *et al.* (2018) Optimized libraries for CRISPR–Cas9 genetic screens with multiple modalities. *Nat. Commun.*, **9**, 5416.
33. Doench,J.G., Fusi,N., Sullender,M., Hegde,M., Vaimberg,E.W., Donovan,K.F., Smith,I., Tothova,Z., Wilen,C., Orchard,R., *et al.* (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR–Cas9. *Nat. Biotechnol.*, **34**, 184–191.
34. Usui,G., Matsusaka,K., Huang,K.K., Zhu,F., Shinozaki,T., Fukuyo,M., Rahmutulla,B., Yogi,N., Okada,T., Minami,M., *et al.* (2023) Integrated environmental, lifestyle, and epigenetic risk prediction of primary gastric neoplasia using the longitudinally monitored cohorts. *eBioMedicine*, **98**, 104844.
35. Li,W., Okabe,A., Usui,G., Fukuyo,M., Matsusaka,K., Rahmutulla,B., Mano,Y., Hoshii,T., Funata,S., Hiura,N., *et al.* (2021) Activation of EHF via STAT3 phosphorylation by LMP2A in Epstein–Barr virus-positive gastric cancer. *Cancer Sci.*, **112**, 3349–3362.
36. Deng,S., Feng,Y. and Pauklin,S. (2022) 3D chromatin architecture and transcription regulation in cancer. *J. Hematol. Oncol.*, **15**, 49.
37. Cicirò,Y. and Sala,A. (2021) MYB oncoproteins: emerging players and potential therapeutic targets in human cancer. *Oncogenesis*, **10**, 19.
38. Mitra,P. (2018) Transcription regulation of MYB: a potential and novel therapeutic target in cancer. *Ann. Transl. Med.*, **6**, 443.
39. Jiang,Q., Chen,Z., Meng,F., Zhang,H., Chen,H., Xue,J., Shen,X., Liu,T., Dong,L., Zhang,S., *et al.* (2023) CD36-BATF2\MYB axis predicts anti-PD-1 immunotherapy response in gastric cancer. *Int. J. Biol. Sci.*, **19**, 4476–4492.
40. Robson,M.I., Ringel,A.R. and Mundlos,S. (2019) Regulatory landscaping: how enhancer–promoter communication is sculpted in 3D. *Mol. Cell*, **74**, 1110–1122.
41. Sigova,A.A., Abraham,B.J., Ji,X., Molinie,B., Hannett,N.M., Guo,Y.E., Jangi,M., Giallourakis,C.C., Sharp,P.A. and Young,R.A.

(2015) Transcription factor trapping by RNA in gene regulatory elements. *Science*, **350**, 978–981.
42. Oudelaar,A.M. and Higgs,D.R. (2021) The relationship between genome structure and function. *Nat. Rev. Genet.*, **22**, 154–168.
43. Kim,S. and Shendure,J. (2019) Mechanisms of interplay between transcription factors and the 3D genome. *Mol. Cell*, **76**, 306–319.
44. Huang,K.K., Ma,H., Chong,R.H.H., Uchihara,T., Lian,B.S.X., Zhu,F., Sheng,T., Srivastava,S., Tay,S.T., Sundar,R., *et al.* (2023) Spatiotemporal genomic profiling of intestinal metaplasia reveals clonal dynamics of gastric cancer progression. *Cancer Cell*, **41**, 2019–2037.
45. Bradner,J.E., Hnisz,D. and Young,R.A. (2017) Transcriptional addiction in cancer. *Cell*, **168**, 629–643.
46. Stadhouders,R., Filion,G.J. and Graf,T. (2019) Transcription factors and 3D genome conformation in cell-fate decisions. *Nature*, **569**, 345–354.
47. Zheng,H. and Xie,W. (2019) The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.*, **20**, 535–550.
48. Field,A. and Adelman,K. (2020) Evaluating enhancer function and transcription. *Annu. Rev. Biochem.*, **89**, 213–234.
49. Sur,I. and Taipale,J. (2016) The role of enhancers in cancer. *Nat. Rev. Cancer*, **16**, 483–493.
50. Kubo,N., Ishii,H., Xiong,X., Bianco,S., Meitinger,F., Hu,R., Hocker,J.D., Conte,M., Gorkin,D., Yu,M., *et al.* (2021) Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nat. Struct. Mol. Biol.*, **28**, 152–161.
51. Bhat,P., Honson,D. and Guttman,M. (2021) Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nat. Rev. Mol. Cell Biol.*, **22**, 653–670.
52. Weintraub,A.S., Li,C.H., Zamudio,A.V., Sigova,A.A., Hannett,N.M., Day,D.S., Abraham,B.J., Cohen,M.A., Nabet,B., Buckley,D.L., *et al.* (2017) YY1 is a structural regulator of enhancer–promoter loops. *Cell*, **171**, 1573–1588.
53. Liu,Y., Guo,B., Aguilera-Jimenez,E., Chu,V.S., Zhou,J., Wu,Z., Francis,J.M., Yang,X., Choi,P.S., Bailey,S.D., *et al.* (2020) Chromatin looping shapes KLF5-dependent transcriptional programs in human epithelial cancers. *Cancer Res.*, **80**, 5464–5477.
54. Xiang,J.F., Yin,Q.F., Chen,T., Zhang,Y., Zhang,X.O., Wu,Z., Zhang,S., Wang,H.B., Ge,J., Lu,X., *et al.* (2014) Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res.*, **24**, 513–531.
55. Asakawa,Y., Okabe,A., Fukuyo,M., Li,W., Ikeda,E., Mano,Y., Funata,S., Namba,H., Fujii,T., Kita,K., *et al.* (2020) Epstein–Barr virus-positive gastric cancer involves enhancer activation through activating transcription factor 3. *Cancer Sci.*, **111**, 1818–1828.
56. Roy,G., Yang,T., Liu,S., Luo,Y.L., Liu,Y. and Zhong,Q. (2023) Epigenetic regulation of MAP3K8 in EBV-associated gastric carcinoma. *Int. J. Mol. Sci.*, **24**, 1964.
57. Clesham,K., Walf-Vorderwülbecke,V., Gasparoli,L., Virely,C., Cantilena,S., Tsakaneli,A., Inglott,S., Adams,S., Samarasinghe,S., Bartram,J., *et al.* (2022) Identification of a c-MYB-directed therapeutic for acute myeloid leukemia. *Leukemia*, **36**, 1541–1549.
58. Liang,J., Liu,X., Xue,H., Qiu,B., Wei,B. and Sun,K. (2015) MicroRNA-103a inhibits gastric cancer cell proliferation, migration and invasion by targeting c-Myb. *Cell Prolif.*, **48**, 78–85.
59. Ducassou,S., Prouzet-Mauléon,V., Deau,M.C., Brunet de la Grange,P., Cardinaud,B., Soueidan,H., Quelen,C., Brousset,P., Pasquet,J.M., Moreau-Gaudry,F., *et al.* (2017) MYB–GATA1 fusion promotes basophilic leukaemia: involvement of interleukin-33 and nerve growth factor receptors. *J. Pathol.*, **242**, 347–357.
60. Andersson,M.K., Afshari,M.K., Andrén,Y., Wick,M.J. and Stenman,G. (2017) Targeting the oncogenic transcriptional regulator MYB in adenoid cystic carcinoma by inhibition of IGF1R/AKT signaling. *J. Natl Cancer Inst.*, **109**, djx017.
61. Qu,X., Yan,X., Kong,C., Zhu,Y., Li,H., Pan,D., Zhang,X., Liu,Y., Yin,F. and Qin,H. (2019) c-Myb promotes growth and metastasis of colorectal cancer through c-fos-induced epithelial–mesenchymal transition. *Cancer Sci.*, **110**, 3183–3196.
62. Li,Y., Jin,K., Van Pelt,G.W., Van Dam,H., Yu,X., Mesker,W.E., Ten Dijke,P., Zhou,F. and Zhang,L. (2016) c-Myb enhances breast cancer invasion and metastasis through the Wnt/β-catenin/Axin2 pathway. *Cancer Res.*, **76**, 3364–3375.
63. Dúckaa,M., Kučeríková,M., Trčka,F., Červinka,J., Biglieri,E., Šmarda,J., Borsig,L., Beneš,P. and Knopfová,L. (2021) c-Myb interferes with inflammatory IL1α–NF-κb pathway in breast cancer cells. *Neoplasia*, **23**, 326–336.
64. Boku,N. (2014) HER2-positive gastric cancer. *Gastric Cancer*, **17**, 1–12.
65. Calcagno,D.Q., Leal,M.F., Assumpçaão,P.P., de Arruda Cardoso Smith,M. and Burbano,R.R. (2008) MYC and gastric adenocarcinoma carcinogenesis. *World J. Gastroenterol.*, **14**, 5962–5968.
66. Huarte,M. (2015) The emerging role of lncRNAs in cancer. *Nat. Med.*, **21**, 1253–1261.
67. Boontanrart,M.Y., Schröder,M.S., Stehli,G.M., Banović,M., Wyman,S.K., Lew,R.J., Bordi,M., Gowen,B.G., DeWitt,M.A. and Corn,J.E. (2020) ATF4 regulates MYB to increase γ-globin in response to loss of β-globin. *Cell Rep.*, **32**, 107993.
68. Choi,Y.S., Gullicksrud,J.A., Xing,S., Zeng,Z., Shan,Q., Li,F., Love,P.E., Peng,W., Xue,H.H. and Crotty,S. (2015) LEF-1 and TCF-1 orchestrate TFH differentiation by regulating differentiation circuits upstream of the transcriptional repressor Bcl6. *Nat. Immunol.*, **16**, 980–990.
69. Zhou,X., Yu,S., Zhao,D.M., Harty,J.T., Badovinac,V.P. and Xue,H.H. (2010) Differentiation and persistence of memory $CD8^+$ T cells depend on T cell factor 1. *Immunity*, **33**, 229–240.
70. Hovanes,K., Li,T.W.H., Munguia,J.E., Truong,T., Milovanovic,T., Lawrence Marsh,J., Holcombe,R.F. and Waterman,M.L. (2001) β-Catenin-sensitive isoforms of lymphoid enhancer factor-1 are selectively expressed in colon cancer. *Nat. Genet.*, **28**, 53–57.
71. Fröhlich,J., Rose,K. and Hecht,A. (2023) Transcriptional activity mediated by β-CATENIN and TCF/LEF family members is completely dispensable for survival and propagation of multiple human colorectal cancer cell lines. *Sci. Rep.*, **13**, 287.
72. Zhang,Y. and Wang,X. (2020) Targeting the Wnt/β-catenin signaling pathway in cancer. *J. Hematol. Oncol.*, **13**, 165.
73. Ji,L., Qian,W., Gui,L., Ji,Z., Yin,P., Lin,G.N., Wang,Y., Ma,B. and Gao,W.Q. (2021) Blockade of β-catenin-induced CCL28 suppresses gastric cancer progression via inhibition of Treg cell infiltration. *Cancer Res.*, **80**, 2004–2016.
74. Ganesan,K., Ivanova,T., Wu,Y., Rajasegaran,V., Wu,J., Ming,H.L., Yu,K., Sun,Y.R., Hyun,C.C., Ylstra,B., *et al.* (2008) Inhibition of gastric cancer invasion and metastasis by PLA2G2A, a novel β-catenin/TCF target gene. *Cancer Res.*, **68**, 4277–4286.
75. Liu,G., Wang,L., Wess,J. and Dean,A. (2022) Enhancer looping protein LDB1 regulates hepatocyte gene expression by cooperating with liver transcription factors. *Nucleic Acids Res.*, **50**, 9195–9211.

Received: December 12, 2023. **Revised:** April 7, 2024. **Editorial Decision:** April 22, 2024. **Accepted:** April 25, 2024