ORIGINAL ARTICLE


C.K.Y. Ng¹,², G. G. Di Costanzo³, N. Tosti¹, V. Paradiso¹, M. Coto-Llerena², G. Roscigno⁴, V. Perrina¹, C. Quintavalle¹, T. Boldanova¹,⁵, S. Wieland², G. Marino-Marsilia⁶, M. Lanzafame¹, L. Quagliata¹, G. Condorelli⁴, M. S. Matter¹, R. Tortora³, M. H. Heim²,⁵, L. M. Terracciano¹ and S. Piscuoglio¹.$

Affiliations

¹Institute of Pathology, University Hospital Basel, Basel, Switzerland;
²Department of Biomedicine, Hepatology Laboratory, University of Basel, Basel, Switzerland;
³Department of Transplantation - Liver Unit, Cardarelli Hospital, Naples, Italy;
⁴Department of Molecular Medicine and Medical Biotechnology, "Federico II" University of Naples, Naples, Italy;
⁵Division of Gastroenterology and Hepatology, University Hospital Basel, Basel, Switzerland;
⁶Pathology Unit, Cardarelli Hospital, Naples, Italy

$Correspondence: Dr. Charlotte KY Ng: Institute of Pathology, University Hospital Basel, Schoenbeinstrasse 40, Basel, 4031, Switzerland. Tel: +41613286874; Fax: +41612653194. E-mail: kiuyancharlotte.ng@usb.ch; Dr. Salvatore Piscuoglio: Institute of Pathology, University Hospital Basel, Schoenbeinstrasse 40, Basel, 4031, Switzerland. Tel: +41613286874; Fax: +41612653194. E-mail: Salvatore.Piscuoglio@usb.ch

Word count (including references, table and figures): 3698; Figures: 2; Table: 1.
ABSTRACT

Background: Hepatocellular carcinomas (HCCs) are not routinely biopsied, resulting in a lack of tumor materials for molecular profiling. Here we sought to determine if plasma-derived cell-free DNA (cfDNA) captures the genetic alterations of HCC in patients who have not undergone systemic therapy.

Patients and methods: Frozen biopsies from the primary tumor and plasma were synchronously collected from 30 prospectively recruited, systemic treatment-naïve HCC patients. Deep sequencing of the DNA from the biopsies, plasma-derived cfDNA and matched germline was performed using a panel targeting 46 coding and non-coding genes frequently altered in HCCs.

Results: In 26/30 patients, at least one somatic mutation was detected in biopsy and/or cfDNA. Somatic mutations in HCC-associated genes were present in the cfDNA of 63% (19/30) patients and could be detected ‘de novo’ without prior knowledge of the mutations present in the biopsy in 27% (8/30) patients. Mutational load and the variant allele fraction of the mutations detected in the cfDNA positively correlated with tumor size and Edmondson grade. Crucially, among the seven patients in whom the largest tumor was ≥5cm or was associated with metastasis, at least one mutation was detected ‘de novo’ in the cfDNA of 86% (6/7) cases. In these patients, cfDNA and tumor DNA captured 87% (80/92) and 95% (87/92) of the mutations, suggesting that cfDNA and tumor DNA captured similar proportions of somatic mutations.

Conclusion: In patients with high disease burden, the use of cfDNA for genetic profiling when biopsy is unavailable may be feasible. Our results support further investigations into the clinical utility of cfDNA in a larger cohort of patients.
**Key message:** Mutations in HCC genes can be detected in the plasma DNA in 86% of therapy-naïve HCC patients with high disease burden using next-generation sequencing. In these patients, plasma DNA captured nearly all mutations found in the matched biopsy. Liquid biopsies may serve as a surrogate of the tumor for molecular profiling in patients with high disease burden and/or advanced grade.
**Introduction**

The invasive nature of biopsy has prompted investigations into the use of plasma-derived cell-free DNA (cfDNA) as a potential minimally invasive surrogate for molecular profiling in several cancer types [1-4]. In contrast to most solid tumor types, HCC diagnosis is frequently on the basis of radiology alone and in the absence of tumor biopsy. Therefore, nucleic acids for genetic profiling of HCC are typically obtained from tumor resection, a procedure that is only performed in patients with limited, early-stage disease. In unresectable HCC patients, should the need for molecular profiling arises, the tumor materials would have to be collected in a non-routine invasive procedure. The lack of routinely collected tumor materials is a hurdle for wider adoption of tumor profiling.

Studies have found that cfDNA concentration in serum or plasma of HCC patients is 3-4 times higher than in patients with chronic hepatitis and is up to 20 times higher than in healthy individuals [5-7]. Moreover, cfDNA concentration was found to be associated with tumor size, portal vein invasion and may be prognostic [5-8]. Molecular studies of circulating tumor DNA (ctDNA) in HCCs have investigated the size profiles of ctDNA [9], or were mutational studies of few cases, of resected materials, performed at very low depth or investigated few mutation hotspots [10-14]. The use of resected materials, however, restricts molecular analyses to patients with early-stage, resectable disease. Given the correlation of cfDNA concentration and tumor size, one may speculate that patients with later stage disease would have higher mutational burden in cfDNA, as has been shown in other cancer types [4, 15].

Restricting molecular studies of ctDNA to mutation hotspots risks missing a substantial number of mutations, as most somatic mutations in HCC, even those in HCC-associated driver genes, do not fall into mutation hotspots [16-21]. Besides **TP53** (p53), **CTNNB1** (β-catenin) and **TERT** promoter, a wide range of HCC-associated driver genes and recurrently mutated promoter regions have been discovered, including those involved in chromatin remodelling (e.g. **ARID1A**, **ARID1B**, **ARID2**, **BAP1**), Wnt/β-catenin pathway (e.g. **AXIN1**, **FGF19**), and
response to oxidative stress (e.g. *KEAP1, NFE2L2*) [16-21]. Additionally, long non-coding RNA genes (IncRNA, e.g. *NEAT1, MALAT1*) and promoter regions of *WDR74, TFPI2* and *MED16* are also recurrently mutated [18, 19, 22].

In this exploratory study, we sought to determine if somatic mutations in HCC driver genes can be detected with high confidence using next-generation sequencing in the plasma-derived cfDNA of HCC patients who have not undergone systemic therapy, and if the repertoire of mutations in the cfDNA is representative of the synchronously collected tumor biopsy. To address these questions, we prospectively recruited 30 HCC patients from whom we synchronously collected diagnostic core needle tumor biopsy and whole blood (*Supplementary Table S1*) and performed deep sequencing targeting HCC driver genes and mutation hotspots (*Supplementary Table S2*).

**Patients and Methods**

**Patients**

Thirty patients diagnosed with HCC at the University Hospital Basel, Basel, Switzerland or at Ospedale Cardarelli, Naples, Italy, were prospectively recruited for this study after written informed consent (*Supplementary Table S1*). Patients who had previous systemic therapy for HCC were excluded. One patient was treated with radio-frequency thermal ablation 21 months prior to sample collection. From each patient undergoing diagnostic liver biopsy, two ultrasound-guided core needle biopsies of the primary tumor and whole blood were collected at diagnosis at the same time. Of the two primary tumor biopsies, one was processed and embedded in paraffin for clinical purposes and the other one was snap-frozen and stored at −80°C for research purposes. 10mL of whole blood was collected in a 10mL Cell-Free DNA Blood Collection Tube (BCT, Streck) and processed immediately (*Supplementary Methods*). Plasma was stored at −80°C until cfDNA extraction.

Tumor size, tumor location, macrovascular invasion, multifocality, and extrahepatic spread of
each patient were assessed radiologically. Clinical staging of the patients was determined according to the Barcelona Clinic Liver Cancer (BCLC) staging system [23]. Sex of the patients, serum alpha-fetoprotein (AFP) levels, primary risk factors (hepatitis B/C virus infection, alcoholic liver disease, non-alcoholic fatty liver disease) were retrieved from clinical files. Histologic grading was performed according to the 4-point scale Edmondson and Steiner system [24] (Supplementary Methods). Approval for the use of these samples has been granted by the ethics committee (Protocol Number EKNZ 2014-099).

**Targeted sequencing and analysis**

Tumor and germline DNA was extracted from fresh frozen biopsies and peripheral blood leukocytes ('buffy coat'). Circulating cfDNA was extracted from 3-6mL of plasma (Supplementary Methods). DNA samples from the tumors, plasma-derived cfDNA and germline DNA were subjected to targeted sequencing using an Ampliseq panel targeting all exons of 33 liver cancer-associated protein-coding genes, all exons of the recurrently mutated IncRNA genes MALAT1 and NEAT1, recurrently mutated promoter region of TERT, WDR74, TFPI2 and MED16, as well as hotspots mutations in an additional seven cancer genes (Supplementary Table S2). Sequencing was performed on an Ion 530 chip using the Ion S5 XL system (Thermo Fisher Scientific, Supplementary Methods and Supplementary Table S3). Sequencing data have been deposited in the Sequence Read Archive under the accession SRP115181.

Sequence reads were aligned to the human reference genome hg19 using TMAP. Somatic mutations were defined using Torrent Variant Caller (TVC) v5.0.3. We filtered out mutations supported by ≤ 8 reads, and/or those covered by <10 reads in the tumor/cfDNA or <10 reads in the matched germline. We only retained mutations for which the tumor variant allele fraction (VAF) was at least 10 times that of the matched normal VAF to ensure we kept only the somatic variants (Supplementary Methods). Due to the repetitive nature and the high GC content of the TERT promoter region, TERT mutation hotspots (chr5:1295228 and
chr5:1295250) were additionally screened, and were considered present if supported by at least 5 reads or VAF of at least 5%. Mutations identified using the above steps are referred to as those found by ‘de novo’ methods.

To account for somatic mutations that may be present at low VAF in either the tumor biopsy or the matched cfDNA samples but not both, all somatic mutations identified using the ‘de novo’ methods in one of the two samples were interrogated for their presence in the matched sample by supplying TVC with their positions as the ‘hotspot list’. Mutations supported by at least 2 reads were considered to be present. Mutations identified using the above steps are referred to as those found by ‘interrogation’. Clinical actionability was assessed using OncoKB [25].

Statistical analysis

All statistical analyses were performed in R v3.3.1. Correlations between the number of mutations, cfDNA concentrations and continuous/ordinal clinical variables (Supplementary Methods) were assessed using the Spearman’s rho. Comparisons of continuous/ordinal clinical variables between patients with and without somatic mutations in the cfDNA were performed using Mann-Whitney U tests. Comparisons of categorical clinical variables and between patients with and without somatic mutations in the cfDNA were performed using Fisher’s exact tests. All statistical tests were two-tailed and \( P<0.05 \) was considered statistically significant. 95% confidence intervals (CIs) were estimated by leaving out 20% of the data points, computed over 100 runs.

Results

Of the 30 patients prospectively recruited into this study, 33% (10/30) had BCLC stages B/C/D disease (Table 1, Supplementary Table S1). Multifocal and metastatic diseases were seen in 11 and one patients, respectively. Median diameter of the largest tumor was 34mm (range 13mm-220mm). At least one primary risk factor was identified for all patients (except HPU025
for whom the information is unavailable. Cirrhosis was seen in 87% (26/30) cases.

From each patient undergoing diagnostic liver biopsy, a core needle biopsy and whole blood were collected at the same time for targeted sequencing. A median of 94.6ng (range 19.8ng-1710ng) of plasma cfDNA was obtained from 10mL of whole blood per patient (Supplementary Table S1). We performed deep sequencing of the HCC biopsies, cfDNA and matched germline using an in-house custom-made panel targeting 46 coding and IncRNA genes frequently altered in HCCs (median 1339x in biopsies and plasma, range 703x-9385x, Supplementary Tables S2-S3). To mimic the potential use of plasma-derived cfDNA in the absence of available resected tumor material or a core needle biopsy in a clinical setting, we defined the somatic mutations for each HCC and cfDNA samples independently without prior knowledge of the repertoire of mutations present in the biopsy/cfDNA counterpart following a stringent set of analysis criteria (or ‘de novo’). Additionally, to account for mutations that may be present at frequencies below the detection limit of the de novo approach, we further examined the sequencing data of the biopsies for all mutations detected in the cfDNA (or ‘by interrogation’), and vice versa. In 26/30 patients, at least one somatic mutation was detected in the biopsy and/or cfDNA (Figure 1 and Supplementary Table S4).

Using the ‘de novo’ approach, we detected at least one somatic mutation in the cfDNA of 27% (8/30) patients (median 3, blue/gold bars, Figure 1). Considering the 7 non-hypermutator cases with at least one detectable mutation in the cfDNA, 81% (17/21) of the mutations detected in the cfDNA were also independently detected in the biopsy counterparts. In the hypermutator case (HPU207), 97% (64/66) of the mutations detected in the cfDNA were also independently detected in the biopsy counterpart (blue bars, Figure 1). All 6 apparently cfDNA-specific mutations were found to be present at low frequencies in their biopsy counterparts by interrogation (gold bars, Figure 1), suggesting that, in accordance with a recent study [13], cfDNA may be useful in overcoming intra-tumor genetic heterogeneity within the biopsies in therapy-naïve HCC patients. On the other hand, of all mutations detected in
the non-hypermutator and the hypermutator cases, 78% (78/100) and 7% (5/71), respectively, were detected only in the HCC biopsies using the de novo approach (dark/light red bars, Figure 1). However, 31% (24/78) and 100% (5/5) of these mutations could in fact be detected in the cfDNA by interrogation (dark red bars, Figure 1). Taken together, these results demonstrate that at least one somatic mutation can be detected in the cfDNA without prior knowledge of the repertoire of mutations in the HCC biopsies in 27% (8/30) of HCC patients and that at least one mutation was present, including those identified by interrogation, in 63% (19/30) cases.

Comparing the clinicopathologic parameters, we found that the 8 cases for whom at least one somatic mutation was detected in the cfDNA using the ‘de novo’ approach were associated with larger tumors (diameter of the largest tumor) and increasing Edmondson grade (P=0.012 and P=0.010, Mann-Whitney U tests, Figure 1, Supplementary Table S5). Across all patients, the number of mutations detected ‘de novo’ in the cfDNA was positively correlated with the diameter of the largest tumor and Edmondson grade (r=0.482, P=0.007 and r=0.470, P=0.012, respectively, Spearman’s rho). The diameter of the largest tumor and Edmondson grade were also correlated with the maximum variant allele fractions of the mutations detected in the cfDNA (r=0.496, P=0.005 and r=0.502, P=0.007, respectively, Spearman’s rho) and cfDNA concentration (r=0.889, P<0.001 and r=0.439, P=0.020, respectively, Spearman’s rho, Supplementary Table S5). Additionally, at least one mutation was detected ‘de novo’ in the cfDNA in 40% (8/20) of male patients compared to 0% (0/10) of female patients, and in 75% (3/4) HCCs not associated with cirrhosis compared to 19% (5/26) of HCCs (P=0.029 and P=0.048, respectively, Fisher’s exact tests, Supplementary Table S5).

Among the seven cases in whom the largest tumor was ≥5cm or was associated with metastasis, at least one mutation was detected ‘de novo’ in the cfDNA of 86% (6/7) cases, with a median of 75% (range 0%-100%) of the mutations detected in the cfDNA (Figure 2). Importantly, 87% (80/92, 95% CI 84%-91%) mutations were detected ‘de novo’, and all but
two remaining mutations could be detected by interrogation in the cfDNA counterparts. Conversely, 95% (87/92, 95% CI 93%-97%) mutations were detected ‘de novo’ in the tumor biopsies, suggesting that mutation profiling of cfDNA in these patients captured similar proportion of mutations as tumor profiling would. By contrast, only 9% (7/78, 95% CI 6%-11%) mutations were detected in the cfDNA of the remaining 23 patients with small (largest tumor ≤5cm), non-metastatic HCC. These results suggest that in most HCC patients with high tumor burden, somatic mutations can be detected in the cfDNA with high confidence and that the repertoire of somatic mutations detected in cfDNA is representative of that in the primary HCC biopsy.

**Discussion**

HCC differs from most other tumor types in that biopsies are rarely performed as they are usually not required for diagnosis. Thus, in patients not eligible for tumor resection (i.e. patients with large or metastatic disease and/or with poor performance status), tumor materials are usually unavailable for molecular profiling. Here we describe a prospective study to investigate the utility of cfDNA collected at the time of biopsy for molecular profiling in HCC patients. Targeting the most significantly mutated genes and regions in HCCs, we found that, even without the prior knowledge of the somatic mutations in the HCCs, high-depth sequencing analysis of plasma-derived cfDNA revealed that at least one somatic mutation in HCC driver genes can be detected in 27% (8/30) of therapy-naïve HCC patients. In an additional 11 cases, cfDNA captured mutations present below ‘de novo’ detection limit in the biopsies, demonstrating that somatic mutations were present in the cfDNA of 63% (19/30) of HCC patients at diagnosis. Importantly, among the patients with high disease burden (large tumor or metastasis) and most likely to be ineligible for resection, cfDNA profiling captured nearly as many mutations as primary tumor biopsy profiling alone. Of note, a TSC2 frameshift mutation detected in the cfDNA and the primary tumor of the metastatic patient HPU209 is targetable by everolimus in cancers of the central nervous system as standard of care (Supplementary Table S4). Taken together, our results demonstrate that the repertoire of mutations in HCC-
associated genes identified in the cfDNA is representative of that in the biopsy.

Many HCC patients present with multifocal or metastatic disease and variable levels of heterogeneity with branched and parallel evolutionary patterns have been detected in HCC patients [13, 14]. Here we found a number of mutations that were detected with high confidence in the cfDNA but could only be detected by interrogation in the biopsy counterparts, reinforcing the notion that genetic analysis of a single diagnostic biopsy of the primary tumor may not be representative of the disease. Studies into the use of cfDNA as a minimally invasive surrogate for molecular profiling in HCC patients are therefore of particular clinical relevance.

Our study was limited in cohort size but as a proof of principle study and interpreted in the context of other tumor types [1-3], we found strong evidence that somatic mutations can be reliably detected in patients with high disease burden. As a prospective study, we have not assessed the prognostic significance of our findings. Furthermore, our filtering steps for the ‘de novo’ approach was deliberately stringent to closely recapitulate a potential clinical scenario. It is plausible that the limited sensitivity in detecting mutations ‘de novo’ in patients with low tumor burden is related to stringent filters. In fact, the number of mutations detected by interrogation suggests that advanced sequencing technologies incorporating molecular barcoding or alternative high-fidelity sequencing techniques will likely increase detection sensitivity in the clinical setting. Despite these limitations, the observed correlation of detectable somatic mutations and disease burden has important implications in the implementation of precision medicine [3] as it points towards the use of cfDNA for genetic profiling in HCC patients ineligible for resection and provides an argument for not subjecting the subset of patients with high disease burden to otherwise diagnostically unnecessary invasive procedure. Our results support further investigations into the clinical utility of cfDNA in a larger cohort of patients.
Funding

This work was supported by the Krebsliga beider Basel [KLbB-4183-03-2017 to C.K.Y.N.]. Additional financial support was provided by the Swiss Cancer League (Oncosuisse) [KLS-3639-02-2015 to L.M.T. and KFS-3995-08-2016 to S.P.]; the Swiss National Science Foundation [Ambizione PZ00P3_168165 to S.P.]; the European Research Council [ERC Synergy Grant 609883 to C.K.Y.N. and M.H.H].

Disclosure

The authors have declared no conflicts of interest.
References


Figure 2

Largest tumor < 5cm

Metastatic HCC

Largest tumor ≥ 5cm

Variant allele frequency

- 0% (°, < 100 reads)
- >0%–1%
- >1%–5%
- >5%–20%
- >20%–40%
- >40%–60%
- >60%–80%
- >80%–100%
- ‘de novo’

Mutation effect

- Hotspot
- Nonsense/frameshift
- Missense/in-frame
- Synonymous
- Non-coding/promoter
FIGURE LEGENDS

Figure 1: Number of somatic mutations detected in plasma-derived cell-free DNA and clinicopathologic information in 30 patients with therapy-naïve hepatocellular carcinoma. The number of somatic mutations were categorized based on whether they were detected ‘de novo’ or ‘by interrogation’ (i.e. without or with prior knowledge of the repertoire of mutations in the biopsy/cfDNA counterpart, respectively, see color key). Clinicopathologic information is color-coded. White indicates unavailable information.

Figure 2: Somatic mutations found in cfDNA and in their primary tumor biopsies. Heatmaps indicate the variant allele fractions of the somatic mutations (blue, see color key) or their absence (grey) in the 26 pairs of tumor biopsy and cfDNA for which at least one somatic mutation was identified. Mutation types are indicated as colored dots. Orange boxes denote the mutations detected using the ‘de novo’ approach (i.e. without prior knowledge of the mutations in the biopsy/cfDNA counterpart). Mutations not detected by the ‘de novo’ approach but were covered by <100 reads are indicated by an asterisk. Cases are grouped according to the diameter of the largest tumor. T: tumor. PL: plasma.