Cellular Oncology (2025) 48:617–635
https://doi.org/10.1007/s13402-024-01033-6

RESEARCH

Three-dimensional genome architecture in intrahepatic cholangiocarcinoma

Youfeng Liang[1] · Cong Li[2,3] · Renchao Zou[4] · Lu Ying[1,5] · Xiaoyang Chen[1] · Zhaohai Wang[2] · Wenjing Zhang[1] · Mingxuan Hao[1] · Hao Yang[2] · Rui Guo[1] · Guanglin Lei[6] · Fang Sun[6] · Kexu Zhao[1] · Yu Zhang[6] · Jia Dai[1] · Shangya Feng[1] · Keyue Zhang[1] · Luyuan Guo[1] · Shuyue Liu[1] · Chuanxing Wan[5] · Lin Wang[4] · Penghui Yang[2,3] · Zhao Yang[1,5]

Accepted: 19 December 2024 / Published online: 20 January 2025

Abstract

Purpose Intrahepatic cholangiocarcinoma (ICC) is a common primary hepatic tumors with a 5-year survival rate of less than 20%. Therefore, it is crucial to elucidate the molecular mechanisms of ICC. Recently, the advance of high-throughput chromosome conformation capture (Hi-C) technology help us look insight into the three-dimensional (3D) genome structure variation during tumorigenesis. However, its function in ICC pathogenesis remained unclear.

Methods Hi-C and RNA-sequencing were applied to analyze 3D genome structures and gene expression in ICC and adjacent noncancerous hepatic tissue (ANHT). Furthermore, the dysregulated genes due to 3D genome changes were validated via quantitative real-time PCR and immunohistochemistry.

Results Primarily, the intrachromosomal interactions of chr1, chr2, chr3, and chr11 and the interchromosomal interactions of chr1-chr10, chr13-chr21, chr16-chr19, and chr19-chr22 were also significantly distinct between ANHT and ICC, which may potentially contribute to the activation of cell migration and invasion via the upregulation of *WNT10A, EpCAM, S100A3/A6, and MAPK12*. Interestingly, 56 compartment regions from 23 chromosomes underwent A to B or B to A transitions during ICC oncogenesis, which attenuated the complement pathway through the downregulation of *C8A/C8B, F7, F10,* and *F13B*. Notably, topologically associated domain (TAD) rearrangements were identified in the region containing *HOPX* (chr4: 57,514,154–57,522,688) and *ACVR1* (chr2:158,592,958–158,732,374) in ICC, which may contribute to the hijacking of remote enhancers that were previously outside the TAD and increased expression of *HOPX* and *ACVR1*.

Conclusions This study reveals relationship between 3D genome structural variations and gene dysregulation during ICC tumorigenesis, indicating the molecular mechanisms and potential biomarkers.

Keywords Intrahepatic cholangiocarcinoma · Three-dimensional (3D) genome architecture · Biomarker · *HOPX* · *ACVR1*

Youfeng Liang, Cong Li and Renchao Zou contributed equally to this work.

✉ Zhao Yang
yangzhao@mail.buct.edu.cn

✉ Penghui Yang
ypenghuiamms@hotmail.com

✉ Lin Wang
wanglinfey@126.com

[1] College of Life Science and Technology, Innovation Center of Molecular Diagnostics, Beijing University of Chemical Technology, Beijing 100029, China

[2] The First Medical Center, Chinese PLA General Hospital, Beijing 100853, China

[3] School of Basic Medical Sciences, Inner Mongolia Medical University, Hohhot, China

[4] Department of Urology, Second Affiliated Hospital of Kunming Medical University, Kunming 650000, China

[5] College of Life Science and Technology, Key Laboratory of Protection and Utilization of Biological Resources in Tarim Basin of Xinjiang Production and Construction Corps, Tarim University, Alar, Xinjiang 843300, China

[6] The Fifth Medical Center, Chinese PLA General Hospital, Beijing 100039, China

Abbreviations

ICC	Intrahepatic cholangiocarcinoma
HCC	Hepatocellular carcinoma
OS	Overall survival
RFS	Relapse-free survival
3D	Three-dimensional
TADs	Topologically associating domains
SVs	Structural variants
ANHT	Adjacent non-cancerous hepatic tissues
RT	Room temperature
ICE	Iterative correction
PCA	Principal component analysis
HMM	Hidden Markov Model
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
UVI	Unique valid interaction
DEG	Digital gene expression tag profiling
GO	Gene ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
TCGA	The Cancer Genome Atlas
CHOL	Cholangiocarcinoma
PAAD	Pancreatic adenocarcinoma
STAD	Stomach adenocarcinoma
COAD	Colon adenocarcinoma
ESCA	Esophageal carcinoma
READ	Rectum adenocarcinoma
DLBC	Diffuse large B-cell lymphoma
HNSC	Head and neck squamous cell carcinoma
THYM	Thymoma
qRT-PCR	Quantitative real-time PCR
TSTADBs	Tumor-specific TAD boundaries
LAML	Acute myeloid leukemia

1 Introduction

Intrahepatic cholangiocarcinoma (ICC), a lethal malignancy originating from epithelial cells of secondary bile duct of the liver, is the second most common primary hepatic tumors after hepatocellular carcinoma (HCC) [1]. The incidence of ICC has increased by more than 140% over the past four decades [1]. ICC represents 8–10% of all cholangiocarcinomas and 10–20% of all primary hepatic tumors [2]. Unfortunately, the vast majority of ICC (70–80%) are diagnosed at an advanced stage [3]. Additionally, the 5-year survival rates of ICC patients ranged from 7 to 20% and the recurrence rates after resection ranged from 60–70% [4, 5]. Specifically, the median overall survival (OS) and relapse-free survival (RFS) of approximately 30 months and 20 months, respectively [6]. The mechanisms underlying ICC tumorigenesis are elusive, and there is a lack of early diagnostic markers and specific treatments, which do not meet the clinical requirements for ICC diagnosis and treatment. It is of great significance to uncover the molecular mechanisms of tumorigenesis in ICC.

Recent research has shown that the three-dimensional (3D) genome architecture of eukaryotic cells plays a crucial role in regulating gene expression [7], impacting various processes such as proliferation, differentiation and disease [8]. Genome organization is a intricate and multi-tiered procedure, encompassing the compacting of chromatin into nucleosomes, chromatin fibers, compartments, and more extensive chromosomal regions [9]. The structural variations from different scales of the 3D genome, including the entire chromosome, A/B compartments, topologically associating domains (TADs) and chromatin loops could potentially affect genome interactions in the 3D nuclear space and then influence the cellular gene expression patterns [10].

More and more evidence suggested the important role of 3D genome organization in cancer development. In 2009, Aiden et al. employed Hi-C technology to investigate the 3D architecture of human lymphoblastoid cells and introduced the A/B compartment concept [11]. Barutcu et al. discovered the transformation of A/B compartments and revealed that the alteration of A/B compartments between normal and breast cancer cells is associated with the corresponding changes in gene expression [12]. Wu et al. detected compartment switching in myeloma and demonstrated its impact on the expression of genes related to myeloma [13]. Some studies indicated that higher-order chromatin structural alterations, including chromosomal interactions, compartment reprogramming, and TAD division and fusion, are associated with tumorigenesis of multiple myeloma [12], prostate cancer [13], breast cancer [14]. Furthermore, Feng et al. performed a comprehensive 3D-epigenomic analyses in cellular models of THLE2 (normal hepatocytes cell line) and HepG2 (HCC cell line), founding that epigenetic variants by DNA hypomethylation at a subset of CTCF motifs proximal to HCC-associated genes could modify chromatin topological configuration, which in turn altered RNAPII-mediated chromatin interactions and contribute to dysregulation of transcription [15]. Additionally, Zhang et al. carried out RNA sequencing and whole-genome sequencing of 1220 cancer cases and found that structural variants (SVs) breakpoint associated with altered gene expression [16]. Specifically, *TERT, MDM2, CDK4*, etc. were upregulated in cancers resulting from the SVs breakpoint. Mechanically, elevated numbers or greater proximity of enhancer regulatory elements adjacent to the gene might contribute to the potential dysregulation of gene expression. However, the 3D genome structure alteration during the tumorigenesis of ICC remained unknown.

Here, Hi-C and RNA-seq techniques were employed to investigate the 3D genome structure variation and gene expression during the tumorigenesis of ICC by using three paired primary ICC samples and their corresponding adjacent non-cancerous hepatic tissues (ANHT). ICC specific chromatin intra/interchromosomal interactions, compartment switch, TAD and related gene dysregulation were demonstrated.

2 Materials and methods

2.1 Human tumor specimens

Primary human ICC and ANHT tissues (Table S1) were obtained from Fifth Medical Center of Chinese PLA General Hospital (Beijing, China) with informed consent and approved by the Research Ethics Board at Fifth Medical Center of Chinese PLA General Hospital (S2016-98-02).

2.2 Tissue preparation and Hi-C experiment

Tumor and normal ICC samples were diced into small (<1 mm) pieces on ice and digested with collagenase to form a cell suspension, filtered and resuspended in ice-cold PBS. When crosslinking, formaldehyde was added to 1% and tissue was rotated at room temperature (RT) for 10 min during which the solution is gently shaken every 1 min. Glycine (2.5 M) was added and samples were rotated for an additional 5 min at RT. After the completion of crosslinking and lysis, chromatin was digested with 400 U Mbo Irestriction enzyme (NEB) at 37 °C. DNA ends were labeled with biotin and incubated at 37 °C for 45 min, and the enzyme was inactivated with 20% SDS solution. DNA ligation was performed by the addition of T4 DNA ligase (NEB) and incubation at 16 °C for 4–6 h. After ligation, proteinase K was added. DNA fragments were purified and dissolved in water. After remove unligated ends, purified DNA was fragmented to a size of 300–500 bp. The fragments labeled by biotin were finally separated on Dynabeads® M-280 Streptavidin (Life Technologies). Hi-C libraries with two replicates were controlled for quality and sequenced on an Illumina sequencer.

2.3 Hi-C data processing, mapping, and ICE normalization

For Hi-C data, we first removed adapter sequences and low-quality reads from a portion of the raw data. The ratio of clean bases (bp) to raw bases (bp), defined as the effective rate (%), was greater than 96.98% for all three sample pairs (Table S2). We then used Bowtie 2 to align the clean reads to the human genome (hg19). Next, Hi-C Pro v2.11.1 was used to filter out dangling end pairs, self-cycles, and invalid data resulting from PCR amplification of duplicate read pairs, resulting in unique valid interactions for analysis. The effective rate (%) of unique valid interactions to total pairs processed was greater than 51.66% for all three sample pairs (effective rate (%) = Unique valid interaction/Total pairs processed) (Table S3). We generated genome-wide raw interaction matrices at resolutions of 1 Mb, 40 kb, and 10 kb, and used iterative correction (ICE) normalization to eliminate biases in the raw matrices. Finally, we confirmed data quality by evaluating the distribution of sequencing quality, GC content, and insert size statistics for deduplicated data [17].

2.4 Hi-C analysis

Hi-C maps of chromosome topology were initially generated by three paired ANHT and ICC samples. We confirmed data quality by assessing the sequencing quality distribution, GC content and insert fragment statistics.

2.5 Definition of differential interactions

Firstly, calculate the interaction probability under each window of each sample; filter out the window with 0 interaction number and the window that does not conform to the binomial distribution, and mark the window where the interaction probability of two-fold difference in the two samples. Then calculate the *p* value of the retained window, and correct the *p* value; Count the number of windows with interaction, the number of windows with significant interaction, and the number of windows where there is a significant interaction with a two-fold difference in the probability of interaction. Finally, according to the above principles, a genome-wide mapping of the differential interaction between the two samples was performed.

2.6 Compartment profiles

The obtained UVI is optimized through the following steps to obtain the correlation coefficient matrix: calculate the data in the xth row and yth column of the original standard matrix according to formula 1 to obtain the xth row and yth column data of the original standard matrix (Pearson correlation coefficient). By calculating the eigenvector values of the matrix through principal component analysis (PCA) at 1Mb, the chromosomes are divided into different compartments according to the first principal component (E1): Compartment A (E1 > 0), Compartment B (E1 < 0).

$$P_{x,y} = \frac{\sum xy - \frac{\sum x \sum y}{N}}{\sqrt{(\sum x^2 - \frac{(\sum x)^2}{N})(\sum y^2 - \frac{(\sum y)^2}{N})}} \tag{1}$$

2.7 Identification of TADs and CTCF bonding site

The chromosome was divided into windows with a fixed length, which was the resolution of the TAD analysis. The directionality index (DI) values of each window were calculated and iterated by Hidden Markov Model (HMM). Finally, TAD was visualized via WashU Epigenome Browser V54.0.0. The motif for CTCF binding is recognized as NCANNAGGNGGC [18]. Combined with Enhancer Atlas 2.0 [19], a database of enhancers, the enhancer targeting *TP53I3* and *MED10* in cells was identified.

2.8 Tissue preparation and RNA-seq

Whole RNA was extracted by the Trizol-based method. For RNA-seq library preparation, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Four fluorescently labeled dNTPs, DNA polymerase and adapter primers are added to the sequencing flow cell for amplification. The sequencer captures the fluorescent signal and converts the light signal into a sequencing peak through computer software to obtain the sequence information of the fragment to be tested.

2.9 RNA-seq raw data pre-processing

Raw data of fastq format were firstly processed through in-house perl scripts. Clean data were obtained by removing reads containing adapter, reads containing N base and low quality reads from raw data. Q20, Q30, GC content and clean data were calculated. All the downstream analyses were based on the clean data. Hisat2 v2.0.5 was used to build index of the reference genome and align clean reads to the reference genome. Quantification of gene expression: Feature Counts v1.5.0-p3 was used to count the reads numbers mapped to each gene. FPKM of each gene was calculated based on the length of the gene and reads count mapped to this gene.

2.10 Differential expression analysis

Differential expression analysis of ANHT and ICC was performed using the DESeq2 R package (1.20.0). DESeq2 provides the statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting *P*-values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted $P\text{-value} < 0.05$ found by DESeq2 were assigned as differentially expressed.

2.11 Quantitative real-time PCR (qTR-PCR)

In the validation step, 63 paired primary ICC and ANHT were obtained from from the Fifth Medical Centre (Beijing, China). Total RNA was extracted using TRIzol reagent (Ambion). RNA concentration was measured by Nanodrop and was subjected to cDNA synthesis using FastKing RT Kit (KR116-02, Tiangen, China) following the manufacturer's instructions. The cDNA was then used as the templates for qRT-PCR of the candidate genes. PCR MasterMix was used to carry out the Real-time PCR according to the manufacturer's protocols. Primers used in this study were obtained from Tsingke and the sequences were listed in Table S18. SYBR Green (Qiagen) was used as the fluorescent probe. Relative expression levels of the target genes were compared with a housekeeping gene, *glyceraldehyde 3-phosphate dehydrogenase* (*GAPDH*). The fold change of differentially expressed genes in ICC compared with ANHT was calculated with the method of $2^{-\Delta\Delta Ct}$.

2.12 Immunohistochemistry

The tissue was sliced and baked at 60 °C for 1 h. The paraffin-embedded tissue sections were then deparaffinized in xylene, rehydrated with graded concentrations of ethanol, and rinsed with tap water and hydrogen peroxide for 1 min each. The slides were incubated in PBS containing 3% H_2O_2 for 10 min at room temperature and rinsed 3 times with distilled water for 3 min. Immerse the sections in 0.01 M citrate buffer, heat to maximum power (98–100 °C) in a microwave oven, and cool for 10 min. This step is repeated twice. The sections were naturally cooled to room temperature and washed with PBS for 3 times for 5 min. Antibody was added dropwise and incubated at 37 °C for 1 h or at 4 °C overnight. After washing with PBS 3 times for 3 min, the surrounding water was absorbed with filter paper. Add the secondary antibody, incubate at 37 °C for 0.5 h and washed with PBS 3 times for 3 min. Use DAB to develop the color, observe the staining under a microscope, stop the staining reaction after 5 min of staining, rinse with tap water, and then rinse with distilled water for several times. Hematoxylin was counterstained for 2 min and rinsed with tap water. Dehydrate with graded ethanol, and treat with xylene for 20 min; seal the slides with gum, and examine under microscope. Antibodies used are listed in Table S19.

2.13 Statistical analysis

The Student *t* test was performed to analyze whether two experimental groups have significant difference using $P < 0.05$ as the significant criteria. A value of *P* less than

0.05 (*, $P<0.05$; **, $P<0.01$; and ***, $P<0.001$) was regarded to indicate a significant difference.

3 Results

3.1 Three-dimensional chromatin profile of ANHT and ICC

To investigate the genome-wide chromatin landscape of ANHT and ICC, paired ANHT and ICC were collected from three diagnosed ICC patients (Table S1). ANHT and ICC were minced and digested with collagenase into single cells, which generated Hi-C and RNA-seq libraries for sequencing (Fig. 1A). Hi-C sequencing produced a total of 2053.738 Gb of Raw reads (Table S2), which were subjected to sequence quality checking and filtering. The sequencing quality distribution, GC content, and insert fragment of all six samples were found to be satisfactory for further analysis (Fig. S1A–C). Furthermore, 2022.167 Gb clean reads were screened after filtering (Fig. S2, Table S2). Specifically, the unique paired alignments, refer to the reads matched to human hg19 genome at both ends were ≥66.00%, and the unique valid interaction (UVI) were ≥63.07% (Table S3). All the data met the requirements for further analysis.

A chromosome versus chromosome heatmap was applied to display a normalized genome-wide interaction landscape in all six samples (Fig. 1B, Table S4). The heatmaps indicated that the frequency of intrachromosomal interactions (~328.036 million, shown as red regions along the diagonal) was significantly greater than that of inter-chromosomal interactions (~248.812 million, shown as green regions off the diagonal) across chromosomes in both ANHT and ICC. Similarly, the normalized genome-wide interchromosomal profiles in all six samples were also displayed (Fig. 1C, Table S4). Specifically, there are ~618 million and ~875 million trans interaction detected in ANHT and ICC, respectively (Table S4). And a significant increase in interchromosomal interactions across chr3 to chrY in ANHT compared to ICC (Fig. 1C). Collectively, this study carried out the Hi-C sequencing of three paired ANHT and ICC successfully, revealing the genome-wide chromatin landscapes of ANHT and ICC.

3.2 The chromosomal interactions between ANHT and ICC are remarkable distinct

In order to explore the unique genome-wide interactions between ANHT and ICC, this study generated a differential interaction heat map (Fig. 2A–C, Table S5). Importantly, the majority of these differential interactions were found to be intrachromosomal, which is represented by yellow regions along the diagonal. Besides, a small fraction of the differential interactions were interchromosomal, indicated by blue regions off the diagonal. This pattern was observed in all three paired ANHT and ICC. Specifically, over 6500 differential intrachromosomal interactions were identified in chr1, chr2, chr3, and chr11 (Fig. 2D, Table S6) and a total of 376 differential interchromosomal interactions were observed between chr1–chr10, chr13–chr21, chr19–chr22, and chr21–chr22, as depicted in the circos map (Fig. 2E, Table S6). Finally, this study provided insight into the genome-wide differential interactions between ANHT and ICC, and identifying the significantly altered chromosomes involved in the pathogenesis of ICC.

3.3 Intrachromosomal interactions activates cell migration signaling in ICC

To explore the potential link between the unique interaction network and gene dysregulation in ICC, RNA-seq were applied to three paired ANHT and ICC. Extracting and combining substantially differentially expressed genes ($|\log2\ \text{foldchange}|>1$, $\text{padj}<0.05$) for each group (ICCs vs. ANHTs) correlated with 4103–5157 upregulated genes and 3690–4671 downregulated genes in ICC compared to ANHT (Fig. S3A, Table S7). Among the three paired ANHT and ICC, the digital gene expression tag profiling (DEG) gene counts were showed (Fig. S3B).

Furthermore, the average expression levels of genes located in the differential chromosomal interaction regions on chr1, chr3, and chr11 were substantially lower in ICC than those in ANHT. However, the average expression levels of genes located in the differential chromosomal interaction regions on chr2 were not significantly changed between ICC and ANHT (Fig. 3A). Differentially expressed genes located in differential chromosomal interaction regions on chr1, chr2, chr3, and chr11 in ICC were also identified (Table S8), which were applied to the enrichment analysis by Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG). Among the GO terms, “collagen fibril organization”, “extracellular matrix disassembly”, “cell-cell adhesion” and “cell adhesion” were significantly enriched (Fig. 3B, Table S8). Similarly, “ECM receptor interaction”, “protein digestion and absorption” and “Cell Cycle” were significantly enriched in KEGG analysis (Fig. 3C, Table S8). Above results indicated that intrachromosomal interactions correlated with the activation of cell migration and invasion in ICC.

Among the above differentially expressed genes enriched in cell migration and invasion related pathways, *WNT10A* (chr2: 219,745,255–219,758,651), *EpCAM* (chr2: 47,596,287–47,614,167), *COL5A2* (chr2: 189,896,641–190,044,605), and *COL16A1* (chr1: 32,117,848–32,169,768)

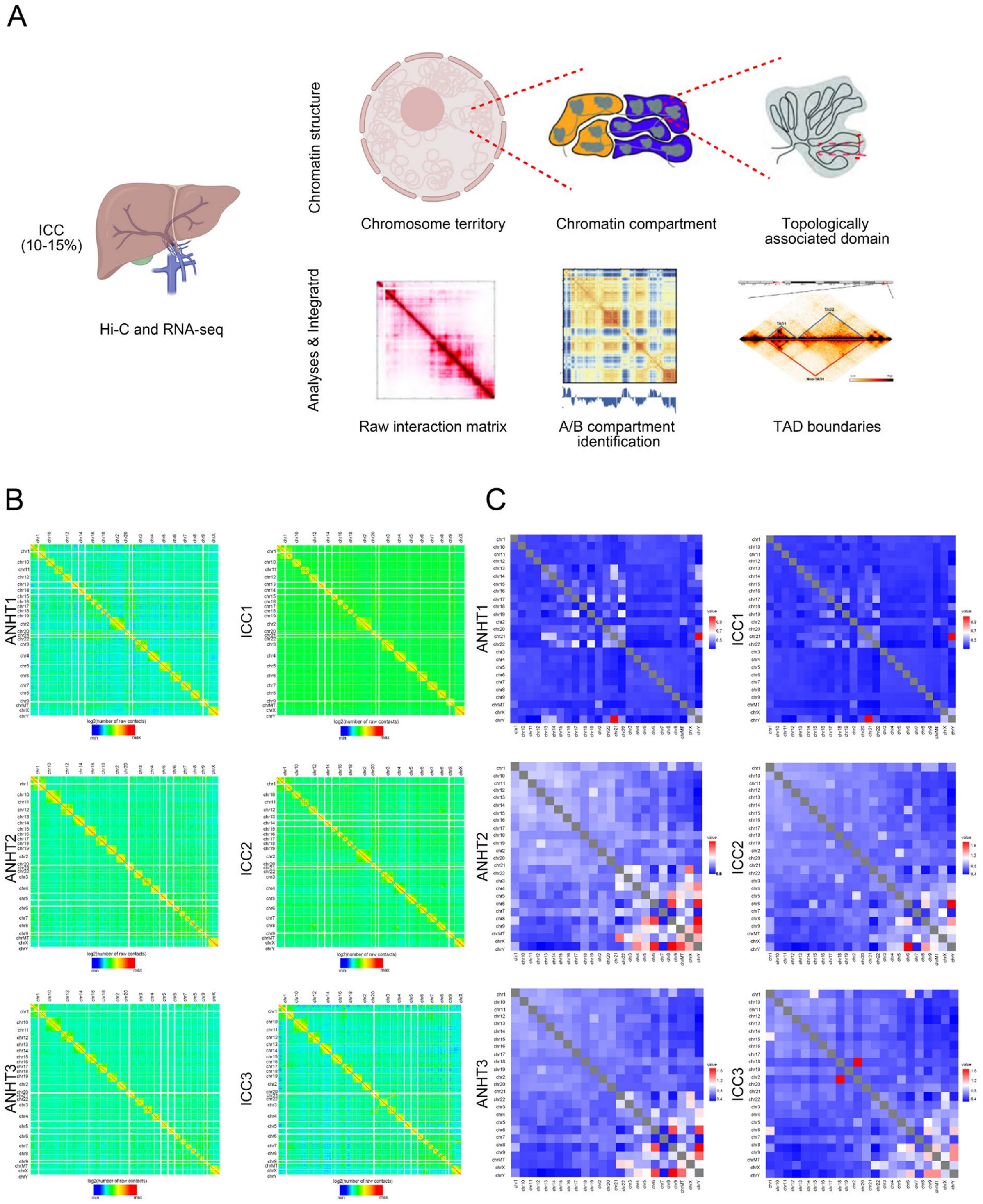

Fig. 1 Three-dimensional chromatin landscapes of ANHT and ICC. **A** The experimental approach is illustrated in a schematic diagram. **B** Genome-wide intrachromosomal contact matrix of ANHT and ICC. The letters and numbers on the left and top of the image represent different chromosomes. The interaction heatmap shows the log2 fold change of normalized interactions for the sample. The intra-chromosomal interaction frequency is depicted from weak to strong in blue-green-yellow-red colors. The intra-chromosomal interaction frequency (indicated by the red diagonal line) is higher than the inter-chromosomal interaction frequency. "chrMT" represents mitochondrial DNA. **C** Interchromosomal contact matrix of ANHT and ICC. The heatmaps are stacked from top left to bottom right in order (chr1, chr10 … chr22 and chrY). The inter-chromosomal interaction frequency is represented from weak to strong in blue-white-red colors. The gray regions indicate repetitive regions in which the sequencing reads could not be mapped

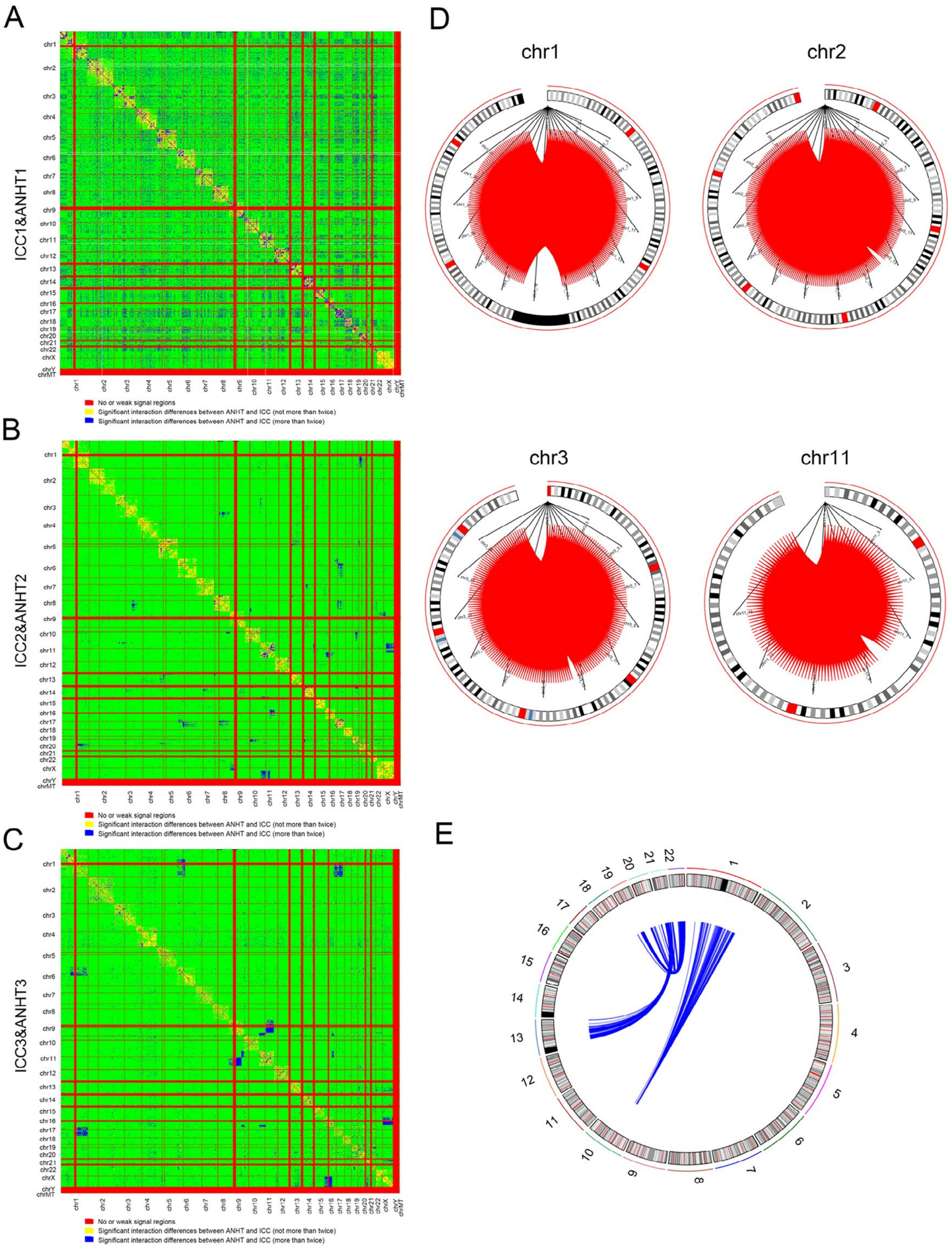

Fig. 2 Chromosomal interactions are significant different between ANHT and ICC. **A**–**C** Genome-wide heatmap of significant differential interactions between ANHT and ICC. Regions with no signal or weak signal value filtered out in the heat map are marked in red (Usually it is a region with a high error rate or contains more repetitive sequences, which appears near the telomeres and centromeres of the chromosomes), regions with significant differences in the probability of interaction between ANHT and ICC but not more than twice are marked with yellow, regions with significant differences and more than twice are marked with dark blue. "chrMT" represents mitochondrial DNA. **D**, **E** Circos diagram made by intrachromosomal, and interchromosomal differential interaction information. The outermost circle represents the chromosomal position coordinates, and the small rectangles on the outermost circle represent different bins, with their colors assigned randomly. **D** the red lines in the inner circle represent interactions occurring at corresponding positions within the same chromosome. The black thin solid lines emanating from the center have no actual significance. **E** the blue lines in the inner circle represent interactions occurring between different regions of different chromosomes. The black thin solid lines in the interior have no actual significance

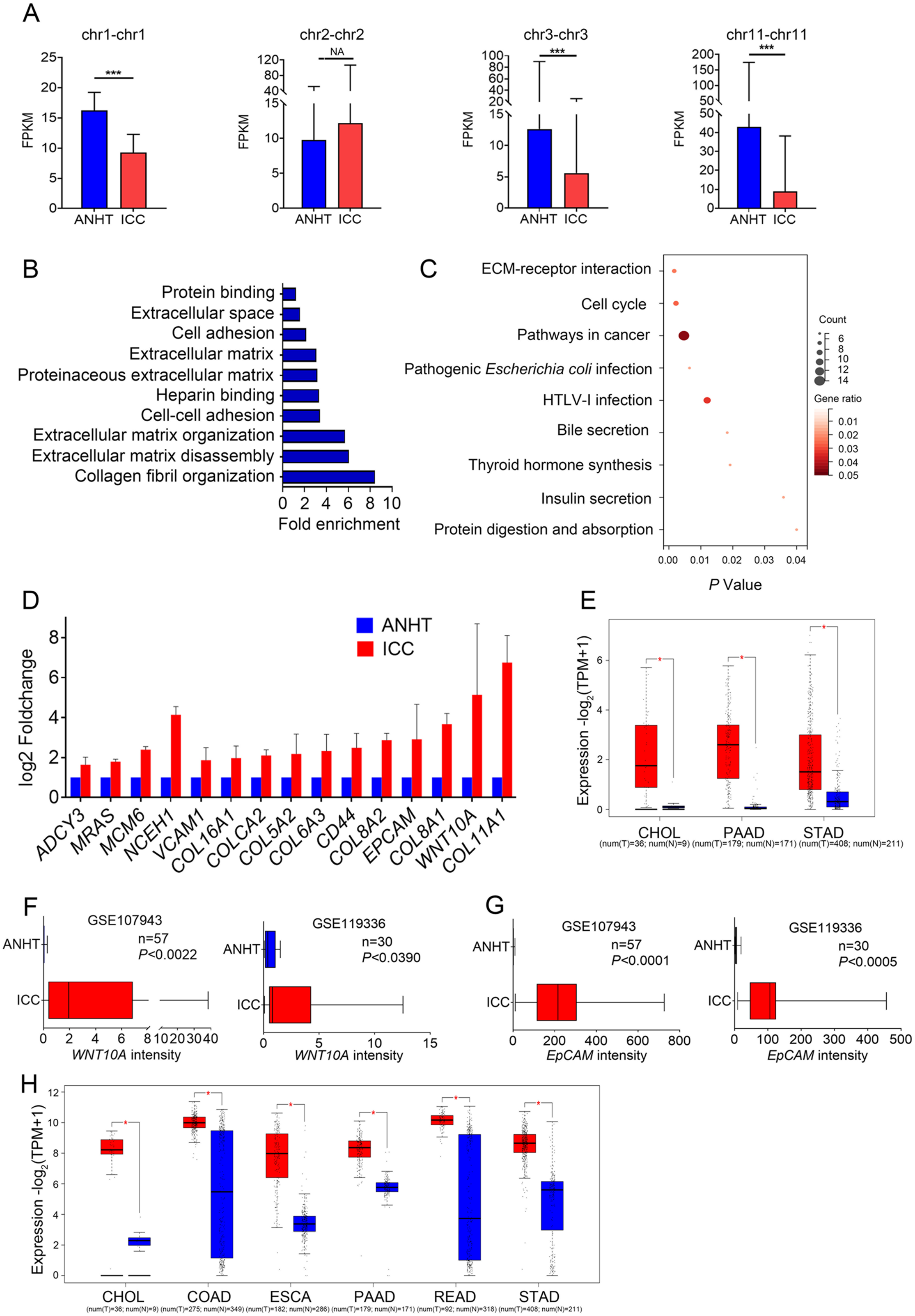
A
chr1-chr1
chr2-chr2
chr3-chr3
chr11-chr11
FPKM
***
NA
***
***
ANHT ICC
B
Protein binding
Extracellular space
Cell adhesion
Extracellular matrix
Proteinaceous extracellular matrix
Heparin binding
Cell-cell adhesion
Extracellular matrix organization
Extracellular matrix disassembly
Collagen fibril organization
Fold enrichment
C
ECM-receptor interaction
Cell cycle
Pathways in cancer
Pathogenic *Escherichia coli* infection
HTLV-I infection
Bile secretion
Thyroid hormone synthesis
Insulin secretion
Protein digestion and absorption
Count
Gene ratio
P Value
D
ANHT
ICC
log2 Foldchange
ADCY3
MRAS
MCM6
NCEH1
VCAM1
COL16A1
COLCA2
COL5A2
COL6A3
CD44
COL8A2
EPCAM
COL8A1
WNT10A
COL11A1
E
Expression -log₂(TPM+1)
CHOL
PAAD
STAD
(num(T)=36; num(N)=9) (num(T)=179; num(N)=171) (num(T)=408; num(N)=211)
F
GSE107943
n=57
P<0.0022
GSE119336
n=30
P<0.0390
WNT10A intensity
G
GSE107943
n=57
P<0.0001
GSE119336
n=30
P<0.0005
EpCAM intensity
H
Expression -log₂(TPM+1)
CHOL
COAD
ESCA
PAAD
READ
STAD
(num(T)=36; num(N)=9) (num(T)=275; num(N)=349) (num(T)=182; num(N)=286) (num(T)=179; num(N)=171) (num(T)=92; num(N)=318) (num(T)=408; num(N)=211)

Fig. 3 Intrachromosomal interactions activates cell migration signaling in ICC. **A** The expression levels of all the genes locating on the differential intrachromosomal regions. **B** Gene ontology (GO) of the significantly up-regulated genes in the differential intrachromosomal regions. GO results indicated that these genes were mainly enriched in collagen fibril organization, extracellular matrix disassembly, cell-cell adhesion and cell adhesion. **C** KEGG analysis of the significantly up-regulated genes in the differential intrachromosomal regions. KEGG analysis indicated that these genes are mainly involved in ECM receptor interaction, protein digestion and absorption, and Cell Cycle. **D** *WNT10A, EpCAM, COL5A2* and *COL16A1* is up-regulated in ICC compared to ANHT. **E** According to TCGA database, the expression of *WNT10A* was elevated in CHOL, PAAD and STAD. **F** The expression of *WNT10A* was analyzed according to the data of GSE107943 and GSE119336. **G** The expression of *EpCAM* was analyzed according to the data of GSE107943 and GSE119336. **H** According to TCGA database, the expression of *EpCAM* was elevated in multiple tumor types

were significantly up-regulated in ICC when compared to ANHT (Fig. 3D), which located in the different chromosomal interaction regions in three pairs of ICCs and ANHTs (Table S9). According to data from The Cancer Genome Atlas (TCGA), *WNT10A* is upregulated in cholangiocarcinoma (CHOL) and overexpressed in pancreatic adenocarcinoma (PAAD) and stomach adenocarcinoma (STAD) (Fig. 3E), and *EpCAM* is upregulated in CHOL and overexpressed in colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), PAAD, Rectum adenocarcinoma (READ) and STAD (Fig. 3H). Additionally, GEO datasets reveal higher expression levels of *EpCAM* and *WNT10A* in ICC compared to ANHT in both Ahn's cohort (GSE107943) and Cao's cohort (GSE119336) (Fig. 3F, G).

COL5A2 was upregulated in CHOL and overexpressed in ESCA, PAAD and STAD (Fig. S4A), *COL16A1* is upregulated in CHOL and overexpressed in lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), head and neck squamous cell carcinoma (HNSC), PAAD and thymoma (THYM) (Fig. S4B). Furthermore, GEO datasets reveals that the expression of *COL5A2* and *COL16A1* were higher in ICC than in ANHT, as observed in both Ahn's cohort (GSE107943) and Cao's cohort (GSE119336) (Fig. S4C–D). Taken together, alterations in intrachromosomal interactions may facilitate the activation of cellular migration signals via the upregulation of *EpCAM, WNT10A, COL5A2*, and *COL16A1* in ICC.

3.4 Interchromosomal interactions enhances the adhesion and infiltration of ICC

The study also analyzed the differentially expressed genes in differential interchromosomal interaction regions between chr1–chr10, chr13–chr21, chr16–chr19 and chr19–chr22 in ICC compared to ANHT. Interestingly, the average expression levels of genes located in differential interchromosomal interaction regions between chr1–chr10, chr13–chr21 and chr16–chr19 were significantly lower in ICCs than in ANHTs (Fig. 4A). However, the average expression levels of genes located in differential interchromosomal interaction regions between chr19–chr22 were significantly higher in ICC than in ANHT (Fig. 4A).

Additionally, the commonly up-regulated and down-regulated genes in differential intrachromosomal interaction regions on chr1–chr10, chr13–chr21, chr16–chr19, and chr19–chr22 in ICC were also identified (Table S10). Among the GO terms, "RAGE receptor binding", "Protein binding" and "Calcium ion binding" were significantly enriched (Fig. 4B, Table S10). Similarly, "TNF signaling pathway" and "Leukocyte trans endothelial migration" were significantly enriched in KEGG analysis (Fig. 4C, Table S10). Above results suggested that interchromosomal interactions could potentially lead to the dysfunction of cell adhesion and infiltration in ICC.

Among the above differentially expressed genes enriched in cell migration and invasion related pathways, *S100A3* (chr1: 153,519,809–153,521,734), *S100A6* (chr1: 153,507,076–153,508,717) and *MAPK12* (chr22: 50,688,492–50,700,089) were significantly up-regulated in ICC when compared to ANHT (Fig. 4D), which located in the distinct chromosomal interaction regions in three pairs of ICCs and ANHTs tissues (Table S11). According to corroborated data from the TCGA database, *S100A3* is consistently upregulated in CHOL and overexpressed in ESCA and PAAD (Fig. 4E). Furthermore, an examination of approximately 80 ICC and ANHT from Ahn's (GSE107943) and Cao's (GSE119336) cohorts in the GEO database confirms the significant differential expression of these genes in ICC in comparison to ANHT (Fig. 4F). which were also verified by qRT-PCR (Fig. 4G).

S100A6 can regulate the proliferation, invasion, migration, and angiogenesis of lung cancer cells. According to data from the TCGA database, *S100A6* is upregulated in CHOL and overexpressed in COAD, LIHC, PAAD, READ and STAD (Fig. 4H). Moreover, an analysis of approximately 80 ICC and ANHT from Ahn's (GSE107943) and Cao's (GSE119336) cohorts in the GEO database confirms the significant differential expression of these genes in ICC compared to ANHT (Fig. 4I). *MAPK12* is upregulated in CHOL and overexpressed in CESC, COAD, PAAD and READ (Fig. 4J). Furthermore, GEO database and qRT-PCR also verified the highly expressino of *S100A6* in ICC (Fig. 4K, L). Collectively, alterations in interchromosomal interactions may activate adhesive and invasive properties through the heightened expression of *S100A3, S100A6* and *MAPK12* in ICC.

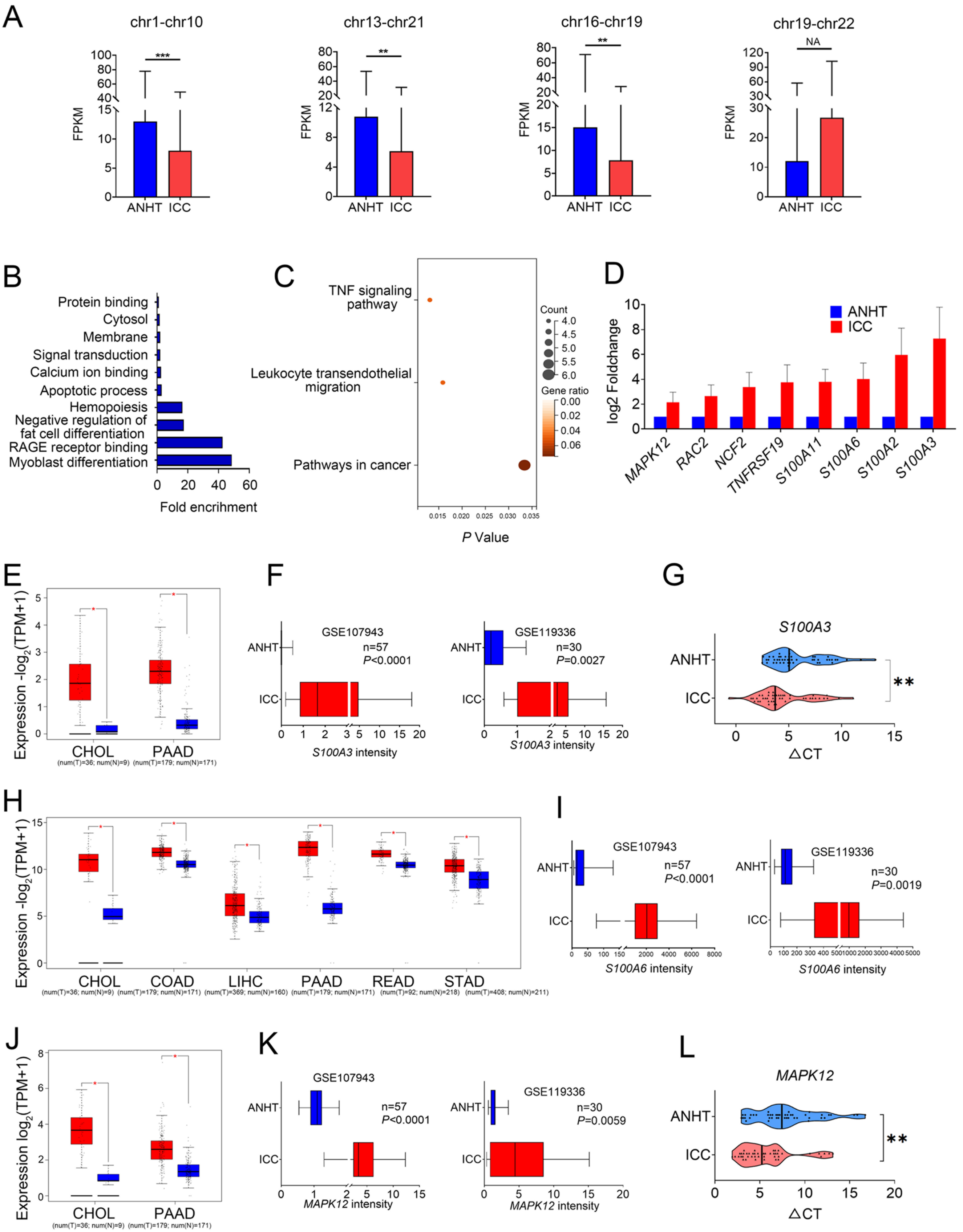
A
chr1-chr10
chr13-chr21
chr16-chr19
chr19-chr22
***
**
**
NA
FPKM
ANHT ICC
B
Protein binding
Cytosol
Membrane
Signal transduction
Calcium ion binding
Apoptotic process
Hemopoiesis
Negative regulation of fat cell differentiation
RAGE receptor binding
Myoblast differentiation
Fold encrihment
C
TNF signaling pathway
Leukocyte transendothelial migration
Pathways in cancer
Count
Gene ratio
P Value
D
log2 Foldchange
ANHT
ICC
MAPK12
RAC2
NCF2
TNFRSF19
S100A11
S100A6
S100A2
S100A3
E
Expression -log2(TPM+1)
CHOL
PAAD
(num(T)=36; num(N)=9)
(num(T)=179; num(N)=171)
F
GSE107943
n=57
P<0.0001
GSE119336
n=30
P=0.0027
ANHT
ICC
S100A3 intensity
G
S100A3
**
△CT
H
CHOL
COAD
LIHC
PAAD
READ
STAD
(num(T)=36; num(N)=9)
(num(T)=179; num(N)=171)
(num(T)=369; num(N)=160)
(num(T)=179; num(N)=171)
(num(T)=92; num(N)=218)
(num(T)=408; num(N)=211)
I
GSE107943
n=57
P<0.0001
GSE119336
n=30
P=0.0019
S100A6 intensity
J
Expression log2(TPM+1)
CHOL
PAAD
K
GSE107943
n=57
P<0.0001
GSE119336
n=30
P=0.0059
MAPK12 intensity
L
MAPK12
**
△CT

Fig. 4 Interchromosomal interactions enhances the adhesion and infiltration of ICC. **A** The expression levels of all the genes locating on the differential interchromosomal regions. **B** GO of the significantly up-regulated genes in the differential interchromosomal regions. GO results indicated that these genes were mainly enriched in RAGE receptor binding, Protein binding and Calcium ion binding. **C** KEGG analysis of the significantly up-regulated genes in the differential interchromosomal regions. KEGG analysis indicated that these genes are mainly involved in TNF signaling pathway, and "Leukocyte trans endothelial migration. **D** *S100A3, S100A6* and *MAPK12* is up-regulated in ICC compared to ANHT. **E** According to TCGA database, the expression of *S100A3* is elevated in CHOL and PAAD. **F** The expression of *S100A3* was analyzed according to the data of GSE107943 and GSE119336. **G** The mRNA expression levels of *S100A3* in ANHT and ICCs were validated by qRT-PCR (n=20), Student *t* test. **H** According to TCGA database, the expression of *S100A3* is elevated in multiple tumor types. **I** The expression of *S100A6* was analyzed according to the data of GSE107943 and GSE119336. **J** According to TCGA database, the expression of *MAPK12* is elevated in CHOL and PAAD. **K** The expression of *MAPK12* was analyzed according to the data of GSE107943 and GSE119336. **L** The mRNA expression levels of *MAPK12* in ANHT and ICCs were validated by qRT-PCR (n=20), Student *t* test. Data are presented as mean ± SD. *$P<0.05$; **$P<0.01$

3.5 Compartment switch inhibits the complement pathway in ICC

A compartment being primarily composed of transcriptionally active euchromatin, while B compartment are characterized by high chromatin density and inactive heterochromatin [20–22]. Compared to ANHT, the majority of compartments in ICC remained unchanged. However, 7.90 to 12.35% of compartments switched from one compartment (A or B) to the other in ICCs (Fig. 5A). The certain regions occurring the A to B or B to A transition overlapped in three paired ANHT and ICC were also identified. In comparison to ANHT, 20 regions from 11 chromosomes underwent A to B transitions, while 36 regions from 12 chromosomes underwent B to A transitions (Table S12). The representative overlapped A to B transitions located on chr1, chr2 and chr5 and chr13 were displayed (Fig. 5B). Significantly overexpressed genes (*CD44, E2F3, PFKP, HOPX*) were identified in regions B to A (Fig. 5C). TCGA data supported their consistent upregulation in CHOL (Fig. 5D). Analysis of ICC and ANHT from Ahn's (GSE107943) and Cao's (GSE119336) cohorts in the GEO database confirmed their significant differential expression (Fig. 5E, F).

To further characterize the relationship between genome-wide different spatial compartmentalization and gene expression, the up-regulated genes located on the B-A transition regions and down-regulated genes located on the A-B transition regions were screened (Table S13). GO analysis indicated that above dysregulated genes was significantly enriched in "regulation of complement activation" (Fig. 6A) and KEGG analysis showed that above dysregulated genes was remarkably enriched in "complement and coagulation cascades" (Fig. 6B). Surprisingly, the dysregulated genes mediated by differential intrachromosomal interactions (Fig. S5A–B) and differential interchromosomal interactions (Fig. S5C–D) were also enriched in the regulation of complement activation. Above results might suggest that compartment switch correlated with the dysregulation of complement system in ICC.

Among all the dysregulated genes resulted from compartment switch (Figs. 5B, 6C), *C8A* (chr1:57,320,443–57,383,894), *C8B* (chr1:57,394,883–57,431,688), *F7* (chr13:113,760,102–113,774,995), *F10* (chr13:113,777,113–113,803,843) and *F13B* (chr1:197,008,321–197,036,397) located in compartment A in ANHT and are observed to switch into compartment B in ICC. This transition could potentially decrease the expression of these genes (Fig. 6C). More importantly, they are all typical genes participating in the complement pathway, indicating the inhibition of the complement function during ICC oncogenesis. Based on clinical and genomic data from (TCGA), it was observed that *C8A, C8B, F7, F10*, and *F13B* exhibited low expression in CHOL (Fig. 6D). Moreover, in Ahn's (GSE107943) and Cao's (GSE119336) cohorts from the GEO database, the expression levels of these genes were lower in ICC compared to ANHT (Fig. 6E, F). In summary, compartment switch in ICC compared with ANHT may weaken complement pathways by downregulating *C8A, C8B, F7, F10,* and *F13B.*

3.6 Identification of alternations in TADs and related DEGs in ICC

Topologically associating domains (TADs) are functional units within the sub-megabase range that play a crucial role in regulating gene expression and often coincide with replication domains [20, 23]. TADs function by constraining interactions between cis-regulatory sequences and their target genes [24, 25]. Chromatin loops within TADs facilitate long-range interactions between enhancers and promoters, thereby modulating gene expression [26]. The genome of ANHTs was composed of 83.98% TADs, 11.64% unorganized chromatin, and 4.37% TAD boundaries, while the genome of ICCs was composed of 83.51% TADs, 12.33% unorganized chromatin, and 4.16% TAD boundaries (Fig. S6). The TAD location maps of ANHT and ICC were merged, and TAD structure alterations on each chromosome were visually compared (Fig. 7A–C). Detailed examination of TAD boundaries, it was revealed that numerous large TADs present in ANHTs were actually composed of smaller domains in ICCs (Fig. 7A–C).

A total of 1324 TADs with a mean size of 1.96 Mb were identified in ANHTs, while 1235 TADs with a mean size of 2.05 Mb were identified in ICCs (Fig. S7A–B, Table S14). However, no significant difference was observed in the

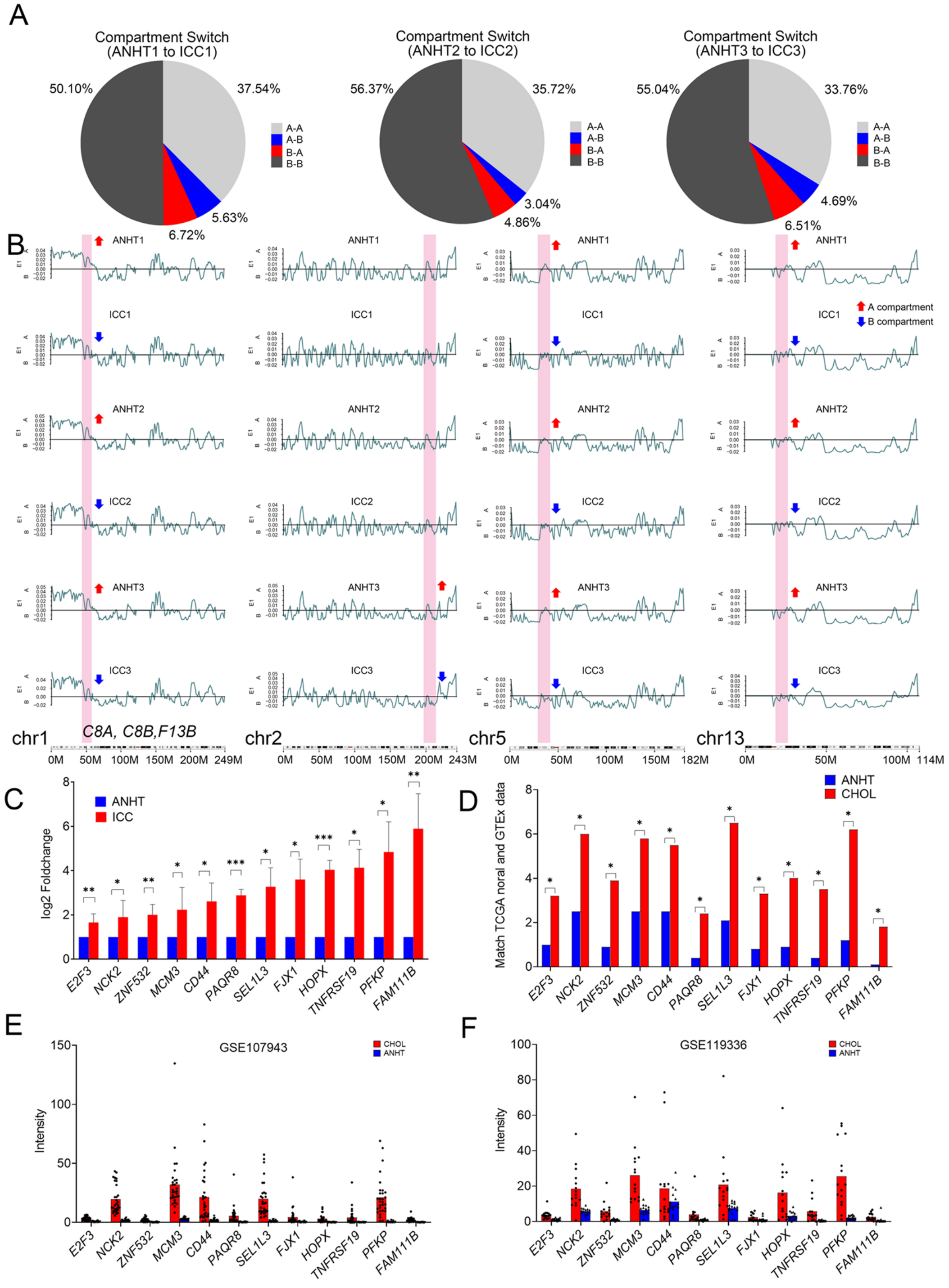

Fig. 5 Compartment reorganization and up-regulated gene in ICC. **A** Pie chart showing the genomic compartment changes between ANHT and HCC genomes. "A" and "B" denote the open and closed compartments, respectively. "A-A" represents compartments that are open in both tissues, "B-B" represents compartments that are closed in both tissues, "A-B" denotes compartments that are open in ANHT but closed in ICC, and "B-A" denotes compartments that are closed in ANHT and open in ICC. **B** Highlighted bars represent examples of regions with differential compartmentalization. The differential compartments are defined as genomic regions in which one type of compartmentalization is observed in ANHT and the other compartment type in ICC. Open type A (E1 > 0, red arrow) and closed type B (E1 < 0, blue arrow). **C** In RNA-seq analysis, the expression of these genes was up-regulated in ICC samples compared to ANHT samples. **D** The analysis of the expression of these genes from TCGA database. **E**, **F** The expression of these genes was analyzed according to the data from GSE107943 and GSE119336

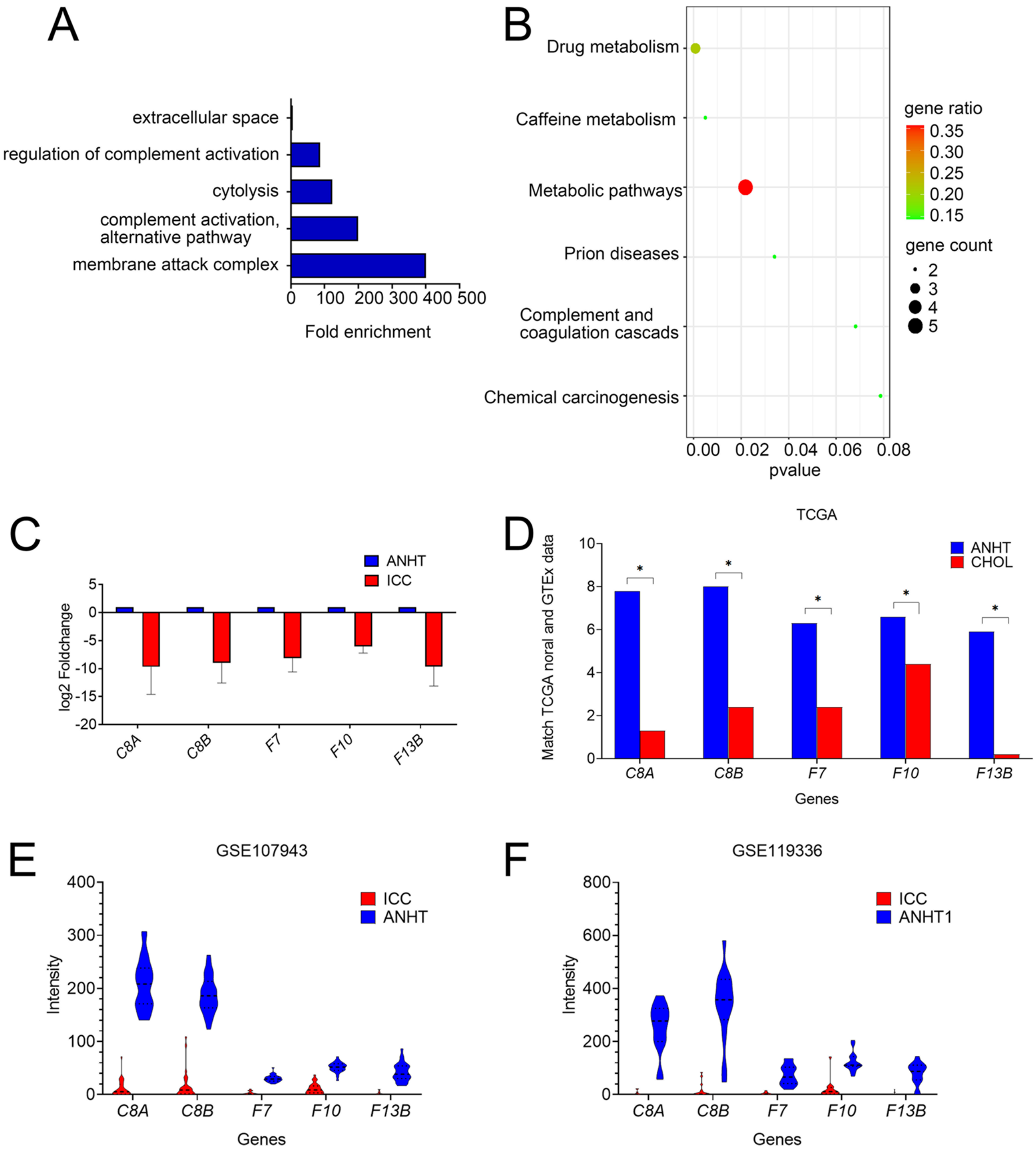

Fig. 6 Compartment switch inhibits the complement pathway in ICC. **A** GO of the significantly up-regulated genes in compartment switch region. Among the GO terms, "regulation of complement activation" and "complement activation, alternative pathway". **B** significantly enriched. "complement and coagulation cascades" was significantly enriched in KEGG analysis. **C** The expression of all complement pathway genes on the A/B transition regions. In ICC, *C8A, C8B, F7, F10* and *F13B* are down-regulated. **D** According to TCGA database, the expression of these genes was down-regulated in CHOL. **E**, **F** The expression of these genes was analyzed according to the data from GSE107943 and GSE119336

number and size of TADs between ANHTs and ICCs. Additionally, an average of 1925 TAD boundaries were detected in ANHTs and 1908.67 TAD boundaries were detected in ICCs (Fig. S7C). The TAD boundaries that were unique to ICCs were referred to as tumor-specific TAD boundaries (TSTADBs) in comparison to ANHTs (Table S15). Obviously, the large TADs present in ICCs were composed of two smaller TADs in ANHTs, which contained *HOPX* (chr4:57,514,154–57,522,688) and *ACVR1* (chr2:158,592,958–158,732,374) gene. Multiple transcription factor binding sites/motifs and promoters were scattered over the region of *HOPX* and *ACVR1* with TAD rearrangement. Mechanically, larger TAD boundaries were examined in the regions containing *HOPX* and *ACVR1* in ICC2 tissues compared to

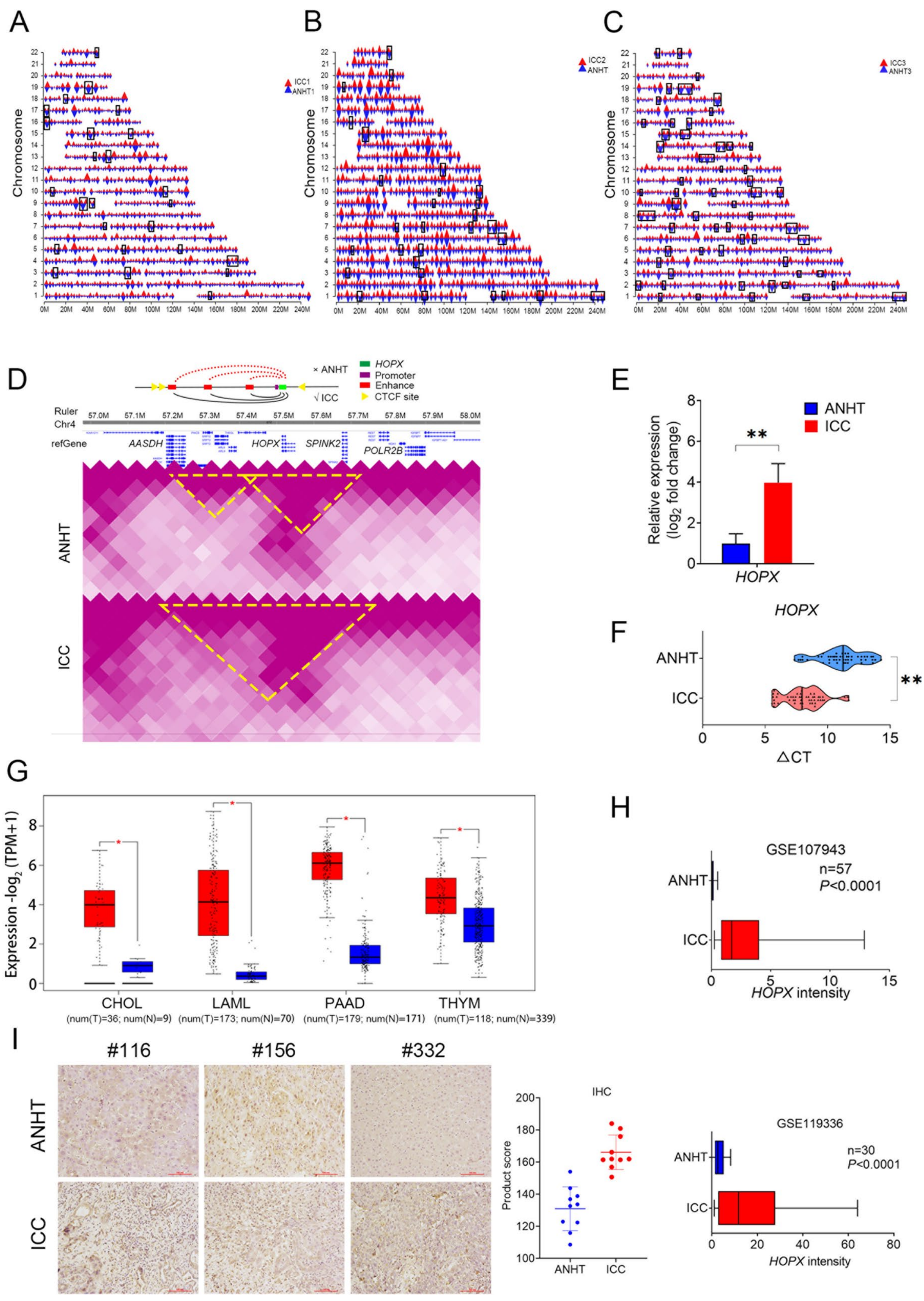
A
B
C
Chromosome
ICC1
ANHT1
ICC2
ANHT
ICC3
ANHT3
0M 20M 40M 60M 80M 100M 120M 140M 160M 180M 200M 220M 240M
D
× ANHT
√ ICC
HOPX
Promoter
Enhance
CTCF site
Ruler
Chr4
57.0M 57.1M 57.2M 57.3M 57.4M 57.5M 57.6M 57.7M 57.8M 57.9M 58.0M
refGene
AASDH
HOPX
SPINK2
POLR2B
ANHT
ICC
E
Relative expression (log2 fold change)
ANHT
ICC
**
HOPX
F
HOPX
ANHT
ICC
**
△CT
G
Expression -log2 (TPM+1)
CHOL (num(T)=36; num(N)=9)
LAML (num(T)=173; num(N)=70)
PAAD (num(T)=179; num(N)=171)
THYM (num(T)=118; num(N)=339)
H
GSE107943
n=57
P<0.0001
ANHT
ICC
HOPX intensity
I
#116
#156
#332
ANHT
ICC
IHC
Product score
ANHT
ICC
GSE119336
n=30
P<0.0001
ANHT
ICC
HOPX intensity

Fig. 7 Fusion of topologically associated domain boundaries upregulates the expression of *HOPX* in ICC. **A**–**C** Whole gene TAD distribution map of three pairs of ANHT and ICC. The large TADs in ANHT (represented by blue triangles) are divided into several small TADs in ICC (represented by red triangles). The areas of variation between the two are marked with black borders. **D** The tumor-specific TAD boundary located on chr4: 57,514,154–57,522,688 was visualized in a two-dimensional interaction matrix via WashU browser. With yellow triangles representing CTCF binding sites, and red, purple and green squares representing enhancers, promoters, and targeted genes, respectively. The black lines indicate normal interactions within the same TAD boundary, while the red lines indicate failed interactions that are isolated by different TAD boundaries. **E** In RNA-seq analysis, the expression of *HOPX* was up-regulated in ICC compared to ANHT. **F** The mRNA expression levels of *HOPX* in ANHT and ICCs were validated by qRT-PCR ($n=20$), Student *t* test. **G** The expression of *HOPX* was not only highly expressed in CHOL, but also in LAML, PAAD and THYM compared to that of corresponding normal tissues. **H** The expression of *HOPX* was higher in ICC than in ANHT in GSE107943 and GSE119336. **I** The expression of HOPX was higher in ICC than that in ANHT as assessed by immunohistochemistry ($n=10$). HOPX staining was measured by multiplying the numerical score of the staining intensity (none = 1, weak = 2, moderate = 3, strong = 4) with the staining percentage (0–100%), resulting in an overall product score, Student *t* test. Data are presented as mean ± SD. $*P<0.05$; $**P<0.01$

ANHT2 tissues. This allowed for the hijacking of remote enhancers that were previously outside the same TAD range through TAD recombination and CTCF binding, correlated with increased expression of *HOPX* (Fig. 7D, Tables S16–17). During RNA-seq analysis, it was observed that *HOPX* was up-regulated in ICC compared to ANHT (Fig. 7E). The TCGA data indicated that *HOPX* was highly expressed not only in CHOL, but also in acute myeloid leukemia (LAML), PAAD and THYM when compared to the corresponding normal tissues (Fig. 7G). Furthermore, the expression of *HOPX* was found to be higher in ICC than in ANHT in GSE107943 and GSE119336 based on the analysis of GEO datasets (Fig. 7H). Furthermore, qRT-PCR and IHC experiments verified that the mRNA and protein expression levels of *HOPX* were also significantly elevated in ICC than those in ANHT (Fig. 7F, I).

Furthermore, enhancer hijacking were also found in the chr2: 158,100,000–158,800,000 area (Fig. S8A). During RNA-seq analysis, it was observed that *ACVR1* was up-regulated in ICC compared to ANHT (Fig. S8B). The TCGA data indicated that *ACVR1* was highly expressed not only in CHOL, but also in ESCA, PAAD, etc. when compared to the corresponding normal tissues (Fig. S8C). Based on the analysis of GEO datasets, the expression of *ACVR1* was higher in ICC than those in ANHT in GSE107943 and GSE119336 (Fig. S8D). Furthermore, qRT-PCR and IHC experiments confirmed that the mRNA and protein expression levels of *ACVR1* were significantly higher in ICC than in ANHT (Fig. S8E–F). These findings suggest that ICC development is linked to changes in transcriptional profiles correlated with the formation of new TAD boundaries and the disappearance of original ones. Taken together, TADs alternations promote the expression of *HOPX* and *ACVR1* in ICC via enhancer hijacking, which may underscore the importance of genomic changes during the tumorigenesis of ICC.

3.7 Discussion

ICC is a subtype of bile duct adenocarcinoma that affecting small ducts of the liver [27]. It ranks as the second most prevalent primary liver malignancy after HCC [28], and is associated with a poor prognosis due to the scarcity of diagnostic markers and treatments options [4]. However, there is a scarcity of research investigating the comprehensive genomic chromatin landscape in ICC patients. This study aims to elucidate the 3D genome architecture and molecular mechanisms underlying ICC using Hi-C and RNA-seq.

Firstly, the differential intrachromosomal interactions of chr1, chr2, chr3, and chr11 resulted in the dysregulation of genes within these regions. These genes were mainly enriched in cell migration and invasion related pathways, which were commonly implicated tumorigenesis [29]. *WNT10A*, a member of the WNT/β-catenin signaling pathway, enhances the transcription of genes associated with cell proliferation and cancer progression. Certain *WNT10A* variants are associated with an increased risk of colorectal adenoma in malignancy [30, 31]. *EpCAM* has the ability to repress contractility of the actomyosin cell cortex, and directly inhibits the novel PKC family members and specific downstream PKD-Erk cascade [32]. In Zhou's study, *EpCAM* overexpression in HCC was associated with poorer clinicopathological features and predicted shorter OS and DFS [33]. *COL16A1* act as a substrate for the adhesion and invasion of connective tissue tumor cells. In glioblastoma, it induces tumor invasiveness by modifying the activation pattern of β1-integrin [34]. Yuen Tan's research indicates that high expression of *COL5A2* is associated with reduced survival and enhanced cell migration in gastric cancer patients [35]. However, the relationship between *WNT10A, EpCAM, COL5A2, COL16A1*, and ICC remains uncertain. Our study suggests that these genes are all up-regulated in ICC as a result intrachromosomal interaction. In summary, intra/interchromosomal interactions may contribute to the dysregulation of cell migration and invasion pathways in ICC, potentially leading to an increase in the expression of *WNT10A, EpCAM, COL5A2,* and *COL16A1.*

Secondly, the differential interchromosomal interactions between chr1–chr10, chr13–chr21, chr16–chr19 and chr19–chr22 also resulted in the dysregulation of genes within these regions. These genes were primarily enriched in adhesion and infiltration pathways, which have been demonstrated to be closely associated with ICC pathogenesis. The S100 protein family is involved in diverse biological activities, encompassing cell invasion, migration, proliferation, differentiation, apoptosis, and calcium homeostasis. Our study suggests that *S100A3* activation is implicated in tumorigenesis and tumor aggressiveness. In Tao's study, *S100A3* was found to be crucial for the development and growth of HCC, and its mRNA and protein expression levels were significantly elevated in HCC [36]. Recently, Mahmood et al. demonstrated that *S100A3* might represent a new target for combating HBV pathogenesis [37]. *S100A6* has been studied in HCC and ICC, showing its role in promoting tumor growth and migration [38, 39]. In HCC, *S100A6* up-regulation is associated with increased cell proliferation and motility through the activation of the PI3K/AKT pathway and down-regulation of E-cadherin, contributing to enhanced β-catenin nuclear accumulation [40]. Similarly, overexpression of *S100A6* promotes cell proliferation and p38/MAPK activity in ICC, while its silencing suppresses proliferation via the inhibition of p38/MAPK. Overall, *S100A6* is implicated in the aggressive behavior of ICC cells [41]. Over the past decade, *S100A3/A6* has emerged as a significant biomarker for tumor progression. However, the impact of interchromosomal interactions on *S100A3/A6* remained unexplored. In our study, we confirmed for the first time the variation of *S100A3/A6* in the interaction between ICC 3D genome chromosomes. Remarkably, it appears that *S100A3/A6* exhibits higher expression levels in ICC compared to ANHT, potentially due to interchromosomal interactions. Additionally, *MAPK12*, known for its role in promoting tumorigenesis in various tissues, is being recognized as a potential prognostic and immunotherapeutic target in several cancers, including thyroid carcinoma (THCA) [42, 43]. Nevertheless, its specific involvement in ICC remains not entirely clear. Our study indicates an up-regulation of *MAPK12* in ICC, attributed to intrachromosomal interactions. This observation positions *MAPK12* as a candidate deserving further investigation in the context of ICC.

Thirdly, significantly overexpressed genes (including *CD44, E2F3, PFKP* and *HOPX*) were screened from the B to A compartments, all of which have been previously reported to be associated with the occurrence and development of cancer [44–47], thus validating the efficacy of this research method. In ANHT and ICC, a relatively small proportion of compartments underwent transitions between A and B compartments in ICCs, ranging from 7.90 to 12.35%, while the majority remained stable. Chr1, chr2, chr5 and chr13 show the highest number of AB compartment transitions. Excitingly, we found enrichment of complement system-related functions involving *C8A* and *C8B* at chromosome and Compartment levels. C8, a crucial part of the complement system, interacts with *C5B, C6, C7*, and *C9* to form the MAC *C5B-9*, causing membrane damage [48, 49]. Our study identifies two *C8* components affected by compartment transitions, potentially impacting MAC formation during complement activation. According to literature reports, *C8A* is highly expressed in prostate cancer and correlates with improved trastuzumab efficacy in HER$^+$ breast cancer patients [50, 51]. High expression of *C8B* benefits HBV-HCC patients with improved overall survival and RFS [52]. "Complement activation" is a vital defense mechanism against cancer invasion and progression. Nevertheless, inhibition of the complement system allows tumor cells to evade complement-mediated elimination, thereby reducing the clinical efficacy of tumor immunotherapy [53]. Lubka's study revealed that malignant cells within the tumor microenvironment produce various complement components, leading to altered complement activation [54]. Notably, high levels of complement regulators on malignant cells inhibit the terminal pathway. Interestingly, *C8A, C8B*, and *C9* genes show remarkably low expression in most tumor types, except for cholangiocarcinoma, suggesting limited activation of the terminal pathway by locally produced components. This unique combination of gene expression supports the idea that malignant cells adapt to evade potential cell lysis caused by the membrane attack complex (MAC) formation. The study revealed a significant reduction in coagulation factors *F7, F10* and *F13B* in ICC, correlating with reduced *C8A/B* levels in the complement system. Previous research has established the potential involvement of the complement and coagulation cascades in disease development [55, 56], yet their specific connection with ICC remained unclear. In a two-year investigation involving 165 cirrhotic patients, Violi et al. observed that non-survivors exhibited lower *F7* levels, and individuals with less than 34% activity faced a staggering 93% mortality rate within ten months [57]. Furthermore, Frank et al.'s six-year study of 111 patients with chronic liver disease unveiled a significant rise in the risk of severe upper gastrointestinal bleeding and heightened mortality among those with *F13* activity below 50%, in contrast to the normal range of 50–75% [58]. Based on the ANHT derived from hepatocytes serving as the control group, we have proposed the potential significance of key genes in the complement and coagulation cascade pathways in the development of ICC.

Finally, we found that several smaller TADs from ANHT tend to merge into larger TADs in ICC, suggesting a significant chromatin reorganization during tumorigenesis. Additionally, we identified two highly expressed genes in ICC, *HOPX* and *ACVR1*. Mechanically, our findings revealed that TAD recombination and CTCF binding in the region of *HOPX* and *ACVR1* enabled the hijacking of remote enhancers that were previously beyond their TAD range. This hijacking mechanism appears to increase the expression of *HOPX* and *ACVR1*, potentially suggesting a contribution to the oncogenic potential of ICC. Building upon these insights, literature supports similar scenarios in other malignancies. Liu et al. identified enhancer co-amplification and hijacking events in liposarcoma, leading to overexpression of oncogenes *MDM4, CDK2*, and *HMGA3*, and promoting liposarcoma development and progression [59]. Wei et al. identified *HOPX, TBX15* and *TFCP2L1* as master regulators critical for cholangiocyte-to-hepatocyte differentiation [60]. Additionally, Zheng et al. discovered that *CD4* neoantigen-reactive tumor-infiltrating lymphocytes (TILs) are enriched in bile duct cancer cells expressing *HOPX* or *ADGRG1* [61]. Wu et al. discovered that 32% of DIPGs exhibit clonal activating mutations and overexpression of *ACVR1*, providing robust evidence that *ACVR1* acts as an oncogenic driver in this disease [62]. However, further investigation on the functional and mechanistic relationship between altered 3D genome structure or gene expression and ICC oncogenesis are needed.

4 Conclusion

In this study, we have successfully revealed the 3D genome landscapes and higher-order chromatin organization alterations in primary ANHT and ICC using Hi-C and RNA-seq for the first time. These alterations involve intra/inter-chromosomal interactions, compartment switching, and TAD rearrangement, which may contribute to the dysregulation of *WNT10A, EpCAM, S100A3*/6, *MAPK12, C8A*/*B, HOPX* and *ACVR1* in ICC. These genes play essential roles in pathways related to cell migration and invasion pathways, and the complement system. These findings offer a valuable genomic resource for investigating 3D genomic variations during the progression of ICC. However, the newly found structural changes and biomarkers have potential implications for ICC treatment and prognosis, which need further investigation on more ICC samples with different stages and grades, compared with noncancerous hepatic tissue and biliary tissues.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s13402-024-01033-6 .

Acknowledgements We thank Fifth Medical Center of Chinese PLA General Hospital who donated samples used in this study. We thank all patients involved in this study. We thank Jian Dong (Beijing University of Chemical Technology), members of the First Medical Center of Chinese PLA General Hospital, Fifth Medical Center of Chinese PLA General Hospital and the Second Affiliated Hospital of Kunming Medical University for the all the work and effort involved.

Author contributions P.Y., L.W. and Z.Y. designed the experiments. P.Y., Z.W., C.L., Z. W., W. Z., H.Y., Y. Z., S. F., L. G., S. L., G.L., F.S., Y.Z., Y.T., and L.Y. prepared sample. Y.L. X.C., W.Z., J.D., T.Y., M.H., R.G. K. Z., and K. Z performed experiments and analyzed data. Y.L., C.L. and R.Z. analyzed data. Y.L. wrote the paper. L.W., C.W., L.W., Z.Y., and P.Y. revised the paper. Z.Y., L.W. and P.Y. initiated the study and organized, designed, and revised the paper.

Funding This work was supported by the Major Research Plan of the National Natural Science Foundation of China (92359202), Scientific and Technological Research Project of Xinjiang Production and Construction Corps (2022AB022), Open Competition to Select the Best Candidates" Key Technology Program for Nucleic Acid Drugs of NCTIB (NCTIB2022HS01016) and the Joint Project of Biomedical Translational Engineering Research Center of Beijing University of Chemical Technology-China-Japan Friendship Hospital (XK2023-21).

Data availability The Hi-C and RNA sequencing data reported in this paper have been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2021) in National Genomics Data Center (Nucleic Acids Res 2022), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: HRA007899) that are publicly accessible at https://ngdc.cncb.ac.cn/gsa-human/submit/hra/submit. Source code and documentation are available at http://github.com/nservant/HiC-Pro.

Declarations

Ethics approval and consent to participate Informed consent signed by all patients. The protocols used in this paper were approved by the Research Ethics Board at Fifth Medical Center of Chinese PLA General Hospital (S2016-98-02).

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

1. D. Moris, M. Palta, C. Kim, et al., Advances in the treatment of intrahepatic cholangiocarcinoma: an overview of the current and future therapeutic landscape for clinicians. CA Cancer J. Clin. **73**, 198–222 (2023)
2. A. Gupta, E. Dixon, Epidemiology and risk factors: intrahepatic cholangiocarcinoma. Hepatobiliary Surg. Nutr. **6**, 101–104 (2017)
3. J.W. Valle, I. Borbath, S.A. Khan, et al., Biliary cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann. Oncol. **27**, v28–v37 (2016)

4. J.M. Banales, J.J.G. Marin, A. Lamarca, et al., Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nat. Rev. Gastroenterol. Hepatol. **17**, 557–588 (2020)
5. Y. Bekki, D. Von Ahrens, H. Takahashi, et al., Recurrent intrahepatic cholangiocarcinoma - review. Front. Oncol. **11**, 776863 (2021)
6. U. Cillo, C. Fondevila, M. Donadon, et al., Surgery for cholangiocarcinoma. Liver Int. **39**(Suppl 1), 143–155 (2019)
7. J. Wang, H. Tao, H. Li, et al., 3D genomic organization in cancers. Quant. Biol. **11**, 109–121 (2023)
8. A. Kloetgen, P. Thandapani, P. Ntziachristos, et al., Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. Nat. Genet. **52**, 388–400 (2020)
9. M. Spielmann, D.G. Lupiáñez, S. Mundlos, Structural variation in the 3D genome. Nat. Rev. Genet. **19**, 453–467 (2018)
10. D.L. Lafontaine, L. Yang, J. Dekker, et al., Hi-C 3.0: improved protocol for genome-wide chromosome conformation capture. Curr. Protoc. **1**, e198 (2021)
11. E. Lieberman-Aiden, N.L. van Berkum, L. Williams, et al., Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science (New York, NY) **326**, 289–293 (2009)
12. A.R. Barutcu, B.R. Lajoie, R.P. McCord, et al., Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. Genome Biol. **16**, 214 (2015)
13. P. Wu, T. Li, R. Li, et al., 3D genome of multiple myeloma reveals spatial genome disorganization associated with copy number variations. Nat. Commun. **8**, 1937 (2017)
14. J. Zuin, J.R. Dixon, M.I. van der Reijden, et al., Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc. Natl. Acad. Sci. U. S. A. **111**, 996–1001 (2014)
15. Y. Feng, P. Wang, L. Cai, et al., 3D-Epigenomic regulation of gene transcription in hepatocellular carcinoma. Adv. Genet. (Hoboken, NJ) **3**, 2100010 (2022)
16. Y. Zhang, F. Chen, N.A. Fonseca, et al., High-coverage whole-genome analysis of 1220 cancers reveals hundreds of genes deregulated by rearrangement-mediated cis-regulatory alterations. Nat. Commun. **11**, 736 (2020)
17. Z. Yang, M. Shi, Y. Liang, et al., Three-dimensional chromatin landscapes in hepatocellular carcinoma associated with hepatitis B virus. J. Gastroenterol. **59**, 119–137 (2024)
18. D. Xu, R. Ma, J. Zhang, et al., Dynamic nature of CTCF Tandem 11 zinc fingers in multivalent recognition of DNA as revealed by NMR spectroscopy. J. Phys. Chem. Lett. **9**, 4020–4028 (2018)
19. T. Gao, J. Qian, EnhancerAtlas 2.0: an updated resource with enhancer annotation in 586 tissue/cell types across nine species. Nucleic Acids Res. **48**, D58–D64 (2020)
20. J.R. Dixon, S. Selvaraj, F. Yue, et al., Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. **485**, 376–380 (2012)
21. J. Dekker, L. Mirny, The 3D genome as moderator of chromosomal communication. Cell. **164**, 1110–1121 (2016)
22. M. Chen, Q. Zhu, C. Li, et al., Chromatin architecture reorganization in murine somatic cell nuclear transfer embryos. Nat. Commun. **11**, 1813 (2020)
23. B.D. Pope, T. Ryba, V. Dileep, et al., Topologically associating domains are stable units of replication-timing regulation. Nature **515**, 402–405 (2014)
24. E. McArthur, J.A. Capra, Topologically associating domain boundaries that are stable across diverse cell types are evolutionarily constrained and enriched for heritability. Am. J. Hum. Genet. **108**, 269–283 (2021)
25. J.R. Dixon, D.U. Gorkin, R. B, Chromatin domains: the unit of chromosome organization. Mol. Cell **62**, 668–680 (2016)
26. S.S. Rao, M.H. Huntley, N.C. Durand, et al., A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell **159**, 1665–1680 (2014)
27. T. Patel, Cholangiocarcinoma—controversies and challenges. Nat. Rev. Gastroenterol. Hepatol. **8**, 189–200 (2011)
28. W. Wu, X. He, D. Andayani, et al., Pattern of distant extrahepatic metastases in primary liver cancer: a SEER based study. J. Cancer **8**, 2312 (2017)
29. O. Ilina, P.G. Gritsenko, S. Syga, et al., Cell–cell adhesion and 3D matrix confinement determine jamming transitions in breast cancer invasion. Nat. Cell Biol. **22**, 1103–1115 (2020)
30. R.L. Galbraith, E.M. Poole, D. Duggan, et al., Polymorphisms in WNT6 and WNT10A and colorectal adenoma risk. Nutr. Cancer **63**, 558–564 (2011)
31. B.J. Doolan, A. Onoufriadis, P. Kantaputra, et al., WNT10A, dermatology and dentistry. Br. J. Dermatol. **185**, 1105–1111 (2021)
32. F. Fagotto, A. Aslemarz, EpCAM cellular functions in adhesion and migration, and potential impact on invasion: a critical review. Biochim. Biophys. Acta Rev. Cancer **1874**, 188436 (2020)
33. L. Zhou, Y. Zhu, The EpCAM overexpression is associated with clinicopathological significance and prognosis in hepatocellular carcinoma patients: a systematic review and meta–analysis. Int. J. Surg. **56**, 274–280 (2018)
34. S. Grässel, R.J. Bauer, Collagen XVI in health and disease. Matrix Biol. **32**, 64–73 (2013)
35. Y. Tan, Q. Chen, Y. Xing, et al., High expression of COL5A2, a member of COL5 family, indicates the poor survival and facilitates cell migration in gastric cancer. Biosci. Rep. **41**, BSR20204293 (2021)
36. R. Tao, Z.F. Wang, W. Qiu, et al., Role of S100A3 in human hepatocellular carcinoma and the anticancer effect of sodium cantharidinate. Exp. Ther. Med. **13**, 2812–2818 (2017)
37. F. Mahmood, R. Xu, M.U.N. Awan, et al., Transcriptomics based identification of S100A3 as the key anti-hepatitis B virus factor of 16F16. Biomed. Pharmacother. **163**, 114904 (2023)
38. D. Song, B. Xu, D. Shi, et al., S100A6 promotes proliferation and migration of HepG2 cells via increased ubiquitin-dependent degradation of p53. Open Med. (Wars.) **15**, 317–326 (2020)
39. J.W. Kim, J.H. Kim, S.Y. Yoon, et al., S100A6 protein as a marker for differential diagnosis of cholangiocarcinoma from hepatocellular carcinoma. Hepatol. Res. **23**, 274–286 (2002)
40. Z. Li, M. Tang, B. Ling, et al., Increased expression of S100A6 promotes cell proliferation and migration in human hepatocellular carcinoma. J. Mol. Med. (Berl.) **92**, 291–303 (2014)
41. S. Zheng, H. Shen, Q. Jia, et al., S100A6 promotes proliferation of intrahepatic cholangiocarcinoma cells via the activation of the p38/MAPK pathway. Future Oncol. **13**, 2053–2063 (2017)
42. J. Wang, Z. Song, L. Ren, et al., Pan-cancer analysis supports MAPK12 as a potential prognostic and immunotherapeutic target in multiple tumor types, including in THCA. Oncol. Lett. **24**, 1–13 (2022)
43. S. Messina, L. Frati, C. Leonetti, et al., Dual-specificity phosphatase DUSP6 has tumor-promoting properties in human glioblastomas. Oncogene **30**, 3813–3820 (2011)
44. G. Liu, X. Ouyang, L. Gong, et al., E2F3 promotes liver cancer progression under the regulation of circ-PRKAR1B. Mol. Ther. Nucleic Acids **26**, 104–113 (2021)
45. J. Shen, Z. Jin, H. Lv, et al., PFKP is highly expressed in lung cancer and regulates glucose metabolism. Cell. Oncol. (Dordrecht) **43**, 617–629 (2020)
46. M. Waraya, K. Yamashita, H. Katoh, et al., Cancer specific promoter CpG Islands hypermethylation of HOP homeobox (HOPX) gene and its potential tumor suppressive role in pancreatic carcinogenesis. BMC Cancer **12**, 397 (2012)

47. M. Hassn Mesrati, S.E. Syafruddin, M.A. Mohtar, et al., CD44: a multifunctional mediator of cancer progression. Biomolecules. **11**, 1850 (2021)
48. H.J. Müller-Eberhard, Molecular organization and function of the complement system. Annu. Rev. Biochem. **57**, 321–347 (1988)
49. A.F. Esser, The membrane attack complex of complement. Assembly, structure and cytotoxic activity. Toxicology **87**, 229–247 (1994)
50. S. Willis, V. Polydoropoulou, Y. Sun, et al., Exploratory analysis of single-gene predictive biomarkers in HERA DASL cohort reveals that C8A mRNA expression is prognostic of outcome and predictive of benefit of trastuzumab. JCO Precis. Oncol. **2** (2018)
51. S.M. Totten, R. Adusumilli, M. Kullolli, et al., Multi-lectin affinity chromatography and quantitative proteomic analysis reveal differential glycoform levels between prostate cancer and benign prostatic hyperplasia sera. Sci. Rep. **8**, 6509 (2018)
52. Y. Zhang, X. Chen, Y. Cao, et al., C8B in complement and coagulation cascades signaling pathway is a predictor for survival in HBV-related hepatocellular carcinoma patients. Cancer Manage. Res. **13**, 3503–3515 (2021)
53. R. Pio, L. Corrales, J.D. Lambris, The role of complement in tumor growth. Adv. Exp. Med. Biol. **772**, 229–262 (2014)
54. L.T. Roumenina, M.V. Daugan, F. Petitprez, et al., Context-dependent roles of complement in cancer. Nat. Rev. Cancer **19**, 698–715 (2019)
55. K. Oikonomopoulou, D. Ricklin, P.A. Ward, et al., Interactions between coagulation and complement–their role in inflammation. Semin. Immunopathol. **34**, 151–165 (2012)
56. A.T. Bauer, C. Gorzelanny, C. Gebhardt, et al., Interplay between coagulation and inflammation in cancer: limitations and therapeutic opportunities. Cancer Treat. Rev. **102**, 102322 (2022)
57. F. Violi, D. Ferro, S. Basili, et al., Prognostic value of clotting and fibrinolytic systems in a follow-up of 165 liver cirrhotic patients. CALC Group. Hepatology (Baltimore, Md). **22**, 96–100 (1995)
58. F. Tacke, K. Fiedler, M. von Depka, et al., Clinical and prognostic role of plasma coagulation factor XIII activity for bleeding disorders and 6-year survival in patients with chronic liver disease. Liver Int. **26**, 173–181 (2006)
59. T. Liu, J. Wang, H. Yang, et al., Enhancer coamplification and hijacking promote oncogene expression in liposarcoma. Cancer Res. **83**, 1517–1530 (2023)
60. J. Wei, G. Ran, X. Wang, et al., Gene manipulation in liver ductal organoids by optimized recombinant adeno-associated virus vectors. J. Biol. Chem. **294**, 14096–14104 (2019)
61. C. Zheng, J.N. Fass, Y.P. Shih, et al., Transcriptomic profiles of neoantigen-reactive T cells in human gastrointestinal cancers. Cancer Cell **40**, 410–23.e7 (2022)
62. G. Wu, A.K. Diaz, B.S. Paugh, et al., The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. Nat. Genet. **46**, 444–450 (2014)