## Title
Genomic and Epigenomic Heterogeneity of Hepatocellular Carcinoma



## Journal
Cancer Research, 77(9)

## ISSN
0008-5472

## Authors
Lin, De-Chen
Mayakonda, Anand
Dinh, Huy Q
et al.

## Publication Date
2017-05-01

## DOI
10.1158/0008-5472.can-16-2822

Peer reviewed

Published in final edited form as:

*Cancer Res*. 2017 May 01; 77(9): 2255–2265. doi:10.1158/0008-5472.CAN-16-2822.

# Genomic and epigenomic heterogeneity of hepatocellular carcinoma

**De-Chen Lin**[1,2,†,*], **Anand Mayakonda**[3,*], **Huy Q. Dinh**[4,*], **Pinbo Huang**[5,*], **Lehang Lin**[1,*], **Xiaoping Liu**[1], **Ling-wen Ding**[3], **Jie Wang**[5], **Benjamin P. Berman**[4,†], **Er-Wei Song**[1,†], **Dong Yin**[1,†], and **H. Phillip Koeffler**[2,3,6]

[1]Guangdong Province Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Research Center of Medicine, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China

[2]Division of Hematology/Oncology, Cedars-Sinai Medical Center, UCLA School of Medicine, Los Angeles, USA

[3]Cancer Science Institute of Singapore, National University of Singapore, Singapore

[4]Center for Bioinformatics and Functional Genomics, Biomedical Sciences, Cedars-Sinai Medical Center, UCLA School of Medicine, Los Angeles, USA

[5]Department of Hepatobiliary Surgery, Sun-Yat-Sen Memorial Hospital, Sun-Yat-Sen University, Guangzhou, China

[6]National University Cancer Institute, National University Hospital Singapore, Singapore

## Abstract

Understanding the intratumoral heterogeneity of hepatocellular carcinoma (HCC) is instructive for developing personalized therapy and identifying molecular biomarkers. Here we applied whole-exome sequencing to 69 samples from 11 patients to resolve the genetic architecture of subclonal diversification. Spatial genomic diversity was found in all 11 HCC cases, with 29% of driver mutations being heterogeneous, including TERT, ARID1A, NOTCH2, and STAG2. Similar with other cancer types, TP53 mutations were always shared between all tumor regions i.e. located on the “trunk” of the evolutionary tree. In addition, we found that variants within several drug targets such as KIT, SYK and PIK3CA were mutated in a fully clonal manner, indicating their therapeutic potentials for HCC. Temporal dissection of mutational signatures suggested that mutagenic processes associated with exposure to aristolochic acid and aflatoxin might play a more important role in early, as opposed to late, stages of HCC development. Moreover, we observed extensive

[†]Correspondence should be addressed to B.P.B (Benjamin.Berman@csmc.edu), D.Y. (yin_dong@yahoo.com), E-W.S (songerwei02@aliyun.com) or D-C.L. (dchlin11@gmail.com).
[*]These authors contributed equally to this work.

The authors have no potential conflicts of interest to disclose.

Web Resources
GATK best practices, https://www.broadinstitute.org/gatk/guide/bp_step.php?p=1.
PHYLIP, http://evolution.genetics.washington.edu/phylip.html.
BWA-MEM, http://arxiv.org/abs/1303.3997v2.
Bam-readcount, (https://github.com/genome/bam-readcount)
TCGA data portal, https://tcga-data.nci.nih.gov/tcga/

intratumoral epigenetic heterogeneity in HCC based on multiple independent analytical methods and showed that intratumoral methylation heterogeneity might play important roles in the biology of HCC cells. Our results also demonstrated prominent heterogeneity of intratumoral methylation even in a stable HCC genome. Together, these findings highlight widespread intratumoral heterogeneity at both the genomic and epigenomic levels in HCC and provide an important molecular foundation for better understanding the pathogenesis of this malignancy.

## INTRODUCTION

As the second leading cause of cancer mortality, the incidence of liver cancer is increasing in almost all countries, representing a major health problem, particularly in Asia (1). Hepatocellular carcinomas (HCC) contribute to over 90% of all liver cancers, with well-studied risk factors including infection of hepatitis B virus (HBV) and hepatitis C virus (HCV), alcohol consumption, aflatoxin exposure and metabolic abnormalities, among others (2, 3). Previous HCC genomic analysis using single sampling per tumor have identified a number of driver mutations in this malignancy (4–11); however, much remains unclear regarding the extent of genomic diversity and clonal evolution of primary HCC. In addition, despite the few exogenous mutational processes (e.g., aflatoxin ingestion) associated with particular mutational signatures (6, 9, 12), the temporal dynamics of these processes and their contribution to subclonal diversification during HCC evolution remain unclear. Moreover, the intratumoral epigenetic heterogeneity of HCC is unexplored and its biological significance remains unknown. Recently, our group addressed these important questions in esophageal tumors and observed profound intratumoral heterogeneity at both genetic and epigenetic levels (13).

Precise understanding of both the genomic and epigenomic architecture of primary HCC tumors is crucial for developing personalizing patient treatment and molecular-based biomarkers (14). In this study, we address these critical issues through integrative molecular approaches, including Multiregion Whole-Exome Sequencing (M-WES), phylogenetic tree construction, as well as multiregion methylation profiling. Furthermore, we analyzed HCC genomic data from TCGA to dissect temporally the mutational processes and clonality of driver mutations.

## MATERIAL AND METHODS

### Patients and sample processing

All of the samples including tumor tissues (n = 52), morphologically nonmalignant liver tissues (n = 6) and matched peripheral blood (Germline, n = 11) from 11 HCC patients were collected after diagnosis from Sun Yat-Sen Memorial Hospital during the years of 2015–2016 (Supplementary Table 1). Samples from the case HCC8031 were collected after transarterial chemoembolization; all other samples were obtained before treatment. All patients presented evidence of HBV infection except HCC8392, and no patient reported a history of alcohol consumption. The median patient age was 48 years (range 32–71). Tumor stages ranged from II to IIIC. Detailed clinical and pathological characteristics of these HCC cases are provided in Supplementary Table 2. This study has been approved by the Ethics

Committee at Sun Yat-Sen Memorial Hospital. To study intratumoral heterogeneity, five spatially-isolated tumor specimens were obtained per patient (except for HCC8392, HCC8716 and HCC5647, where only 4 different tumor regions from each individual were available). Each specimen was at least 3–4 cm away from the others. In 6 out of 11 patients, we also collected morphologically nonmalignant liver tissues which were at least 3–4 cm away from the nearest tumor region. We carefully reviewed available hematoxylin and eosin (H&E) slides of each tumor region to verify that the tumor cell content of the selected regions were comparable and were at least greater than 70%. All of the tissue samples were snap frozen in liquid nitrogen and preserved at −80 °C before analysis.

## Whole exome sequencing (WES)

DNA was extracted using Qiagen DNeasy Blood & Tissue Kit. Whole exome capture was performed on 2–3 μg of genomic DNA with the Agilent SureSelect Human All Exon v4 (51 Mb) kit according to the manufacturer's instructions, and captured nucleotides were subjected to massively parallel sequencing using Illumina HiSeq4000 platform at Beijing Genome Institute (BGI), China.

## Construction of phylogenetic tree

To increase the sensitivity of detecting mutations that had been called from one tumor section in other tumor sections from the same case at a low VAF, we adopted the method published by Stachler et al (15). Specifically, read counts for all somatic variants across all tumor sections were extracted using Bam-Readcount (URL). A variant with VAF less than 0.02 were considered as absent. A binary table was then generated across all tumor regions for each HCC case. The binary tables were used to construct phylogenetic trees, using the method of Discrete Character Parsimony, implemented in PHYLogeny Inference Package (URL). Germline sample was set as an outgroup root. Inferred trees were manually recolored, with branch/trunk lengths proportional to the number of mutations.

## DNA methylation analysis

DNA methylation profiles of 22 tumor regions and 4 morphologically nonmalignant liver tissues (3 out of 4 were matched nonmalignant livers) from 5 HCC individuals were performed using the Illumina Infinium HumanMethylation450K (HM450) platform (Illumina, San Diego, CA) at the University of Southern California Norris Comprehensive Cancer Center Genomics Core. The identical DNA extracted from these 22 tumor regions were also subjected to M-WES as described earlier. For data analysis, pre-processing steps including background correction and normalization to get corrected beta values were performed using the TCGA pipeline (methylumi R package as described in Triche et al. (16)). Probes with a detection p-value over 0.01 were removed, as were probes overlapping with dbSNP SNPs, and probes on the X chromosome. We defined low methylated probes in nonmalignant tissues as those for which beta values of all nonmalignant samples were less than or equal to 0.3; highly methylated probes in nonmalignant tissues were defined as those for which all nonmalignant samples were larger than 0.6. For intratumoral analysis, computational methods were performed as we have described recently (13). Specifically, we defined the variable probes as those where the difference in beta values of at least 1 pair of

tumor regions was at least 0.3, and invariable probes as those where the differences in beta values of all the pairs were less than 0.1.

### Accession codes

Digital sequencing and HM450 Bead array files have been deposited into Sequence Read Archive (SRP076602) and Gene Expression Omnibus (GSE83691), respectively.

Other Materials and Methods are described in the Supplementary Information.

## RESULTS

### HCC Genomic Architecture and Evolution

To investigate HCC genomic evolution, we first performed M-WES on a total of 52 tumor regions and 11 matched germline DNA (peripheral blood), collected from 11 HCC cases (Supplementary Tables 1–2). In 6 out of these 11 patients, we also sequenced whole exome of the DNA from morphologically nonmalignant liver tissues which were 3–4 cm away from cancerous regions, to investigate whether cancer "field effect" exists in HCC. Field effect refers to the presence of clonal expansions in nonmalignant tissues that provide a background for malignancy development, which was observed by us and others in the tumors from colon (17), prostate (18) and head and neck (19).

M-WES achieved an average target depth of 158X (range, 72X–232X; Supplementary Table 1), and identified an average of 129 somatic variants per tumor region (Supplementary Tables 3–5), a mutational rate comparable to recent sequencing reports on HCC (4–6, 9, 20). Notably, the case-wise mutational rate (159 mutations/case) was significantly higher ($P < 0.001$, Supplementary Table 3), underscoring the stronger power and improved resolution of our multi-biopsy approach for genomic analysis. To evaluate intratumoral heterogeneity and to illustrate HCC genome evolution, we constructed phylogenetic trees with the trunk representing mutations present in all tumor regions, branches representing mutations present in some but not all regions. Spatial intratumoral heterogeneity was evident in all 11 HCC cases, showing a median of 38.9% of variants being heterogeneous (range, 5.5%–91.7%; Fig. 1, Supplementary Fig. 1, Supplementary Table 3). The phylogenetic tree structure varied considerably between HCC cases (Fig. 1). For example, HCC8392 had a longer branch than its trunk, whereas HCC8716 displayed a very homogenous mutational pattern. Importantly, we validated the intratumoral heterogeneity at the protein level by immunohistochemistry (IHC) staining of the tumor regions which had available matched formalin-fixed paraffin-embedded (FFPE) samples. In HCC5647, *FAT4* mutation was only observed in T4 region, and correspondingly FAT4 protein expressed at markedly lower level compared with other tumor regions (Fig. 2A). In contrast, in HCC6690 which harbored a truncal *TP53* mutation, we observed strong staining of nuclear p53 protein in all 5 tumor regions (Fig. 2A).

### Spatial Heterogeneity and Clonal Status of Driver Mutations

Deciphering intratumoral heterogeneity of driver mutations is instructive for both better understanding of the cancer genome evolution and improving precision medicine strategies.

We, therefore, identified potential driver mutations in this HCC cohort by comparing against COSMIC gene census (21), Pan-cancer analysis (22), as well as recent large-scale HCC sequencing results (4–7, 20, 23) on the basis of a set of criteria evaluating the functional impact of the variant (See Methods). These candidate driver events were then mapped on the phylogenetic tree structures. Similarly as observed in breast (24) and esophageal adenocarcinoma (25), 29% (10/34) of putative driver mutations were present in the branches, including *TERT, ARID1A, NOTCH2, STAG2,* etc. Consistent with previous reports in other cancer types, *TP53* mutations were always truncal. These results suggest that single-biopsy genome analysis may not be sufficient to evaluate comprehensively HCC driver events.

We next calculated cancer cell fraction (CCF) of all mutations to assess their clonal status through integrative analysis of tumor cellularity, variant copy number and variant allele frequency (VAF) as described previously (26, 27). Notably, several driver mutations were fully clonal within some tumor regions, yet could not be detected in other areas. For example, *TERT* promoter hotspot mutation, which has frequently been identified in HCC, was absent in T1 region in case HCC8031 but appeared to be clonally dominant in all of the rest tumor regions (Fig. 2B). Likewise, *ARID1A* mutation was assessed to be present in 100% tumor cells in T3 and T4 regions, yet was undetectable in T1 and T2 regions in HCC8257 (Fig. 2B). Overall, although a number of mutations targeting driver genes were estimated to be fully clonal in all tumor regions sequenced (e.g., *TERT* and *CTNNB1* in HCC6046, *TP53* and *SYK* in HCC6952), it is evident that many of the driver mutations can be subclonal, indicating that they occur as relatively late events during the development of HCC.

To corroborate further these findings, we calculated the relative timing and clonal status of driver mutations using TCGA HCC datasets (23) (See Methods). We observed that 45.5% (87 out of 191) TCGA HCC patients had more than 10% subclonal mutations (Supplementary Fig. 2). Importantly, in line with our earlier results, we found that in TCGA HCC samples a number of driver genes (defined by MutSig (28) in TCGA cohort) harbored a significantly higher proportion of subclonal mutations compared with that of total driver variants, such as *CTNNB1, PTEN NOTCH2* and *PTPRB* (Fig. 3B). Overall, compared with passenger genes, driver genes still had significantly more clonal mutations in the TCGA cohort, in accordance with the observations in other types of cancers (23) (Fig. 3A).

However, we did not find compelling evidence of a “field effect” in our cohort of HCCs. Only an average of 2 nonsilent mutations per sample (range 0–5) occurred in nonmalignant liver tissues from these HCC cases; and none of them seemed to be functionally relevant to cancer biology (Fig. 1). CCF analysis of these variants also suggested that many of them exhibited lack of clonal expansions within nonmalignant liver tissue (Supplementary Table 5).

We next examined HCC intratumoral heterogeneity at the copy number level, and compared our somatic copy number alteration (SCNA) calls with GISTIC (Genomic Identification of Significant Targets in Cancer) results from the TCGA HCC project. Consistent with TCGA and other previous HCC SCNA studies (4, 9), a number of important copy number changes were noted in the present cohort, including copy number gain of 1q21, 5p15 (encompassing

*TERT*), 8q24 (harboring *MYC*) and 11q13 (harboring *CCND1*) as well as loss of 1p36, 10q26 and 14q32. Overall, SCNA profiles were more similar within the same individual than between different ones. Nevertheless, extensive copy number intratumoral heterogeneity was observed, such that the majority of recurrent SCNAs were spatially heterogeneous in one or more cases (Supplementary Fig. 3). For instance, gain of 8q24 (harboring *MYC*) was ubiquitous in some cases but occurred as a heterogeneous aberration in other cases. Gain of 5p15 (encompassing *TERT*) was detected to be often spatially heterogeneous. Interestingly, our deep sequencing results also found a branch hotspot mutation in *TERT* promoter region (Fig. 1). These data suggest that the hyper-activation of telomerase resulting from both genetic mutations and copy number gains are likely late events during HCC pathogenesis. We also noticed that two significant SCNAs, copy number gain of 11q13 (harboring *CCND1*) and loss of 11q14 were consistently ubiquitous across intratumoral regions, underscoring the importance of these SCNAs as initiating forces during HCC transformation. Together, these results demonstrate that similar to DNA single nucleotide mutations, SCNAs also exhibit significant intratumoral heterogeneity in HCC, congruent with the observations in other types of cancers (25, 29, 30).

## Temporal Dissecting Mutational Spectra and Signatures

We next analyzed the mutational spectra of both truncal and branch mutations to determine the dynamics of the mutagenic processes operative in HCC. Although we only observed moderate changes in the six mutational classes between truncal and late branch mutations (Fig. 4A, Supplementary Fig. 4), conspicuous differences in the 96 trinucleotide mutational signatures were noted (Fig. 4B). We next calculated the contributions of individual mutational signatures to each tumors, and identified several dominant signatures in these tumors, including Signature 1 (associated with age), Signatures 22 (associated with aristolochic acid exposure) and Signature 24 (associated with aflatoxin exposure), and Signatures 6 and 15 (associated with DNA mismatch repair), which were in agreement with previous results from other groups (6, 9, 12). The robustness and stability of the analysis was confirmed by extensive permutation tests (Supplementary Fig. 5). Interestingly, a number of tumors exhibited prominent decreases of the contribution of Signatures 22, 24 and 25 (unknown aetiology) in the branch compared with truncal mutations, albeit without obtaining statistical significance due to the relatively small number of tumors analyzed. These results indicate that exogenous factors, such as exposure to aristolochic acid and aflatoxin, contribute significantly to the mutagenic processes in the early stage of HCC development, whereas other endogenous factors might play more important roles in shaping the mutational spectrum after the most recent common tumor ancestor is established.

## Intratumoral Heterogeneity of DNA Methylation in HCC

As in other common cancers, epigenetic alterations contribute significantly to HCC pathogenesis; and DNA methylation has been found to alter the expression of critical HCC cancer genes (31–33). To investigate intratumoral heterogeneity at the epigenetic level, we profiled the global DNA methylation of a total of 22 tumor tissues and 4 nonmalignant liver tissues from 5 HCC patients (all of which had matched WES results) using Illumina Human Methylation 450k (HM450) Bead array. First, we identified variably-methylated CpG sites (See Methods) across different tumor regions within each individual and found intratumoral

heterogeneity in all individuals (Supplementary Fig. 6), a phenomenon that was recently observed in glioma, prostate and esophageal cancers (34–37). All of the tumor regions had numerous differences from the methylation pattern of matched nonmalignant liver tissues (Supplementary Fig. 6; in two cases where matched nonmalignant samples were unavailable, tumor samples were compared against nonmalignant tissues from other individuals). Importantly, while different tumor regions showed variable methylation patterns, nonmalignant liver tissues were very homogenous, as demonstrated by the tight clustering; this high degree of inter-individual homogeneity of nonmalignant liver tissues is consistent with previous epigenetic studies in HCC (31–33), and was further confirmed by our analysis of TCGA nonmalignant liver samples (Supplementary Fig. 7).

To explore potential biological relevance of the intratumoral heterogeneity of DNA methylation in HCC, we examined how the differentially-methylated probes were distributed across different genomic contexts. To this end, we defined both hyper- and hypo-methylated CpG sites in these five individuals compared with matched nonmalignant liver tissues, and further separated them into ones with variable methylation between tumor regions (Fig. 5A) and ones with invariable methylation (Supplementary Fig. 8). This classification revealed a wide range in the degree of intratumoral methylation heterogeneity as measured by the number of variable probes in these 5 HCCs (Fig. 5A and Supplementary Table 6). For example, only 1.8% of total hypermethylated probes exhibited heterogeneity in HCC6952; in sharp contrast, the majority (63.4%) of hypermethylated CpG sites in HCC8257 were variable among different tumor regions (Supplementary Table 6). In order to rule out and mitigate the confounding effects from non-tumor DNA contamination, we performed approaches described previously by us and others to infer and correct for non-cancer cells (37–39), using genome-wide Infinium methylation profiles of various flow-sorted immune cell types (40) to adjust beta values of each probe accordingly. Importantly, the extent of intratumoral methylation heterogeneity in these HCCs was almost identical after this correction (Supplementary Table 6), suggesting that the intratumoral methylation heterogeneity was largely driven by the cancer cells themselves. Moreover, we validated the relative degree of methylation heterogeneity using a recently-described metric based on Shannon entropy (41), after carefully controlling for the number of regions per case. This independent analysis agreed strongly with the degree of methylation heterogeneity inferred from the number of variably methylated probes (Supplementary Fig. 9), and confirmed intratumoral methylation heterogeneity is a highly robust feature of these samples.

We next compared variable and invariable loci to different genomic contexts, including relevant regulatory features such as promoters, CpG Islands (CGIs), and enhancers (Fig. 5B). Invariably-methylated CpG sites showed known patterns of cancer-specific methylation changes (42), including hypermethylated probes that were strongly enriched in CGI promoter regions, and depleted in both partially methylated domains (PMDs) and enhancer regions after removing CGIs. Conversely, invariably-hypomethylated probes were markedly depleted in CGI promoters while enriched in PMDs as well as enhancer regions (Fig. 5B, Supplementary Table 7). Strikingly, variably-methylated CpG sites strongly resembled the distribution pattern of their invariable counterpart (Fig. 5B, Supplementary Table 7) in all 5 cases. Given the rich body of evidence showing the direct contribution of cancer-specific methylation to tumorigenesis and progression (e.g., suppression of tumor suppressors

through promoter hypermethylation, activation of oncogenes through long-range hypomethylation (43, 44)), our results suggest that heterogeneity of intratumoral methylation is likely to play important roles in the biology of HCC cells. In support of this postulate, gene ontology (GO) functional analysis of the genes with variably-hypermethylated promoters showed that they were significantly enriched in a number of cancer-related processes, including cell proliferation, differentiation, death, migration, adhesion and transcriptional regulation (Fig. 5C). As expected, invariably-hypermethylated promoters also exhibited stronger enrichment in most processes (Fig. 5C). Close examination of the methylation targeted genes showed that a number of critical cancer genes (e.g., *APC, IRF6, RUNX3, CDKN2A*) could be either invariably or variably hypermethylated in their promoter regions, indicating that their expression might be differentially suppressed in different tumor regions (Supplementary Tables 8–9).

## Prognostic value of genomic and epigenomic intratumoral heterogeneity

Recent studies in other cancer types have found a high degree of genomic intratumoral heterogeneity was associated with worse clinical outcome (45). While clinical outcome information was unavailable for the cases studied here, we attempted to gain insights into this question using the TCGA HCC dataset. TCGA cohort contains only single-region samples, and is thus not ideal for studying intratumoral heterogeneity. Nevertheless, a recent mathematical algorithm has been introduced to estimate genomic intratumoral diversity from bulk-tumor WES data, using the Mutant-Allele Tumor Heterogeneity (MATH) approach (46, 47).

We first determined the MATH score of each individual (see Methods), then performed Log-rank analysis to test the association between MATH scores and patients' survival probability. HCC individuals with higher MATH scores (> median) had worse outcome compared with those with lower scores (< median), albeit the result did not achieve statistical significance (Supplementary Figure 10B, $P = 0.19$). We reasoned that more extreme difference in MATH scores (corresponding to larger variation in the degree of genomic heterogeneity) might reflect more difference in cancer malignancy, and therefore we compared patients with the highest 10% MATH scores to those with the lowest 10%. Here, cases with low MATH values exhibited significantly better survival rate (Supplementary Figure 10A, $P = 0.05$).

Inspired by this result, we defined a Methylation Intratumoral Heterogeneity (mITH) score similar in form as the MATH metric but using beta values from the HM450 array. Briefly, the mITH measures the median absolute deviation of probe beta values (see Methods for details). We restricted our analysis to hypermethylated probes, since the process of hypermethylation (which occurs primarily at CpG island promoters) is a more constrained and well-understood process in cancer than hypomethylation. When TCGA cases were divided into two equal groups based on mITH score, we observed the same trend as for the MATH comparison -- high-mITH cases had worse survival rates than low-mITH ones; like the MATH comparison, this trend did not reach statistical significance (Supplementary Fig. 10D, $P = 0.17$). However, when we did the analysis of the mITH extremes (10% highest mITH v.s. 10% lowest mITH), the result was highly significant (Supplementary Fig. 10C, $P = 0.001$). While we believe that multiregion sampling will ultimately be required to validate

fully these inference-based results, they do suggest that intratumoral heterogeneity may have prognostic value in HCC, and reinforce the importance of fully capturing intratumoral heterogeneity using a multiregion approach.

## DISCUSSION

HCC is a worldwide leading malignancy with surgical resection and liver transplantation being the current mainstay treatment options. Morphologic and immunophenotypic intratumoral heterogeneity of HCC have been observed previously (48–50). By focusing recently on *TP53* and *CTNNB1* mutations in HCC, Friemel *et al.* (50) also reported that those variants could be either clonal or subclonal, which is in line with our findings. In contrast, hepatocellular adenoma which is another form of rare, benign tumor of the liver, was found to exhibit homogenous histology and morphology (51, 52). Accordingly, genotypic and phenotypic classifications of hepatocellular adenoma has been established, and begin to have implications for clinical management (53). However, unresolved intratumoral heterogeneity impedes the development of both molecular classifications and targeted therapies in HCC. Sorafenib, a small-molecular inhibitor of several tyrosine protein kinases, is the only targeted therapeutic drug available for advanced HCC patients, yet it only offers limited survival benefit (54). Therapeutic merit of targeting Wnt/β-catenin pathway has not been established for HCC treatment, which could partially be attributed to the heterogeneous alterations of the activation of this signaling pathway. Clearly, an urgent need exists to decipher HCC genomic architecture, clonal evolution as well as subclonal diversification.

This study represents the first comprehensive genomic and epigenomic investigations of HCC intratumoral heterogeneity. Notably, we found evidence of spatial diversity with respect to DNA mutations in all 11 HCC cases. In support of this finding, we noted that 45.5% (87/191) TCGA-sequenced HCC patient harbored at least 10% subclonal mutations. The extent of subclonal diversification of HCC was comparable with esophageal and lung adenocarcinoma (25, 29, 55), but lower than clear cell renal cell carcinoma (30).

We demonstrated that many driver mutations (10/34) were mapped in the branches, suggesting that they occurred after the founding clone was established and contributed to the growth advantage of sub-populations. Further CCF analysis of each variant in each tumor region showed an illusion of clonal dominance of a number of driver events, where they appeared fully clonal in some tumor regions but were undetectable in others. This drastic CCF variation was not caused by tumor cellularity as skewed clonality was not observed in any of HCC samples, and overall the cancer cell content was comparable among intratumoral regions (Supplementary Fig. 11). This result challenges the accuracy of the approaches that estimate the clonal status of driver events from a single biopsy for developing biomarkers and therapeutic targets. Although independently-acquired somatic variants in distinct subclones were observed in some other tumor types (29, 30), this phenomenon was not present in this HCC cohort.

Conceivably, therapeutic targeting of truncal drivers might provide more anti-neoplastic efficacy than targeting branch divers. Indeed, it has been shown that inhibiting subclonal

drivers in multiple myeloma only achieved partial treatment efficacy at best (26). In some cases targeting subclonal drivers even resulted in growth acceleration of non-mutated subpopulations, underscoring the importance of analysis of intratumoral heterogeneity for treatment designing and stratification (26). Our findings that *TP53* mutations are always truncal might have clinical relevance for HCC treatment, albeit p53-based cancer therapies have yet to achieve clinical success. In addition, variants in several potential drug targets, including c-KIT and SYK, also ubiquitously occurred in all tumor regions in this cohort. Analysis of TCGA data also revealed that most of *KIT* mutations in HCC were clonal (Fig. 3), suggesting its potential therapeutic merit for this malignancy. In addition to the implications for targeted therapy, our analysis of the single-biopsy WES data from the TCGA HCC cohort revealed the potential prognostic value of intratumoral genomic heterogeneity, which is in agreement with the findings in other cancer types suggesting that higher magnitude of intratumoral heterogeneity is associated with more aggressive malignancy (45).

Although M-WES is able to resolve more subclonal diversity than single-biopsy strategy, to determine the true extent of intratumoral heterogeneity in HCC is still difficult. In addition, the prevailing view that Darwinian selection drives cancer evolution was challenged by recent reports (56, 57). Future studies through multiregional whole-genome sequencing of more tumor regions with deeper coverage are required to decipher the molecular mechanisms associated with the clonal expansion of HCC (56, 58). In the present cohort, 10/11 individuals were HBV positive (Supplemental Table 2), and future work with virus free HCCs will shed light on the intratumoral heterogeneity resulting from viral integrations. Moreover, longitudinal multiregional genomic studies will help unravel whether and how intratumoral heterogeneity contribute to drug resistance and distal metastasis of HCC cells.

HCC methylomes have been characterized using a single-sampling approach; however, little is known regarding intratumoral diversity at the epigenetic level or its relationship to changes in the genome. Here, we have shown profound epigenetic intratumoral heterogeneity in HCC. In order to minimize and alleviate the confounding effects from non-tumor DNA contamination, additional approaches were performed to account for the potential influence from immune cells, and similar results of intratumoral methylation heterogeneity were observed (Supplementary Tables 6, see Methods). Nevertheless, one should keep in mind that while intratumoral methylation heterogeneity possibly reflects complex heterogeneous populations of cancer cells, it may partially result from “phenotypic plasticity”, where biological processes (e.g., epigenetic modifications) of the same cells can change depending on their own differentiation state as well as the microenvironment. Close inspection of heterogeneously altered methylation sites revealed that a number of known cancer-related genes, such as *APC, IRF6, RUNX3, CDKN2A*, were differentially methylated between different tumor regions. These changes may have an impact on the cellular biology between different sub-populations. Importantly, albeit the number of differentially-methylated probes varied substantially between different individuals, our enrichment analysis (Fig. 5B) and GO pathway annotation (Fig. 5C) strongly suggest that intratumoral methylation heterogeneity likely plays important roles in the biology of HCC. Interestingly, although a clear interplay and co-dependence was found between phylogenetic and phyloepigenetic evolution in our investigations on esophageal tumors (13), we did not

observe such a relationship in HCC in the present study (data not shown). It is also worthwhile noting that even tumors with “intratumoral-quiet” genome could still exhibit profound epigenetic heterogeneity (as demonstrated by the cases HCC5647 and HCC8010), suggesting that complex intratumoral heterogeneity can occur on multiple layers. Finally, we inferred methylation intratumoral heterogeneity based on the TCGA HCC cohort using a new analytic metric, mITH, and found that a higher degree of methylation heterogeneity was associated with a worse clinical outcome in HCC. These results, along with GO functional classes associated with variably methylated genes, suggest that epigenetic diversity could have functional relevance for HCC biology.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Financial Support

This research was supported by the National Research Foundation Singapore under its Singapore Translational Research Investigator Award (NMRC/STaR/0021/2014) and administered by the Singapore Ministry of Health’s National Medical Research Council (NMRC), the NMRC Centre Grant awarded to National University Cancer Institute of Singapore, the National Research Foundation Singapore and the Singapore Ministry of Education under its Research Centres of Excellence initiatives to H.P. Koeffler. D-C. Lin was supported by National Center for Advancing Translational Sciences UCLA CTSI Grant UL1TR000124 and NSFC (81672786). D. Yin was supported by grants from the NSFC (81572484, 81420108026, 81621004) and Guangdong Science and Technology Department (2014A050503026, 2015B050501004). B.P. Berman and H.Q. Dinh were partially supported by Cedars-Sinai Department of Biomedical Sciences and Samuel Oschin Comprehensive Cancer Institute.

We thank Dr. Hui Shen (Van Andel Institute) for providing a R script to analyze immune cell fraction, and Dr. Dan Weisenberger (University of Southern California) for the help on methylation analysis. We are grateful for the generous donation from Blance and Steven Korgler as well as the Melamed family.

## References

1. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380:2095–128. [PubMed: 23245604]
2. Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. Lancet. 2012; 379:1245–55. [PubMed: 22353262]
3. El-Serag HB. Hepatocellular carcinoma. The New England journal of medicine. 2011; 365:1118–27. [PubMed: 21992124]
4. Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nature genetics. 2012; 44:694–8. [PubMed: 22561517]
5. Kan Z, Zheng H, Liu X, Li S, Barber TD, Gong Z, et al. Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. Genome research. 2013; 23:1422–33. [PubMed: 23788652]
6. Totoki Y, Tatsuno K, Covington KR, Ueda H, Creighton CJ, Kato M, et al. Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. Nature genetics. 2014; 46:1267–73. [PubMed: 25362482]
7. Fujimoto A, Totoki Y, Abe T, Boroevich KA, Hosoda F, Nguyen HH, et al. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. Nature genetics. 2012; 44:760–4. [PubMed: 22634756]

8. Li M, Zhao H, Zhang X, Wood LD, Anders RA, Choti MA, et al. Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. Nature genetics. 2011; 43:828–9. [PubMed: 21822264]
9. Schulze K, Imbeaud S, Letouze E, Alexandrov LB, Calderaro J, Rebouissou S, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nature genetics. 2015; 47:505–11. [PubMed: 25822088]
10. Huang J, Deng Q, Wang Q, Li KY, Dai JH, Li N, et al. Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. Nature genetics. 2012; 44:1117–21. [PubMed: 22922871]
11. Totoki Y, Tatsuno K, Yamamoto S, Arai Y, Hosoda F, Ishikawa S, et al. High-resolution characterization of a hepatocellular carcinoma genome. Nature genetics. 2011; 43:464–9. [PubMed: 21499249]
12. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013; 500:415–21. [PubMed: 23945592]
13. Hao J-J, Lin D-C, Dinh HQ, Mayakonda A, Jiang Y-Y, Chang C, et al. Spatial intratumor heterogeneity of genetic, epigenetic alterations and temporal clonal evolution in esophageal squamous cell carcinoma. Nat Genet. 2016; doi: 10.1038/ng.3683
14. Xue R, Li R, Guo H, Guo L, Su Z, Ni X, et al. Variable Intra-Tumor Genomic Heterogeneity of Multiple Lesions in Patients With Hepatocellular Carcinoma. Gastroenterology. 2016; 150:998–1008. [PubMed: 26752112]
15. Stachler MD, Taylor-Weiner A, Peng S, McKenna A, Agoston AT, Odze RD, et al. Paired exome analysis of Barrett's esophagus and adenocarcinoma. Nature genetics. 2015; 47:1047–55. [PubMed: 26192918]
16. Triche TJ Jr, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic acids research. 2013; 41:e90. [PubMed: 23476028]
17. Fleischhacker M, Lee S, Spira S, Takeuchi S, Koeffler HP. DNA aneuploidy in morphologically normal colons from patients with colon cancer. Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc. 1995; 8:360–5.
18. Cooper CS, Eeles R, Wedge DC, Van Loo P, Gundem G, Alexandrov LB, et al. Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. Nature genetics. 2015; 47:367–72. [PubMed: 25730763]
19. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. Nature reviews Cancer. 2011; 11:9–22. [PubMed: 21160525]
20. Ahn SM, Jang SJ, Shim JH, Kim D, Hong SM, Sung CO, et al. Genomic portrait of resectable hepatocellular carcinomas: implications of RB1 and FGF19 aberrations for patient stratification. Hepatology. 2014; 60:1972–82. [PubMed: 24798001]
21. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, et al. A census of human cancer genes. Nature reviews Cancer. 2004; 4:177–83. [PubMed: 14993899]
22. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature. 2014; 505:495–501. [PubMed: 24390350]
23. McGranahan N, Favero F, de Bruin EC, Birkbak NJ, Szallasi Z, Swanton C. Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. Sci Transl Med. 2015; 7:283ra54.
24. Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat Med. 2015; 21:751–9. [PubMed: 26099045]
25. Murugaesu N, Wilson GA, Birkbak NJ, Watkins TB, McGranahan N, Kumar S, et al. Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. Cancer Discov. 2015; 5:821–31. [PubMed: 26003801]
26. Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D, et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer cell. 2014; 25:91–101. [PubMed: 24434212]

27. Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell. 2013; 152:714–26. [PubMed: 23415222]
28. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013; 499:214–8. [PubMed: 23770567]
29. de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science. 2014; 346:251–6. [PubMed: 25301630]
30. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nature genetics. 2014; 46:225–33. [PubMed: 24487277]
31. Revill K, Wang T, Lachenmayer A, Kojima K, Harrington A, Li J, et al. Genome-wide methylation analysis and epigenetic unmasking identify tumor suppressor genes in hepatocellular carcinoma. Gastroenterology. 2013; 145:1424–35.e1. [PubMed: 24012984]
32. Shen J, Wang S, Zhang YJ, Kappil M, Wu HC, Kibriya MG, et al. Genome-wide DNA methylation profiles in hepatocellular carcinoma. Hepatology. 2012; 55:1799–808. [PubMed: 22234943]
33. Calvisi DF, Ladu S, Gorden A, Farina M, Lee JS, Conner EA, et al. Mechanistic and prognostic significance of aberrant methylation in the molecular pathogenesis of human hepatocellular carcinoma. J Clin Invest. 2007; 117:2713–22. [PubMed: 17717605]
34. Brocks D, Assenov Y, Minner S, Bogatyrova O, Simon R, Koop C, et al. Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. Cell Rep. 2014; 8:798–806. [PubMed: 25066126]
35. Mazor T, Pankov A, Johnson BE, Hong C, Hamilton EG, Bell RJ, et al. DNA Methylation and Somatic Mutations Converge on the Cell Cycle and Define Similar Evolutionary Histories in Brain Tumors. Cancer cell. 2015; 28:307–17. [PubMed: 26373278]
36. Mazor T, Pankov A, Song JS, Costello JF. Intratumoral Heterogeneity of the Epigenome. Cancer cell. 2016; 29:440–51. [PubMed: 27070699]
37. Hao JJ, Lin DC, Dinh HQ, Mayakonda A, Jiang YY, Chang C, et al. Spatial intratumoral heterogeneity and temporal clonal evolution in esophageal squamous cell carcinoma. Nature genetics. 2016; 48:1500–7. [PubMed: 27749841]
38. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. Nature biotechnology. 2012; 30:413–21.
39. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. Nature genetics. 2013; 45:1134–40. [PubMed: 24071852]
40. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlen SE, Greco D, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PloS one. 2012; 7:e41361. [PubMed: 22848472]
41. Zhang Y, Liu H, Lv J, Xiao X, Zhu J, Liu X, et al. QDMR: a quantitative method for identification of differentially methylated regions by entropy. Nucleic acids research. 2011; 39:e58. [PubMed: 21306990]
42. Bergman Y, Cedar H. DNA methylation dynamics in health and disease. Nature structural & molecular biology. 2013; 20:274–81.
43. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nature reviews Cancer. 2011; 11:726–34. [PubMed: 21941284]
44. Yao L, Shen H, Laird PW, Farnham PJ, Berman BP. Inferring regulatory element landscapes and transcription factor networks from cancer methylomes. Genome biology. 2015; 16:105. [PubMed: 25994056]
45. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. Cancer cell. 2015; 27:15–26. [PubMed: 25584892]
46. Mroz EA, Tward AD, Hammon RJ, Ren Y, Rocco JW. Intra-tumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from the Cancer Genome Atlas. PLoS medicine. 2015; 12:e1001786. [PubMed: 25668320]

47. Mroz EA, Tward AD, Pickering CR, Myers JN, Ferris RL, Rocco JW. High intratumor genetic heterogeneity is related to worse outcome in patients with head and neck squamous cell carcinoma. Cancer. 2013; 119:3034–42. [PubMed: 23696076]

48. Kenmochi K, Sugihara S, Kojiro M. Relationship of histologic grade of hepatocellular carcinoma (HCC) to tumor size, and demonstration of tumor cells of multiple different grades in single small HCC. Liver. 1987; 7:18–26. [PubMed: 3033422]

49. An FQ, Matsuda M, Fujii H, Tang RF, Amemiya H, Dai YM, et al. Tumor heterogeneity in small hepatocellular carcinoma: analysis of tumor cell proliferation, expression and mutation of p53 AND beta-catenin. Int J Cancer. 2001; 93:468–74. [PubMed: 11477549]

50. Friemel J, Rechsteiner M, Frick L, Bohm F, Struckmann K, Egger M, et al. Intratumor heterogeneity in hepatocellular carcinoma. Clin Cancer Res. 2015; 21:1951–61. [PubMed: 25248380]

51. Bioulac-Sage P, Rebouissou S, Thomas C, Blanc JF, Saric J, Sa Cunha A, et al. Hepatocellular adenoma subtype classification using molecular markers and immunohistochemistry. Hepatology. 2007; 46:740–8. [PubMed: 17663417]

52. Bioulac-Sage P, Cubel G, Balabaud C, Zucman-Rossi J. Revisiting the pathology of resected benign hepatocellular nodules using new immunohistochemical markers. Semin Liver Dis. 2011; 31:91–103. [PubMed: 21344354]

53. Zucman-Rossi J, Jeannot E, Nhieu JT, Scoazec JY, Guettier C, Rebouissou S, et al. Genotype-phenotype correlation in hepatocellular adenoma: new classification and relationship with HCC. Hepatology. 2006; 43:515–24. [PubMed: 16496320]

54. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, et al. Sorafenib in advanced hepatocellular carcinoma. The New England journal of medicine. 2008; 359:378–90. [PubMed: 18650514]

55. Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. Science. 2014; 346:256–9. [PubMed: 25301631]

56. Ling S, Hu Z, Yang Z, Yang F, Li Y, Lin P, et al. Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution. Proceedings of the National Academy of Sciences of the United States of America. 2015; 112:E6496–505. [PubMed: 26561581]

57. Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, et al. A Big Bang model of human colorectal tumor growth. Nature genetics. 2015; 47:209–16. [PubMed: 25665006]

58. Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature. 2014; 512:155–60. [PubMed: 25079324]

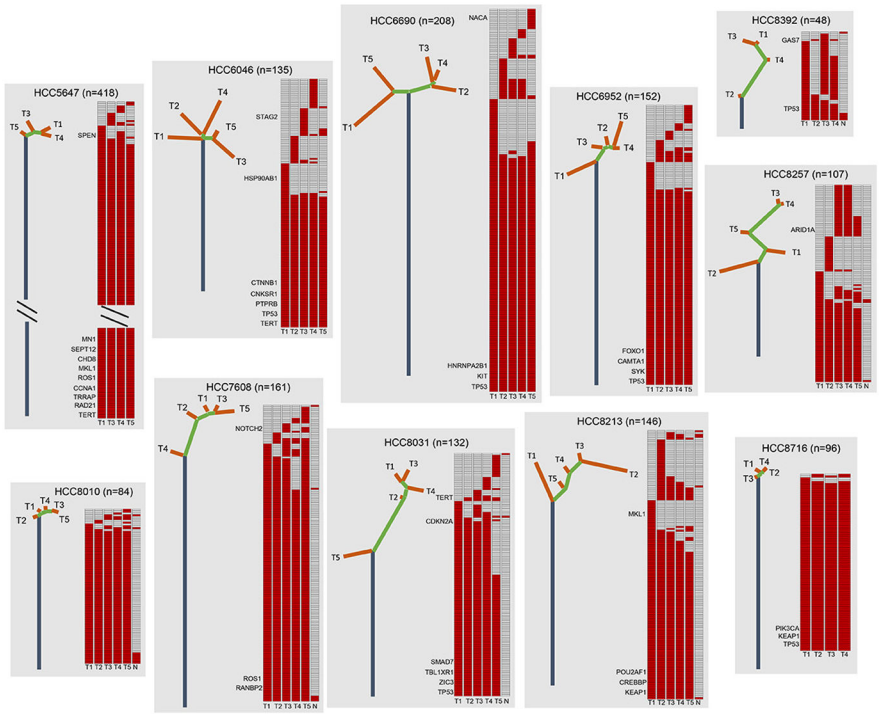


**Figure 1. Phylogenetic trees of 11 HCCs constructed on the basis of M-WES**

Phylogenetic trees were constructed from all somatic variants by Wagner parsimony method using PHYLIP (See Method). Blue, green, and orange lines represent trunk, shared branch, and private branch, respectively. Lengths of trunk and branch are proportional to their number of mutations. Putative driver events are mapped along the trees as indicated. Heat maps indicate the presence (red) or absence (grey) of a mutation in each tumor region (T) or matched nonmalignant liver tissue (N). For case HCC5647, trunk length was reduced for display purpose.

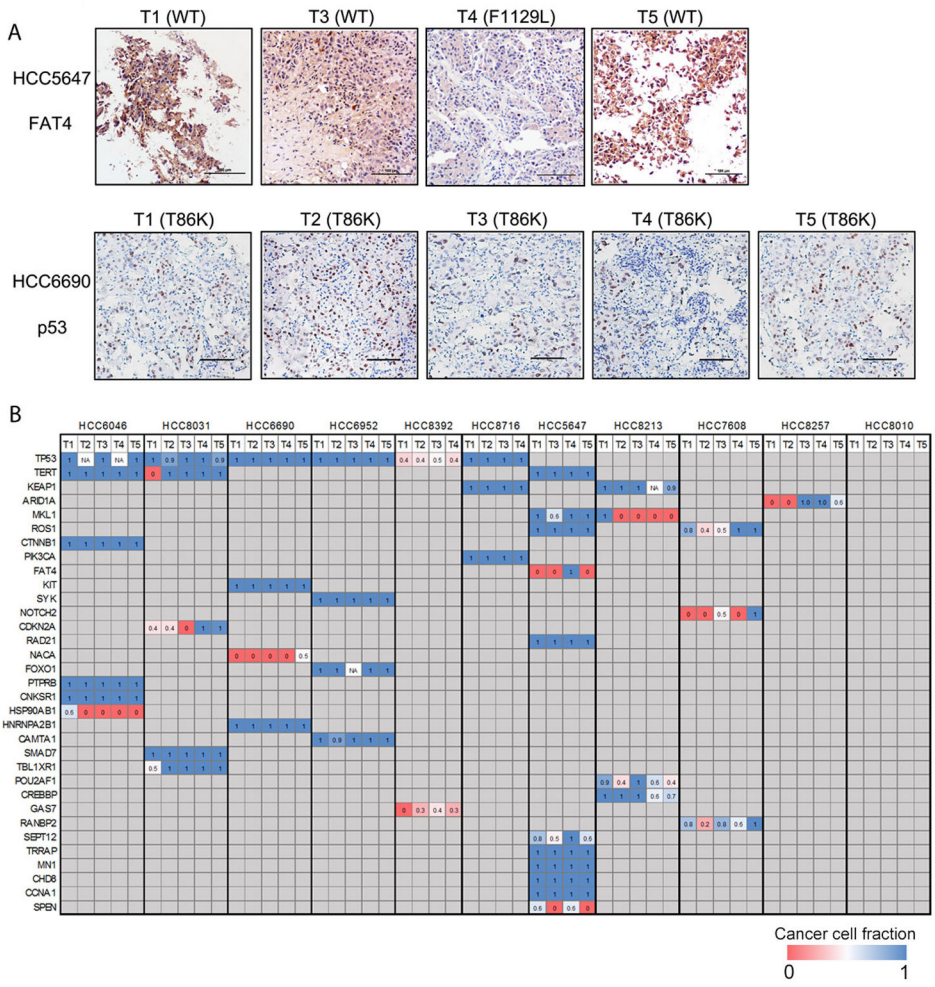


**Figure 2. Spatial heterogeneity and CCF of putative driver mutations in HCC**

(A) IHC staining of FAT4 and p53 in different tumor regions which have been profiled by M-WES. Mutational status is indicated on top of each region. Scale bars, 100 μm. WT, wildtype. (B) Heat map of the CCF of putative driver mutations. Numeric number in each square shows the CCF. Columns, tumor regions; Rows, genes.

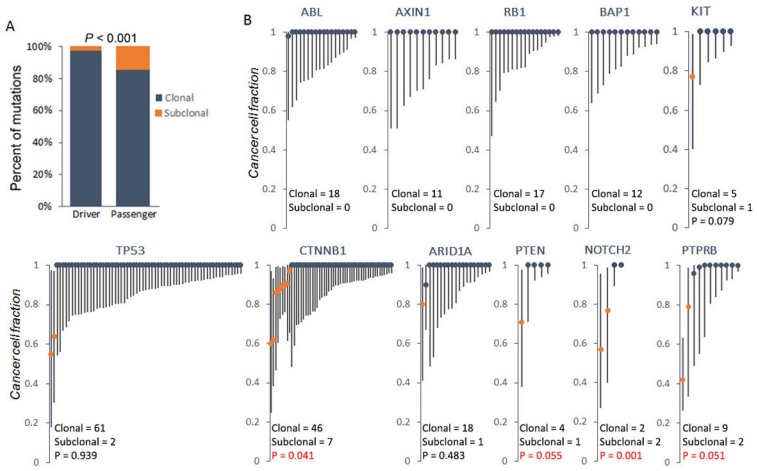


**Figure 3. Clonal status of HCC driver mutations sequenced by TCGA**

(A) The proportion of clonal and subclonal mutations in both driver and passenger events. (B) CCF analysis of mutations in representative driver genes from TCGA HCC dataset. Each line represents an individual mutation. Round dot, upper and lower end of each line represents CCF, upper and lower bound of confidence interval, respectively. Clonal and subclonal CCF are shown as dark blue and orange, respectively. Driver genes were identified using MutSig algorithm (28) on the basis of the TCGA data (FDR $q < 0.1$). P values were derived by hypergeometric tests comparing the frequency of subclonal mutations in each gene against that in all driver genes. In those genes without any subclonal mutations, P value was not calculated.

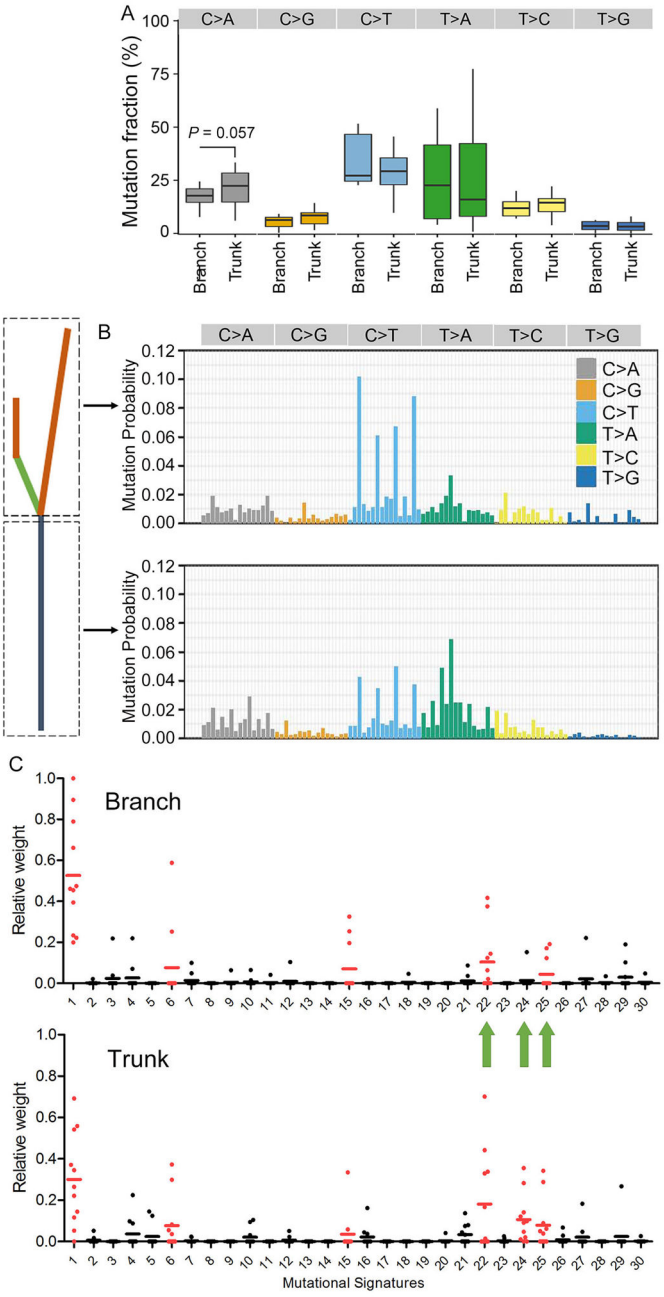


**Figure 4. Temporal dissection of HCC mutational spectra and signatures**

(A) Pairwise fraction analysis of truncal and branch variants on the basis of the six mutation classes. P values were derived by Paired Students' T-test upon verification of normality and variance within each group, with those over 0.1 not shown. (B) Mutational signatures of all truncal and branch variants was inferred by deconstructSigs. Signatures are displayed according to the 96-substitution classification defined by the substitution class and sequence context (12). (C) Dot plots display the contributions of individual mutational signatures to individual cases, with each dot representing one case. Signatures 1–30 were based on the Wellcome Trust Sanger Institute Mutational Signature Framework (12).

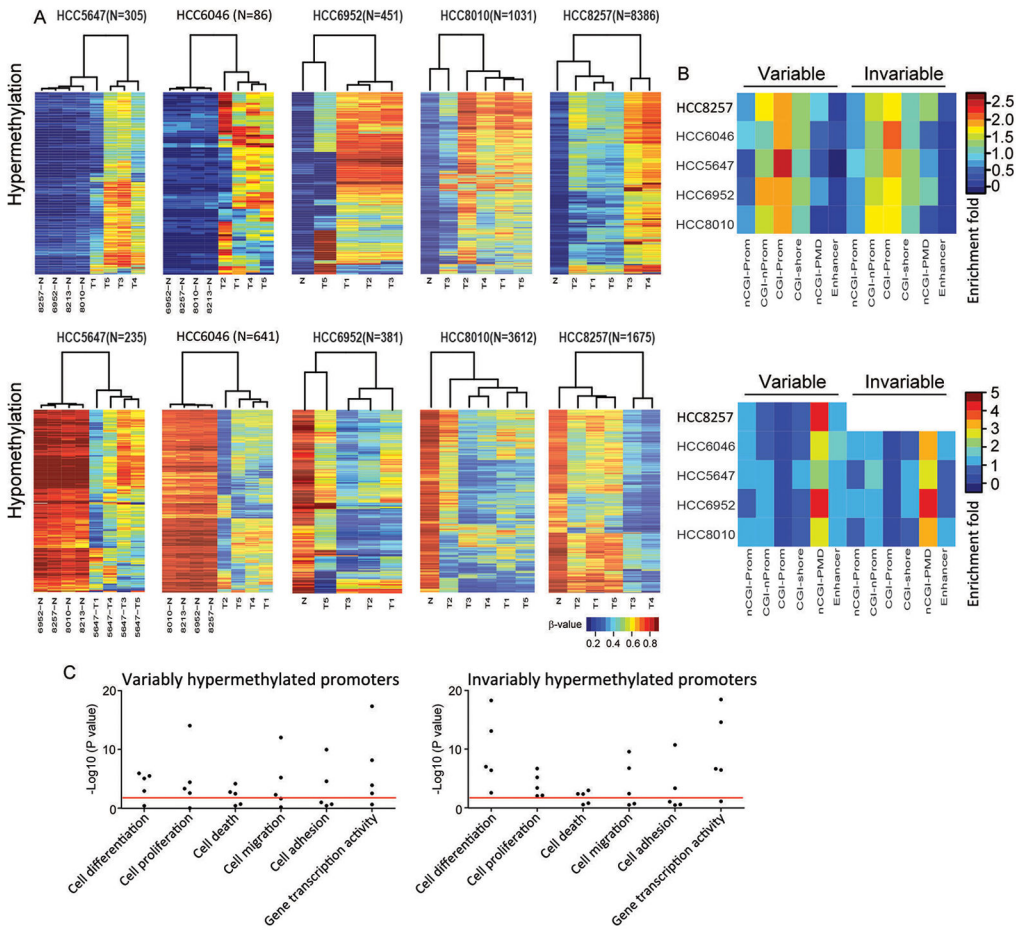


**Figure 5. Epigenetic intratumoral heterogeneity in HCC**

(A) Unsupervised hierarchical clustering of intratumoral methylation profiling of five HCC cases. Rows of the heat maps denote the methylation levels of variably hypermethylated (upper panel) or hypomethylated (upper panel) CpG sites across different tumor regions. Columns represent samples. (B) Enrichment plots showing the distribution of both variably and invariably hypermethylated (upper panel) or hypomethylated (lower panel) CpG sites across a variety of functional genomic domains. nCGI-Prom (non CG island Promoter); CGI-nProm, (CG islands not in promoter regions); CGI-Prom, CG island promoters; nCGI-PMD, PMD excluding CGI probes. All *P* values of hypergeometric enrichment test (comparing the frequency of each variable and invariable probe set category to that of array background) were shown in the Supplementary Table 6. (C) Dot plots displaying enriched GO biological processes for the genes associated with variably and invariably hypermethylated promoters, with each dot representing one individual case. Red line indicates P value of 0.01.