Supplemental Information

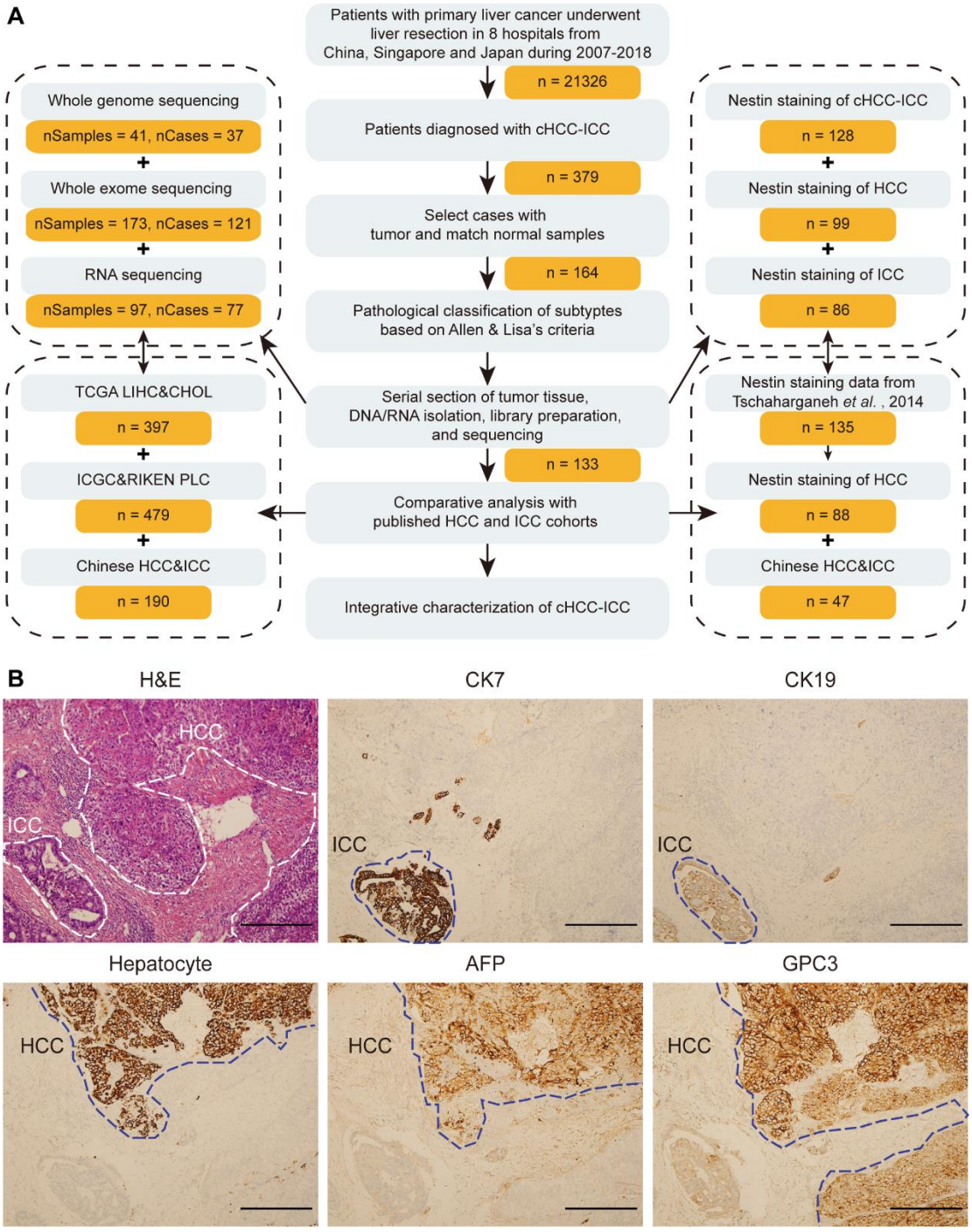

Figure S1. Research strategy and pathological evidence of cHCC-ICC. Related to Figure 1.

(A) Schematic diagram depicting research strategy, including case screening, multi-omics sequencing, IHC staining and comparative analysis with published datasets (data source indicated). TCGA, the Cancer Genome Atlas, LIHC, TCGA's designation for hepatocellular carcinoma, CHOL, TCGA's designation for cholangiocarcinoma, ICGC, International Cancer Genome Consortium.

(B) A set of diagnostic slides for a combined type cHCC-ICC case. Staining markers are indicated. The dashed line divides the HCC and ICC components. Scale bar = 400 μm.

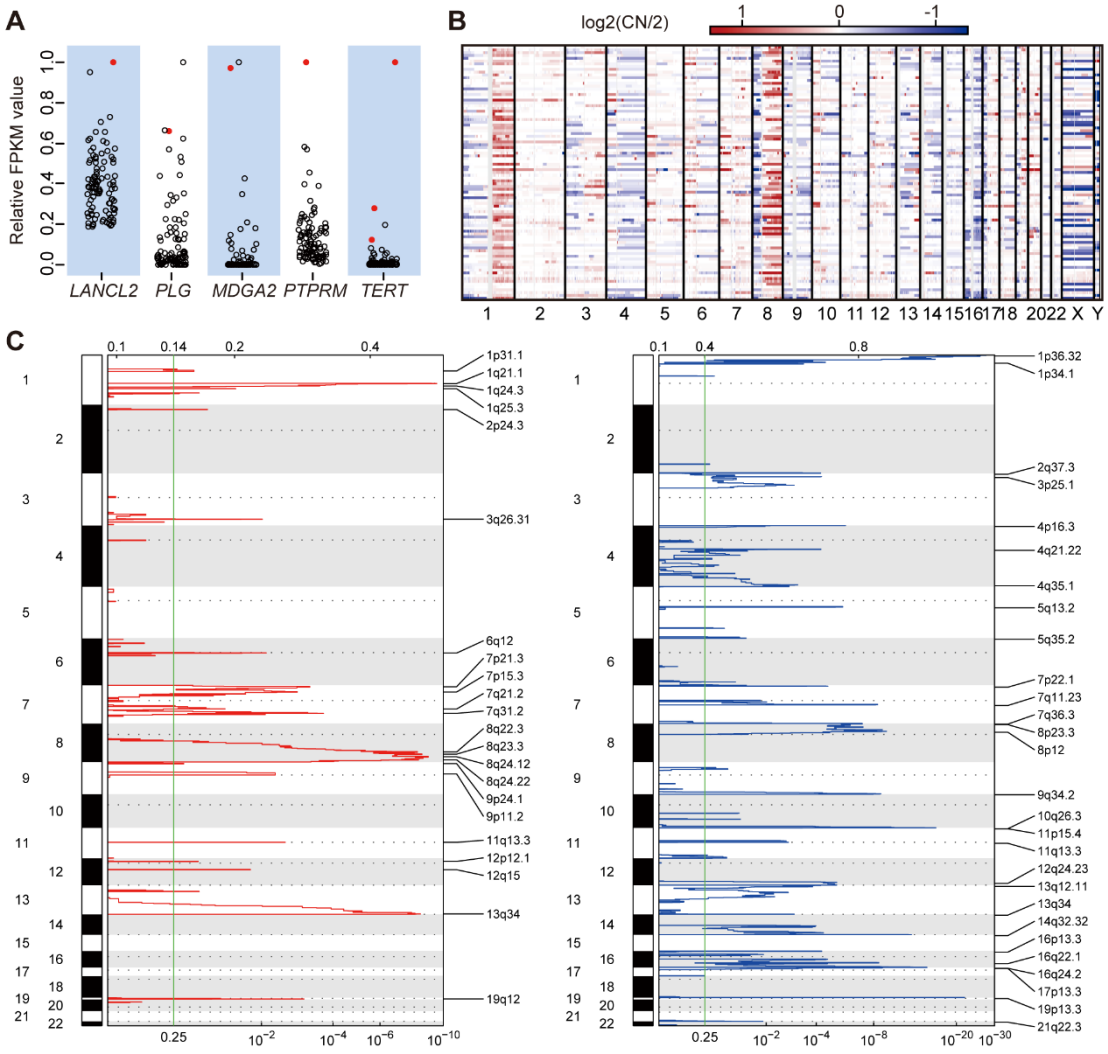

Figure S2. Genomic alterations in cHCC-ICC. Related to Figure 2.

(A) Expression level of genes with HBV integrated into promoter regions. Red denotes the presence of integration, black denotes the absence of integration.

(B) CNA landscape of cHCC-ICC. Colors in the heatmap represent log2 (CN/2) values. Red for copy number gains and amplifications, blue for copy number loss and deletions.

(C) Focally amplified (red) and deleted (blue) genomic regions identified by GISTIC analysis. The genome is oriented vertically from top to bottom, and GISTIC q values at each locus are plotted from left to right on a log scale. The green line represents the significance threshold (q value = 0.25).

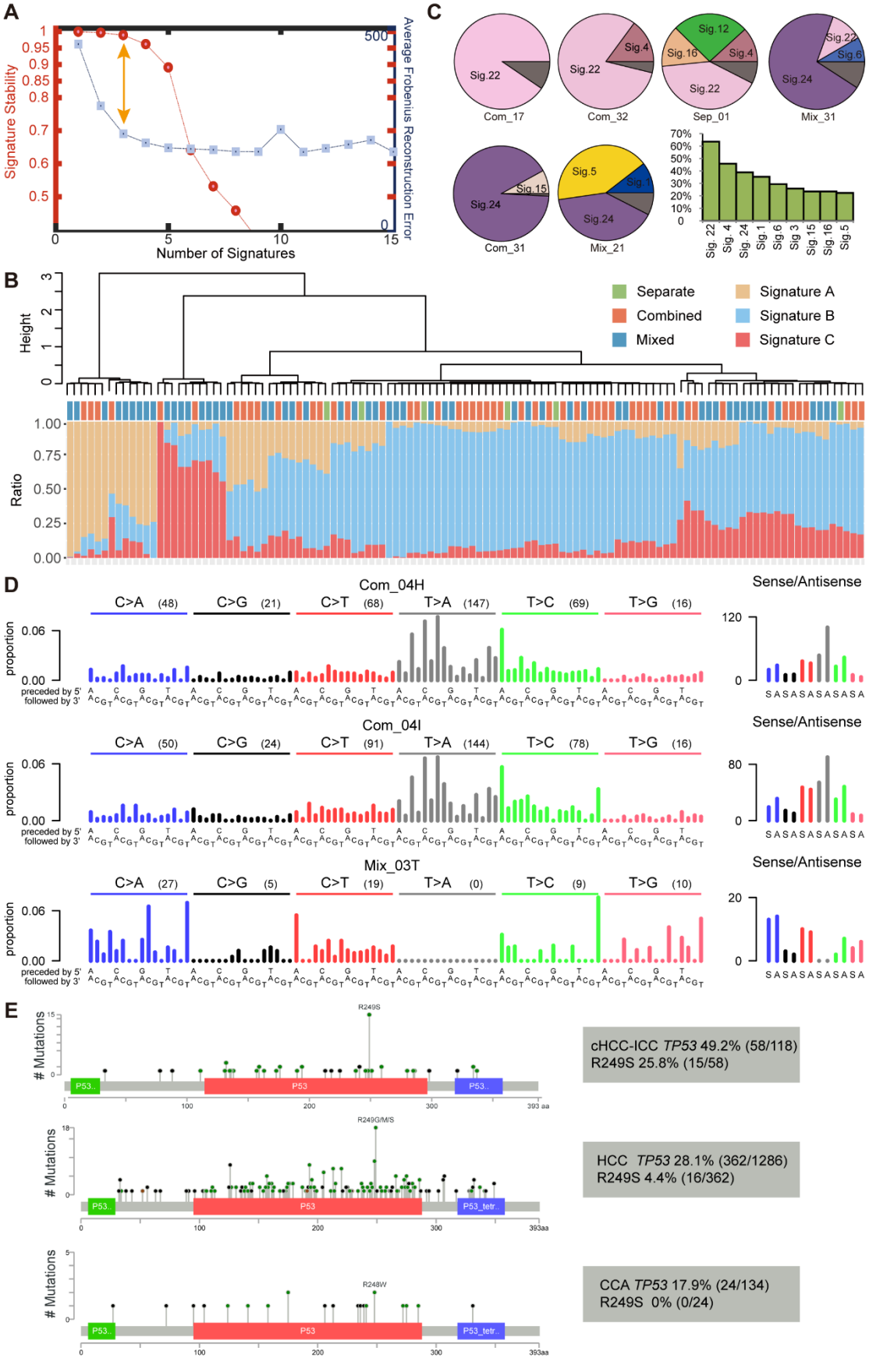
A
Signature Stability
Average Frobenius Reconstruction Error
Number of Signatures
C
Com_17
Com_32
Sep_01
Mix_31
Com_31
Mix_21
B
Height
Ratio
Separate
Combined
Mixed
Signature A
Signature B
Signature C
D
Com_04H
Com_04I
Mix_03T
C>A
C>G
C>T
T>A
T>C
T>G
proportion
preceded by 5'
followed by 3'
Sense/Antisense
SASASASASASA
E
Mutations
R249S
R249G/M/S
R248W
P53
P53_tetr..
cHCC-ICC TP53 49.2% (58/118)
R249S 25.8% (15/58)
HCC TP53 28.1% (362/1286)
R249S 4.4% (16/362)
CCA TP53 17.9% (24/134)
R249S 0% (0/24)

Figure S3. Mutational signatures of cHCC-ICC. Related to Figure 2.

(A) Stability and reconstruction error for the non-negative-matrix factorization analysis. The number of signatures was estimated to be three (indicated by yellow arrow) due to the high signature stability and relatively low reconstruction error.

(B) Clustering of the cHCC-ICC cases based on the three signatures.

(C) COSMIC signatures extracted by deconstructSigs in 6 cHCC-ICC cases. Signatures are indicated on the pie charts. Brown sectors indicate unknown signatures. The frequencies of extracted COSMIC signatures across all cHCC-ICC cases are indicated.

(D) Ninety-six substitution patterns of selected cases from the mSigAct analysis. The Y-axis indicates the frequency of the 96 substitution patterns.

(E) Hotspot mutation analysis for *TP53* across 3 PLCs by Mutation Mapper from cBioPortal. The percentages of *TP53* R249S mutations are indicated. HCC and ICC data are from the cBioPortal website.

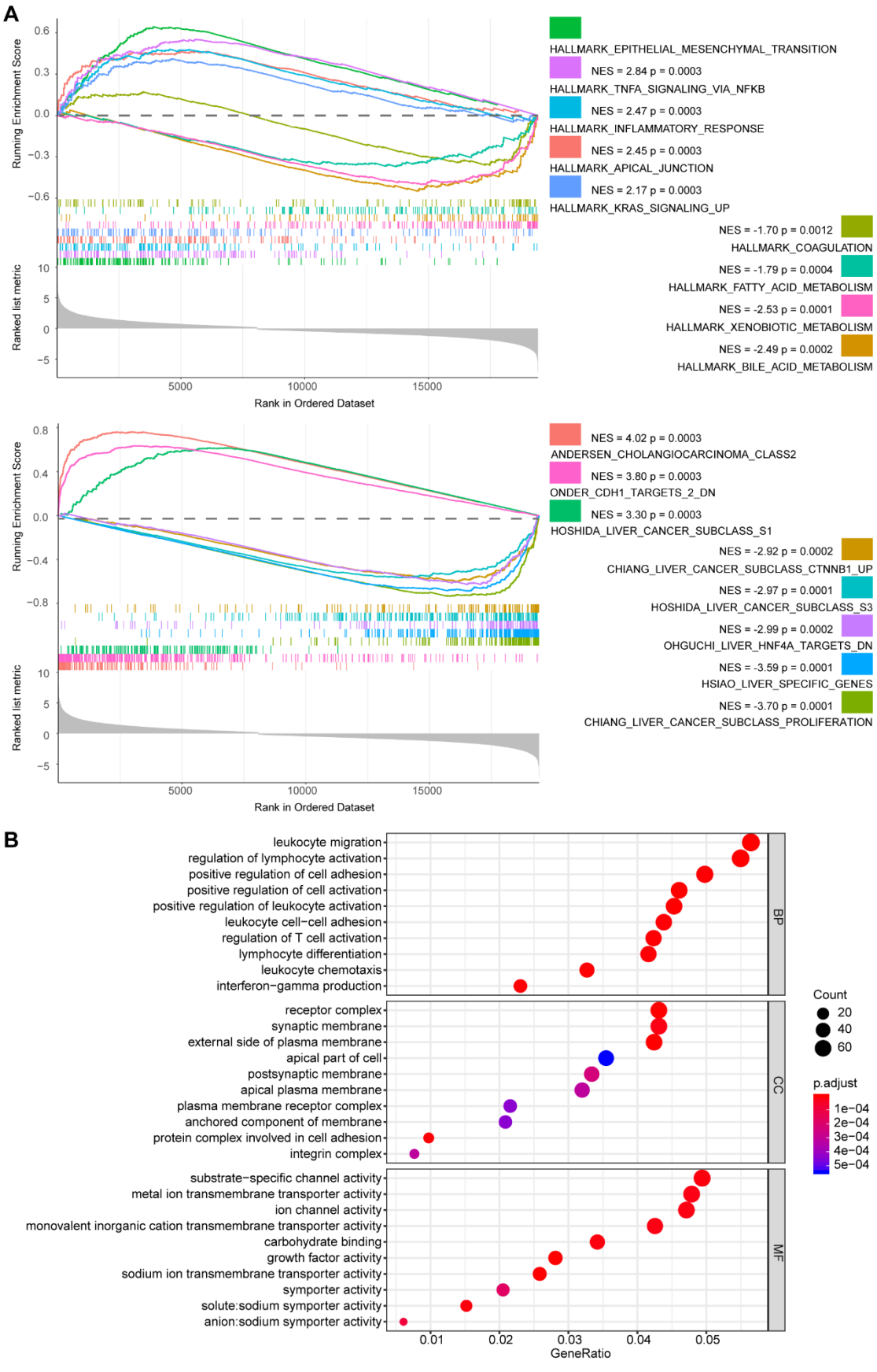
A
Running Enrichment Score
Ranked list metric
Rank in Ordered Dataset
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
NES = 2.84 p = 0.0003
HALLMARK_TNFA_SIGNALING_VIA_NFKB
NES = 2.47 p = 0.0003
HALLMARK_INFLAMMATORY_RESPONSE
NES = 2.45 p = 0.0003
HALLMARK_APICAL_JUNCTION
NES = 2.17 p = 0.0003
HALLMARK_KRAS_SIGNALING_UP
NES = -1.70 p = 0.0012
HALLMARK_COAGULATION
NES = -1.79 p = 0.0004
HALLMARK_FATTY_ACID_METABOLISM
NES = -2.53 p = 0.0001
HALLMARK_XENOBIOTIC_METABOLISM
NES = -2.49 p = 0.0002
HALLMARK_BILE_ACID_METABOLISM
NES = 4.02 p = 0.0003
ANDERSEN_CHOLANGIOCARCINOMA_CLASS2
NES = 3.80 p = 0.0003
ONDER_CDH1_TARGETS_2_DN
NES = 3.30 p = 0.0003
HOSHIDA_LIVER_CANCER_SUBCLASS_S1
NES = -2.92 p = 0.0002
CHIANG_LIVER_CANCER_SUBCLASS_CTNNB1_UP
NES = -2.97 p = 0.0001
HOSHIDA_LIVER_CANCER_SUBCLASS_S3
NES = -2.99 p = 0.0002
OHGUCHI_LIVER_HNF4A_TARGETS_DN
NES = -3.59 p = 0.0001
HSIAO_LIVER_SPECIFIC_GENES
NES = -3.70 p = 0.0001
CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION
B
leukocyte migration
regulation of lymphocyte activation
positive regulation of cell adhesion
positive regulation of cell activation
positive regulation of leukocyte activation
leukocyte cell−cell adhesion
regulation of T cell activation
lymphocyte differentiation
leukocyte chemotaxis
interferon−gamma production
receptor complex
synaptic membrane
external side of plasma membrane
apical part of cell
postsynaptic membrane
apical plasma membrane
plasma membrane receptor complex
anchored component of membrane
protein complex involved in cell adhesion
integrin complex
substrate−specific channel activity
metal ion transmembrane transporter activity
ion channel activity
monovalent inorganic cation transmembrane transporter activity
carbohydrate binding
growth factor activity
sodium ion transmembrane transporter activity
symporter activity
solute:sodium symporter activity
anion:sodium symporter activity
BP
CC
MF
GeneRatio
Count
p.adjust

Figure S4. Comparison of gene expression between clusters P1 and P2. Related to Figure 4.

(A) GSEA analysis results comparing clusters P1 and P2 with H_hallmark gene sets (upper graph) and C2_curated gene sets (lower graph) from MSigDB. Normalized enrichment score (NES) and p values of each gene set are indicated.

(B) Gene ontology (GO) enrichment analysis of differentially expressed genes (fold change >1) between combined and mixed type cHCC-ICC.

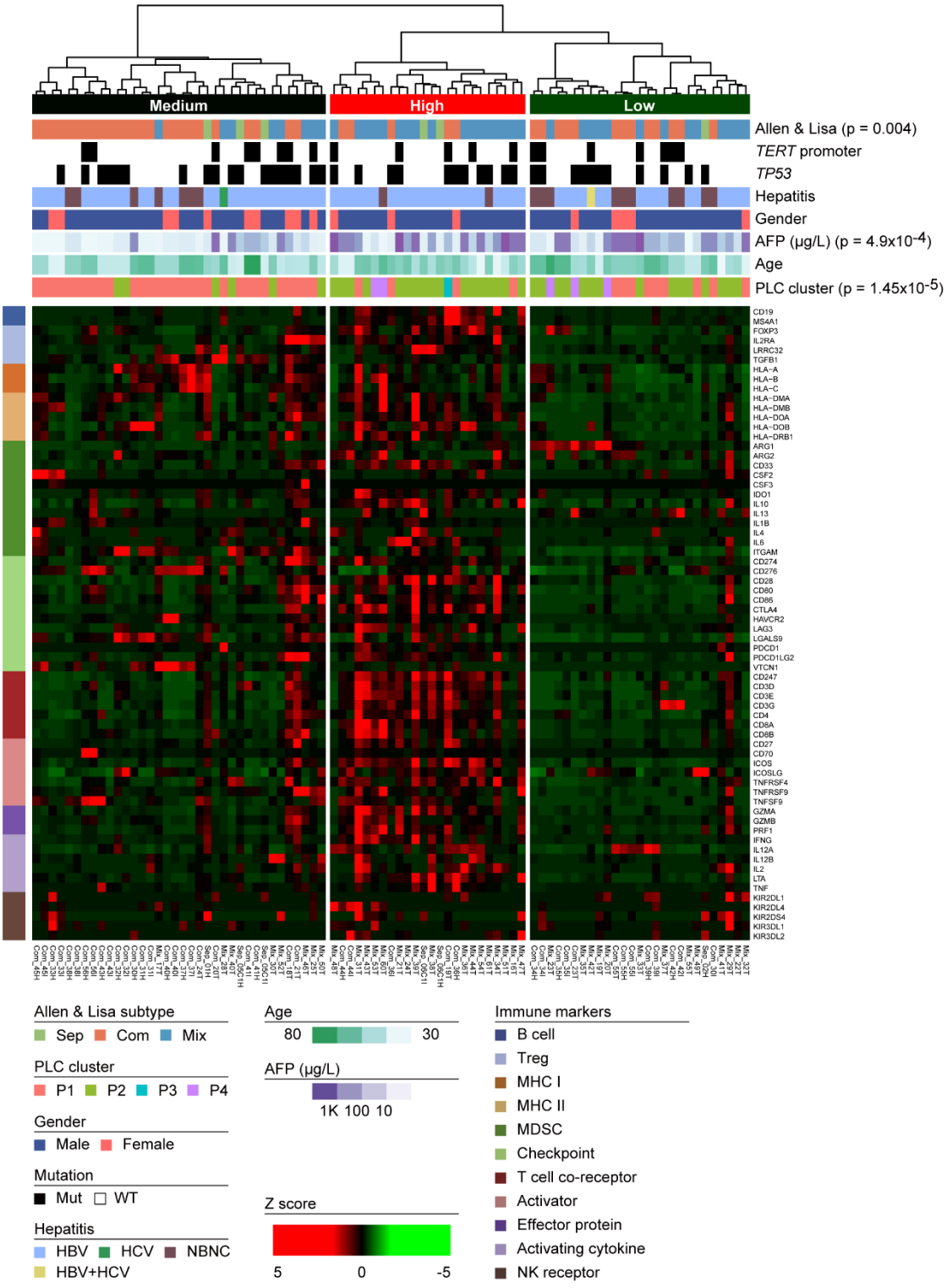

Figure S5. Immune phenotype of cHCC-ICC. Related to Figure 4.

Unsupervised clustering analysis of the gene expression levels of selected immune-related markers identified 3 immune clusters. Clinical and molecular features are annotated above the heatmap with details shown in the legend to the right of the heatmap. P values indicate significant non-random distributions for each attribute.

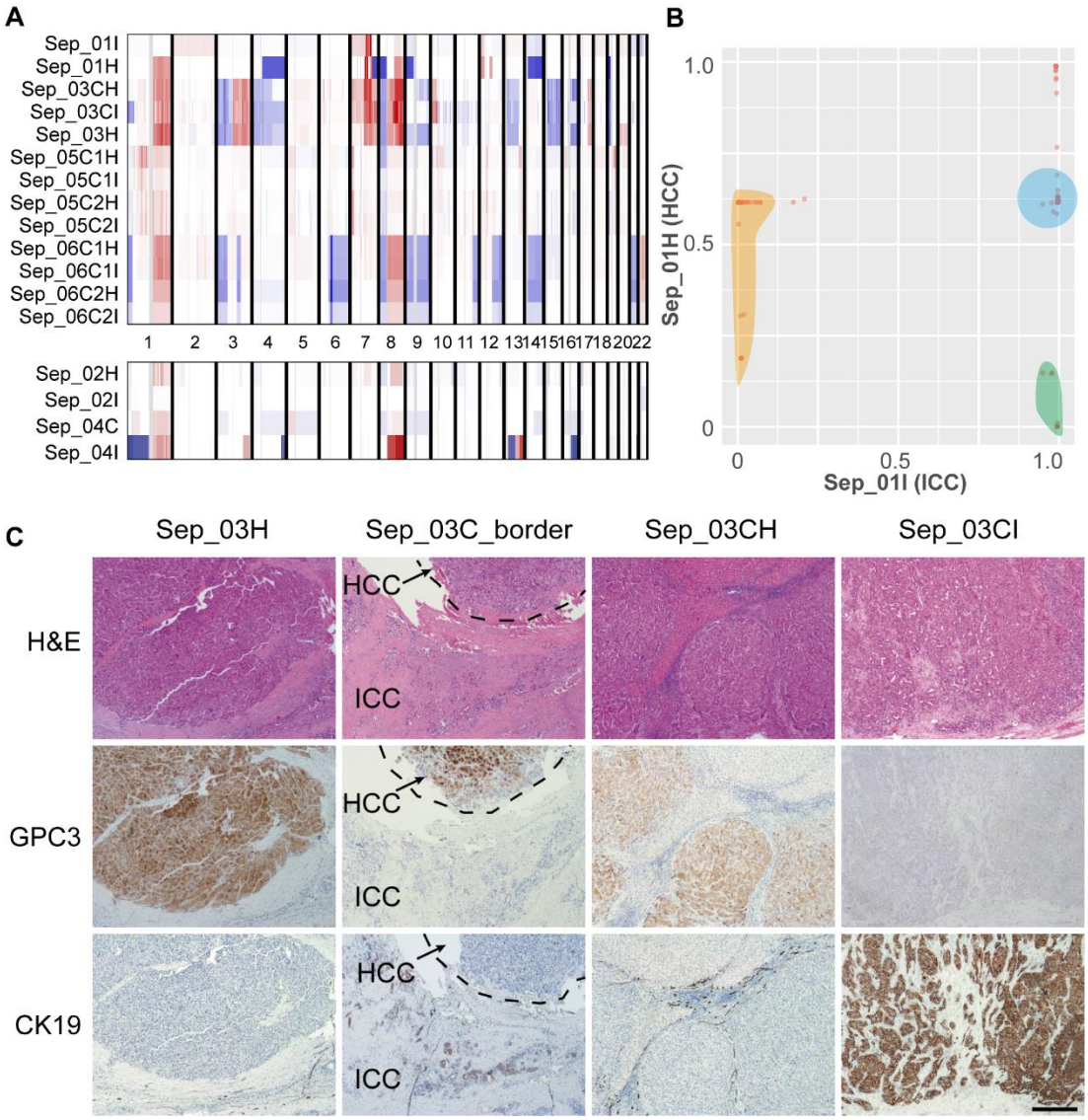

Figure S6. Genomic analysis of separate type cHCC-ICC. Related to Figure 5.

(A) CNA landscape of separate type cHCC-ICC. Colors in the heatmap represent log2 (CN/2) values. Red for copy number gains and amplifications, blue for copy number loss and deletions. The upper panel shows samples from mono-clonal cases. The lower panel shows samples from multi-clonal cases.

(B) Pairwise cancer cell fraction plot of mutations in the two lesions from Sep_01. Each point denotes a mutation: its X coordinate equals the CCF value of this mutation in Sep_01I, its Y coordinate equals the CCF value of this mutation in Sep_01H. The orange and green shades mark mutations private to Sep_01H and Sep_01I, respectively. The blue circular shade marks the subclone that spreads out from the primary (Sep_01H) and later becomes the founder clone of the metastases (Sep_01I).

(C) A set of pathological slides for Sep_03. In the slides of Sep_03C_border, HCC and ICC components are indicated. Scale bar = 400 μm.

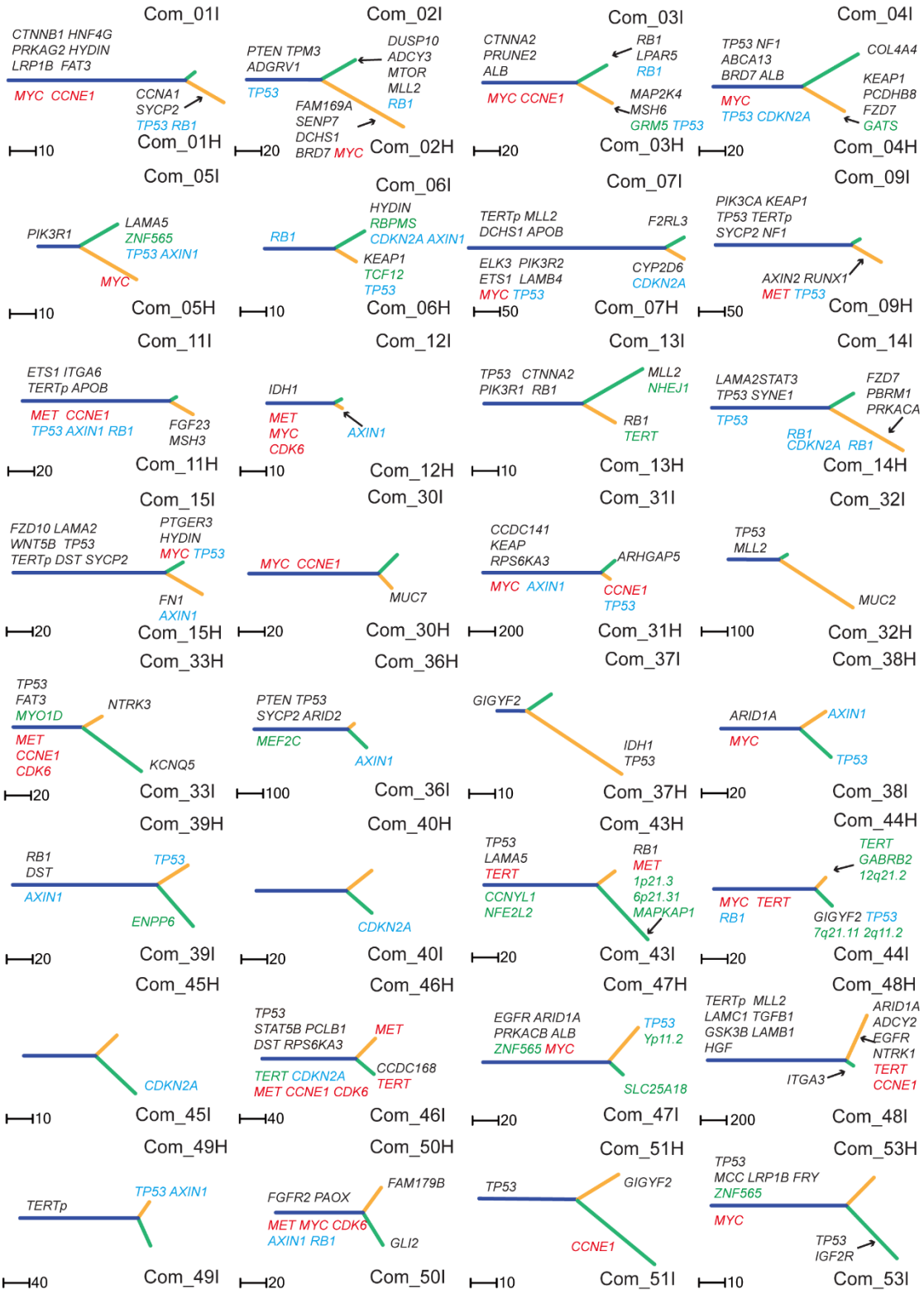

Figure S7. Phylogenetic trees of combined type cHCC-ICC. Related to Figure 6.

In the phylogenetic trees, the length of each line is proportional to the number of mutations. Scale bar of each tree is indicated. Blue lines denote mutations shared by an H-I pair. Orange and green lines denote mutations identified only in HCC or ICC components, respectively. Arrows indicate the acquisition of

potential driver events during tumor evolution. For the genes, black, green, red and blue denote somatic mutation, HBV integration, copy number gain and copy number loss, respectively.