Spatial and temporal expansion of intrahepatic metastasis by molecularly-defined clonality in multiple liver cancers

Shogo Yamamoto1 | Yutaka Midorikawa1,2 | Genta Nagae1 | Kenji Tatsuno1 | Hiroki Ueda1 | Mitsuhiko Moriyama3 | Tadatoshi Takayama2 | Hiroyuki Aburatani1

Abstract
Multiple hepatocellular carcinoma (HCC) is divided into two categories: intrahepatic metastasis (IM), which is a true relapse of HCC, and multicentric origin (MO), which is a second primary tumor. Clinical diagnosis of multiple HCC is usually made based on tumor location and/or time to recurrence; however, it is often difficult to distinguish the two types of multiple HCC. Using 41 matched pairs of multiple HCC specimens, we confirmed the accuracy of clinical diagnoses using exome sequence data and investigated the importance of discriminating the type of multiple HCC. Genomic analysis revealed that 18 (43.9%) patients diagnosed as having genomic IM had common mutations in a pair of HCC tumors with the main tumor of these patients being more progressive compared to those with genomic MO. The accuracy of clinical diagnosis based on lobe (Definition 1) and segment (Definition 2) were 68.3% and 78.0%, respectively. Intriguingly, recurrence ≥2 years after initial surgery for 3 patients was IM. The survival of patients with clinical IM was significantly shorter than for those with clinical MO based on both Definition 1 (P = 0.045) and Definition 2 (P = 0.043). However, mean survival was not different between the patients with genomic IM and those with MO (P = 0.364). Taken together, genomic analysis elucidated that liver cancer may spread more extensively and more slowly than previously thought. In addition, distinguishing multiple HCC as IM or MC may have provided biological information but was not of clinical importance with respect to patient prognosis.

KEYWORDS
diagnosis, exome sequence, intrahepatic metastasis, multicentric origin, multiple hepatocellular carcinoma

1 INTRODUCTION
Multiplicity is a major clinical features of hepatocellular carcinoma (HCC). Hepatocellular carcinoma can spread to the other regions of the liver via portal vein invasion,1,2 which is referred to as intrahepatic metastasis (IM). In contrast, multiple tumors may originate from different clones that develop as a result of underlying chronic liver disease due to viral hepatitis, alcohol abuse or non-alcoholic steatohepatitis.3

Shogo Yamamoto and Yutaka Midorikawa contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
which are considered to be of multicentric origin (MO). Thus, multiple HCC can be divided into two categories according to the mechanism of carcinogenesis.

Clinical classification of multiple HCC is based on tumor location, time to recurrence, and background liver; however, it is difficult to clearly distinguish IM from MO using only those clinical parameters. Pathologically, well-differentiated HCC rarely develops IM, while small recurrent tumors are interpreted to be IM if their differentiation grade is moderate or poor, even though the possibility of MO cannot be completely excluded. Hepatocellular carcinoma exhibits a higher incidence of IM via portal vein invasion when macroscopic findings of the tumor are of simple nodular type with extranodular growth or of multinodular confluent type, even if in the early stage. Therefore, the pathological diagnosis of multiple HCC is done based on differentiation grade or the gross type of the tumor.

In contrast to clinical and pathological diagnoses, molecular biological approaches are able to provide a definite diagnosis of IM or MO in multiple HCC. For instance, comparing the integration pattern of hepatitis B virus into the HCC genome by Southern blotting and/or PCR analysis allows for the differentiation of IM from MO. In addition, comparative genomic hybridization analysis, loss of heterozygosity analysis of DNA microsatellites and identification of mitochondrial DNA mutations within the D-loop control region have also been useful for identifying tumor clonality. Other diagnostic criteria that have been used are based on the promoter hypermethylation status of tumor suppressor genes.

Next-generation sequencing (NGS) technology is currently being used to identify specific mutation signatures of HCC. This approach has been applied to accurately diagnose types of multiple HCC and has shown that whole-exome sequencing can be used to confidently classify all cases of multiple HCC, unlike targeted sequencing. In addition to discriminating IM and MO, physical changes during cancer progression can also be elucidated in individuals. However, the number of patients in these studies was relatively small due to the difficulty of molecular analysis, and, therefore, the clinical significance of differentiating IM from MO could not be determined.

In the current study, we performed whole-exome sequence analysis of paired HCC tumor specimens and adjacent non–tumorous liver specimens obtained from the same patients to diagnose multiple HCC as IM or MO. We also analyzed the discrepancies between genomic and clinical diagnoses. Furthermore, we compared the survival of patients with IM and those with MO and evaluated the value of accurate multiple HCC diagnosis.

2 | MATERIAL AND METHODS

2.1 | Patients

The study group was comprised of patients with multiple HCC who underwent liver resection from January 2011 to December 2017 at the Department of Digestive Surgery, Nihon University, Tokyo, Japan. The study was approved by the institutional review boards of the Nihon University School of Medicine (protocol number; 131) and all participants provided written informed consent. In the case with three tumors in metachronous multiple HCC, the larger two tumors were analyzed for sequencing. Surgical specimens were immediately dissected into small pieces after liver resection, snap frozen in liquid nitrogen and stored at −80°C.

2.2 | Diagnosis of multiple hepatocellular carcinoma

Clinically, two definitions of multiple HCC were used in this study based on tumor location and time for recurrence. All cases of patients with metachronous multiple HCC were classified as MO if HCC relapsed 2 years or more after the initial operation. Patients with synchronous multiple HCC or recurrence within 2 years after the initial operation were considered to have IM if the tumors were located in the same lobe (Definition 1) or if the tumors were in the same segment proposed by Healey et al (Definition 2). For HCC tumors located in the caudate lobe, the caval portion and process portion were defined as right lobe while the Spiegel portion was defined as left lobe. The main tumor in metachronous multiple HCC was defined as the tumor resected at initial resection or as the larger tumor in synchronous multiple HCC. Based on DNA sequencing data, two tumors from the same patient with common mutations were diagnosed as IM (genomic IM), while MO was diagnosed if there were no common mutations (genomic MO).

2.3 | DNA extraction

Genomic DNA was isolated from HCC tumor specimens and adjacent non–tumorous liver specimens using a QIAamp DNA Mini Kit (Qiagen) according to the manufacturer’s protocols. Genomic DNA concentrations were determined using a Qubit dsDNA BR Assay Kit (Life technologies). One microgram of genomic DNA from each sample was used for the whole-exome sequencing procedure.

2.4 | Exome sequencing

Exome sequencing was performed on 41 pairs of HCC tumor specimens and matched adjacent non–tumorous liver specimens according to the manufacturer’s protocol. Briefly, DNA was fragmented using a Covaris SS Ultrasonicator. Exome capture was performed using Agilent SureSelect V4/V5 (Agilent Technologies) or HGSC VCRome 2.1 design1 (42 Mb, NimbleGen). Each sample was sequenced with a HiSeq 2000 system (Illumina) as 100-bp paired-end reads. Burrows-Wheeler Aligner and NovoAlign software (Novocraft Technologies) were used to align the reads to the human reference genome hg19. After removal of PCR duplicates, the Short-Read Micro re-Aligner was used to improve variant discovery through local realignments.
2.5 | Mutation calling

To identify somatic mutations, we used the integrated genotyper software karkinos (http://github.com/genome-rcastr/karkinos) as described previously. Briefly, variant allele frequencies of somatic mutation were adjusted using estimated tumor content ratios and filtered with a heuristic filtering algorithm and Fisher’s exact test. Variants with allele frequency ≥15%, read depth of tumor ≥15, depth of normal ≥10 and variant read ≥3 were retained. The contribution of known mutational signatures for each sample was determined using the deconstructSigs Catalogue of Somatic Mutations in Cancer (COSMIC) mutational signatures v2.23

2.6 | P-values for driver mutations

Driver mutation probabilities were calculated for each gene. An initial probability for observing a recurrent somatic mutation on the same gene was calculated using a binomial probability equation with gene length and background mutation rates. False discovery rates were calculated by simulation as described previously.

2.7 | Statistical analysis

Fisher’s exact test for categorical variables and Student’s t-test and Wilcoxon rank sum test for continuous variables were used to assess the statistical significance of the data collected from early and classical HCC groups. For patients with metachronous HCC, survival time was defined as the period from the second operation to the date of death. Survival curves were generated using the Kaplan-Meier product-limit estimator and were compared using the log-rank test. Statistical significance was set at $P < 0.05$.

3 | RESULTS

3.1 | Patient enrollment

Between 2011 and 2017, 968 patients underwent curative liver resection for HCC: 825 patients for the initial resection, 103 patients for the second resection and 40 patients for third or further resection (Figure 1). Among the 825 patients who underwent the initial resection, 194 patients had multiple tumors, from which we collected “synchronous multiple HCC” samples ($n = 20$). In contrast, we collected “metachronous multiple HCC” samples ($n = 21$) from 103 patients who underwent the second resection. During the follow-up period at the outpatient clinic (from January 2011 to June 2019), 404 patients (48.9%) out of 825 HCC patients relapsed.

3.2 | Clinical diagnosis

Of the 41 patients that underwent curative liver resection for multiple HCC and were included in our study (Figure 2), 21 (51.2%) had metachronous HCC, of which 10 (24.3%) underwent a second surgery 2 years or more after initial resection and 11 (26.8%) relapsed within 2 years of the initial resection. All 10 (24.3%) patients that underwent a second surgery ≥2 years after initial resection were clinically diagnosed as having MO HCC. A total of 20 (48.8%) patients were diagnosed as having synchronous HCC. For the patients diagnosed as having clinical IM from the 11 patients that relapsed within 2 years after resection and the 20 patients with synchronous multiple HCC, 20 (48.7%) had two tumors in the same lobe (Definition 1) and 14 (34.1%) had two tumors in the same segment (Definition 2). From this same group that was diagnosed as having clinical MO, 11 (26.8%) had two tumors in the same lobe (Definition 1) and 17 (41.5%) had two tumors in the same segment (Definition 2).

FIGURE 1 Flowchart of the enrollment of patients. LR, liver resection

FIGURE 2 Flowchart of the clinical diagnosis of multiple hepatocellular carcinoma (HCC)
3.3 Genomic diagnosis

Exome sequencing results revealed that there were 81.7 ± 30.4 common mutations in 18 (43.9%) patients diagnosed as having genomic IM (Figure 3). There were 127.0 ± 54.4 and 122.0 ± 57.7 somatic point mutations per tumor in the genomic IM and MO samples, respectively ($P = 0.577$), and 11.9 ± 7.1 and 11.2 ± 6.4 somatic indels per tumor, respectively ($P = 0.638$). Most of the somatic substitutions in the genomic IM and MO samples were transitions and the distribution did not significantly differ between the two types of HCC ($P = 0.343$) (Table 1).

Based on Definition 1, 12 of 18 patients with clinical IM (66.6% sensitivity) and 16 of 23 patients with clinical MO (69.5% specificity) were diagnosed as having genomic IM and MO, respectively, with an accuracy of 68.2%. In contrast, 11 of 18 patients with clinical IM by Definition 2 were diagnosed as having genomic IM (61.1% sensitivity) and 20 of 23 patients with clinical MO were diagnosed as having genomic MO (86.9% specificity), with an accuracy of Definition 2 of 75.6% (Table S1). Histologically, among 13 (31.7%) patients with different differentiation grades of two nodules, 12 patients were diagnosed as having genomic MO. Interestingly, among the 10 patients who had recurrence ≥2 years after initial resection and had clinical diagnoses of MO, 3 were diagnosed with genomic IM.

**Figure 3** Genomic diagnosis of multiple hepatocellular carcinoma (HCC). Top, Vertical bars represent the number of mutations. Red bars and the green graph line represent numbers and rates, respectively, of common mutations of multiple HCC from the same patients. Middle, Tumor location and time to recurrence of multiple HCC. Bottom, Clinicopathological characteristics of patients with multiple HCC. *P* value less than 0.05

**Table 1** Mutation pattern of HCC (non–synonymous)

<table>
<thead>
<tr>
<th>Mutation type</th>
<th>Genomic IM (n = 18)</th>
<th>Genomic MO (n = 23)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C&gt;G&gt;T:A (%)</td>
<td>21.6 ± 7.1</td>
<td>21.5 ± 11.0</td>
<td></td>
</tr>
<tr>
<td>T&gt;A&gt;C:G (%)</td>
<td>31.6 ± 17.6</td>
<td>33.8 ± 20.9</td>
<td></td>
</tr>
<tr>
<td>C&gt;G&gt;G:C (%)</td>
<td>10.4 ± 5.7</td>
<td>9.7 ± 5.6</td>
<td></td>
</tr>
<tr>
<td>C&gt;G&gt;A:T (%)</td>
<td>21.6 ± 7.1</td>
<td>21.5 ± 11.0</td>
<td></td>
</tr>
<tr>
<td>T&gt;A&gt;G:C (%)</td>
<td>9.0 ± 5.3</td>
<td>9.5 ± 7.7</td>
<td></td>
</tr>
<tr>
<td>T&gt;A&gt;A:T (%)</td>
<td>10.0 ± 5.5</td>
<td>10.7 ± 6.9</td>
<td></td>
</tr>
<tr>
<td>SNV</td>
<td>127.0 ± 54.4</td>
<td>120.0 ± 57.7</td>
<td>0.577</td>
</tr>
<tr>
<td>Indel</td>
<td>11.9 ± 7.1</td>
<td>11.2 ± 6.4</td>
<td>0.638</td>
</tr>
<tr>
<td>Total</td>
<td>138.9 ± 57.9</td>
<td>131.2 ± 61.9</td>
<td>0.565</td>
</tr>
</tbody>
</table>

Note: Data was shown as average with standard deviation. Abbreviations: HCC, hepatocellular carcinoma; IM, intrahepatic metastasis; MO, multicentric origin; SNV, single nucleotide variation.
3.4 | Clinical characteristics

The main tumor was more progressive in patients with genomic IM compared to those with genomic MO based on significantly higher alpha-fetoprotein levels \( (P = 0.044) \), larger tumor size \( (P = 0.019) \) and frequent vascular invasion \( (P = 0.027) \) (Table 2). In contrast, liver cirrhosis and liver function, including indocyanine green clearance rate at 15 minutes and frequency of Child-Pugh A classification, were not significantly different between the two groups.

3.5 | Mutation landscape

Exome sequence analysis revealed that several driver genes in genomic IM and MC, including CTNNB1 (33.3% vs 36.9%, \( P = 0.817 \)), TP53 (27.7% vs 28.2%, \( P = 1.000 \)), ARID2 (13.8% vs 10.8%, \( P = 0.742 \)), MLL3 (11.1% vs 10.8%, \( P = 1.000 \)) and ARID1A (13.8% vs 6.5%, \( P = 0.290 \)), were more frequently mutated than expected by chance (Figure 4A). Ten patients with genomic IM had mutations in these driver genes in only one of the tumors (4 in CTNNB1, 2 in TP53, 1 in ARID2, 1 in MLL3 and 1 in ARID1A). The mutations were detected in metastatic tumors of five of the 10 patients (Table S2). However, one or two copies, which were below threshold, were detected in the main nodules, suggesting that a small number of cancer cells with these driver mutations metastasized to the other sites of the liver parenchyma (Figure 4B). In contrast, 4 of 23 patients with genomic MO had the same driver gene mutations in both tumors (2 in CTNNB1, 1 in TP53 and 1 in MLL3 genes), but the chromosomal positions of the mutation differed between the two tumors (Figure 4C).

3.6 | Mutational signatures

Non-negative matrix factorization analysis was applied to 96 substitution patterns using 82 pair-matched HCC samples from 41 patients. Three groups of HCC were identified by hierarchical clustering of the samples, which was based on the contributions of the mutational signatures in each sample (Figure 5). Group 1 consisted of 34 HCC enriched in signature 16 and was characterized by the same categorization from the patients in both genomic IM and genomic MO (70.5%, \( P = 0.039 \)). Genomic MO tended to be frequent in this group (\( P = 0.080 \)). Group 2 consisted of 13 HCC samples and correlated with signature 5. Group 3 was enriched in signatures 3, 6 and 12 and had a relatively high prevalence of tumor thrombus of the hepatic vein (37.5%, \( P = 0.049 \)). Despite the fact that paired tumors had common origins, five pairs of genomic IM samples (27.7%) were classified into different groups (Figure S1). In contrast, nine pairs of genomic MO (39.1%) that originated from different clones were clustered in the same group.

3.7 | Survival

After a median follow up of 2.5 years (range, 0.5–6.3 years), the median overall survival was 2.3 years (95% confidence interval [CI], 1.2–not available [NA]) and 5.4 years (3.2–NA, \( P = 0.045 \)) for the patients with clinical IM and clinical MO, respectively, based on Definition 1 (Figure 6A). Overall survival at 3 years was 46.8% and 79.1%, respectively, in the two groups. By contrast, the median overall survival of patients with clinical IM based on Definition 2 was 2.0 years (95% CI, 1.1–NA) and those with clinical MO did not reach the median (3.2–NA, \( P = 0.043 \)) (Figure 6B). Overall survival at 3 years was 38.4% and 78.8%, respectively, in the two groups.

Based on genomic diagnoses, the median overall survival for patients with genomic IM and genomic MO was 3.3 years (95% CI, 1.2–NA) and 5.4 years (2.0–NA, \( P = 0.364 \)), respectively (Figure 6C). The 3-year rates of overall survival were 50.7% and 75.0%, respectively (Figure 6A).
Clinically, the determination of IM and MO for patients with multiple HCC is based on tumor location and time to recurrence. However, our current findings showed that the clinical diagnosis of multiple HCC was not very accurate, which was verified by molecular biological analysis. Consequently, we found that liver cancer cells were able to spread more extensively and more slowly than previously known. We also found that the type of multiple HCC did not necessarily reflect patient survival following tumor resection.

It is generally assumed that liver cancer cells metastasize along the portal vein to other sites of the liver parenchyma and, therefore, satellite lesions within the same segment are likely to be IM. However, genomic diagnosis revealed that tumor size was larger and vascular invasion was more frequent in IM, suggesting that IM could result due to tumor progression. Our data also showed that liver cancer cells frequently metastasized, regardless of the region of the liver. Given that IM could be detected even 2 years post-surgery, it is difficult to determine whether multiple HCC is actually IM or if it is a result of the clinical course. In contrast, most pairs of multiple nodules with different differentiation grade were diagnosed as genomic MO; therefore, histological findings after surgery were available for diagnosis of multiple HCC only when differentiation grades were different between the two nodules.

It was previously reported that MO is associated with poor liver function, and, therefore, such lesions may not be a target for locoregional therapy but rather may be a signaling lesion for advanced stage chronic liver disease because patient survival depends in part on liver function. In contrast to the clinical observation, in the current study liver function, including the frequency of liver cirrhosis, was not significantly different between the two groups, and patient survival after the operation was not significantly different between the MO and IM groups. Therefore, patients with multiple HCC may still be candidates for resection, even for patients with MO HCC.

Similar to previous sequence data, driver genes such as CTNNB1, TP53, ARID2, MLL3 and ARID1A were frequently mutated in multiple HCC.

**FIGURE 4** Landscape of significantly mutated genes. A, Genes that were mutated in eight or more samples are shown. The matrix represents individual mutations in individual patient samples. The color of the box indicates the proportion of mutated cancer cells within the tumor sample according to the scale indicated below. Red indicates that all tumor cells in the specimen contained a mutation in the indicated gene. Blue indicates that half of the tumor cells in the specimen contained a mutation in the indicated gene. For each pair of hepatocellular carcinoma (HCC) specimens from the same patient, the left column represents the main tumor. The green frame (genomic intrahepatic metastasis [IM]) indicates that the driver mutation was present in only one of the patient samples. The orange frame (genomic multicentric origin [MO]) indicates that the driver mutation was present in both patient tumors. B, Genome viewer revealed that a metastatic nodule had multiple copies of a mutation, beyond the threshold, while the main tumor harbored only one copy, which was below the threshold for a mutation-positive signal at the same position in CTNNB1. C, Genome viewer revealed that a pair of samples had mutation at different positions in CTNNB1.
mutated in multiple HCC and there was no significant difference in the frequency of gene mutations between IM and MO samples. Consequently, we failed to identify any genes responsible for IM. The frequency of common mutations in 18 genomic IM samples was only 43%. However, several primary lesions harbored one or two copies of the same mutation that were found in metastatic lesions, which indicated that a small population of cancer cells with malignant potential could metastasize to other sites of the liver. Thus, advanced cancer harbors a high level of heterogeneity of cells. However, if both tumors harbored mutations in the same driver gene in MO samples, the chromosomal positions of the driver gene mutations were different. Given that gene mutations were concentrated on several driver genes, including \textit{CTNNB1} and \textit{TP53} in most samples based on the previous reports on sequence of HCC,\textsuperscript{16,18} genomic diagnosis by clinical sequence, not by whole-exome sequence, is the future for clinical practice in the treatment of multiple HCC.\textsuperscript{20} Specifically, target sequencing using liquid biopsy specimens might be applicable for this purpose in the case of metachronous recurrence.\textsuperscript{27,28}

Based on gene mutational signatures, HCC samples can be divided into three groups.\textsuperscript{29} Consistent with our previous data, Group 1, which consisted of the largest number of samples in this study, was characterized by signature 16. This mutational signature is frequent in Japanese males with HCC\textsuperscript{18} and is associated with \textit{ALDH2} and \textit{ADH1B} single nucleotide polymorphisms,\textsuperscript{30} although frequency of inactivation type of single nucleotide polymorphisms of these genes were not significant in Group 1. Given the high frequency of MO, signature 16 might also be associated with the late recurrence of HCC. Mutational signatures 3 and 6, which are associated with \textit{BRCA1/2} mutations\textsuperscript{31} and microsatellite unstable tumors,\textsuperscript{32} respectively, in addition to signature 16, were representative of Group 3. In contrast, signature 5 was frequent in Group 2 and is non-specific in cancer.

It was interesting that a pair of HCC samples were divided into other clusters in both MO and IM samples, which may be attributed to tumor heterogeneity in hematogenous metastasis.\textsuperscript{33} Consistent with the previous report that cancer development such as metastasis contributes to the mutation burden and causes.
a clonal expansion, metastatic lesions harbored more mutations compared to primary lesions in five pairs of the IM group. Consequently, these samples were also divided into different clusters in mutational signatures.

Whether the prognosis for patients with IM differs from those with MO is controversial. Based on the clinical diagnosis (Definition 1 and Definition 2), overall survival of patients with IM was significantly shorter than that of patients with MO, which may be attributed to tumor malignancy being related to poorer prognosis. By contrast, survival of patients with multiple HCC in the same liver section is longer than that of patients with multiple HCC in different sections. Finally, genomic diagnosis demonstrated in this study that there was no difference in survival between patients with IM and those with MO. This discrepancy between genomic and clinical diagnoses may be because the prognosis of patients with HCC depends on liver function, which tended to be worse in patients with MO. Thus, high-grade tumor malignancy by IM and poorer liver function by MO may offset each other. Therefore, we suggest that it is not clinically useful to determine whether multiple HCC is IM or MO.

The current study had several limitations. This study was a retrospective study and was affected by several biases. That is, the number of HCC specimens which were available was limited (selection bias), and tumor samples included both synchronous and metachronous multiple HCC. Consequently, categorization applied in the present study led to survivor treatment bias; the time to recurrence and survival analysis were affected by immortal time bias and time-dependent bias. To avoid the bias, tumor samples should have been collected prospectively, or the tumor stage of the main tumor should have been established.

In conclusion, exome sequencing was able to be used to discriminate multiple HCC as IM or MO and it became clear that liver cancer cells could spread more extensively and more slowly than previously thought. Given that the overall survival of patients with IM and MO did not differ, the accurate classification of multiple HCC is not clinically important.

ACKNOWLEDGMENTS
We thank Kaori Shiina and Hiroko Meguro for valuable technical assistance and Vincent Stanton for helpful discussions. This research was supported by AMED under grant number JP18hk0102049h0001 and by a grant-in-aid of The 106th Annual Congress of The Japan Surgical Society (JSS) Memorial Surgical Research Fund, Tokyo, Japan.

DISCLOSURE
The authors have no conflict of interest to declare.

ORCID
Shogo Yamamoto https://orcid.org/0000-0002-5564-2567
Yutaka Midorikawa https://orcid.org/0000-0002-6662-0494

REFERENCES

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.