Diagnostic Targeted Sequencing Panel for Hepatocellular Carcinoma Genomic Screening

Viola Paradiso,* Andrea Garofoli,* Nadia Tosti,* Manuela Lanzafame,* Valeria Perrina,* Luca Quagliata,* Matthias S. Matter,* Stefan Wieland,† Markus H. Heim,†† Salvatore Piscuoglio,* Charlotte K.Y. Ng,* and Luigi M. Terracciano*

From the Institute of Pathology,* University Hospital Basel, Basel; the Department of Biomedicine,† University of Basel, Basel; and the Department of Gastroenterology and Hepatology,‡ University Hospital Basel, Basel, Switzerland

CME Accreditation Statement: This activity (“JMD 2018 CME Program in Molecular Diagnostics”) has been planned and implemented in accordance with the accreditation requirements and policies of the Accreditation Council for Continuing Medical Education (ACCME) through the joint providership of the American Society for Clinical Pathology (ASCP) and the American Society for Investigative Pathology (ASIP). ASCP is accredited by the ACCME to provide continuing medical education for physicians. The ASCP designates this journal-based CME activity (“JMD 2018 CME Program in Molecular Diagnostics”) for a maximum of 18.0 AMA PRA Category 1 Credit(s)™. Physicians should claim only credit commensurate with the extent of their participation in the activity.

CME Disclosures: The authors of this article and the planning committee members and staff have no relevant financial relationships with commercial interests to disclose.

Accepted for publication July 2, 2018.

Address correspondence to Luigi M. Terracciano, M.D., Institute of Pathology, University Hospital Basel, Schonbeinstrasse 40, 4031 Basel, Switzerland. E-mail: luigi.terracciano@usb.ch.

Sequencing technologies have allowed the discovery of genetic alterations essential in the diagnosis and treatment of human cancer or approval of new targeted therapies.1 In addition, the presence of subclonal mutations has direct implications in the development of drug resistance.2,3 In the era of precision medicine, the development of rapid, accurate, high-throughput, and cost-effective genomic assays to accommodate the increasingly genotype-based therapeutic approaches is required.4,5 Currently, the costs of whole-genome and whole-exome sequencing (WES) are still prohibitive in the clinical setting, especially for small institutions. Furthermore, although DNA from fresh-frozen samples can be directly sequenced, the prohibitive costs might prevent its use in routine clinical screening.

Commercially available targeted panels miss genomic regions frequently altered in hepatocellular carcinoma (HCC). We sought to design and benchmark a sequencing assay for genomic screening of HCC. We designed an AmpliSeq custom panel targeting all exons of 33 protein-coding and two long non-coding RNA genes frequently mutated in HCC, TERT promoter, and nine genes with frequent copy number alterations. By using this panel, the profiling of DNA from fresh-frozen (n = 10, 1495×) and/or formalin-fixed, paraffin-embedded (FFPE) tumors with low-input DNA (n = 36, 530×) from 39 HCCs identified at least one somatic mutation in 90% of the cases. Median of 2.5 (range, 0 to 74) and 3 (range, 0 to 76) mutations were identified in fresh-frozen and FFPE tumors, respectively. Benchmarking against the mutations identified from Illumina whole-exome sequencing (WES) of the corresponding fresh-frozen tumors (105×), 98% (61 of 62) and 100% (104 of 104) of the mutations from WES were detected in the 10 fresh-frozen tumors and the 36 FFPE tumors, respectively, using the HCC panel. In addition, 18 and 70 somatic mutations in coding and noncoding genes, respectively, not found by WES were identified by using our HCC panel. Copy number alterations between WES and our HCC panel showed an overall concordance of 86%. In conclusion, we established a cost-effective assay for the detection of genomic alterations in HCC. (J Mol Diagn 2018, 20: 836–848; https://doi.org/10.1016/j.jmoldx.2018.07.003)

Supported in part by the Swiss Cancer League (Oncosuisse) grants KLS-3639-02-2015 (L.M.T.) and KFS-3995-08-2016 (S.P.), Krebsliga beider Basel project KLbB-4183-03-2017 (C.K.Y.N.), Swiss National Science Foundation Ambizione grant PZ00P3_168165 (S.P.), the Swiss Centre for Applied Human Toxicology (SCAHT; V.Pa.), and the European Research Council ERC Synergy grant 609883 (C.K.Y.N. and M.H.H.). V.Pa. and A.G. contributed equally to this work. C.K.Y.N. and L.M.T. contributed equally to this work as senior authors. Disclosures: None declared.

Funding bodies had no role in the design of the study, collection, analysis, and interpretation of the data or the writing of the manuscript.

Copyright © 2018 American Society for Investigative Pathology and the Association for Molecular Pathology. Published by Elsevier Inc. All rights reserved.

https://doi.org/10.1016/j.jmoldx.2018.07.003
tissue is ideal for genomic screening, it is not part of routine diagnostic practice at most hospitals and institutions. Instead, DNA from formalin-fixed paraffin-embedded (FFPE) material is frequently the only option. Moreover, DNA from small tumors, after reserving materials for histopathologic analyses, may be extremely limited. For research institutes, being able to exploit and revisit archival materials associated with long-term follow-up but whose DNA may potentially be degraded is also highly desirable. Given these limitations, PCR-based sequencing panels may be more broadly applicable than capture-based solutions.

Existing commercial sequencing panels, such as the amplicon-based Ion Torrent Oncomine Comprehensive Assay version 3 (Thermo Fisher Scientific, Waltham, MA) and the capture-based Foundation Medicine FoundationOne assay, are broadly applicable to common cancer types. Compared with other common cancer types, however, hepatocellular carcinoma (HCC) has a distinct mutational profile. Although HCC driver genes TP53 and CTNNB1 are also frequently mutated in cancers such as those of the lungs, the breasts, and colon,6 genes such as APOB, ALB, HNF1A, and HNF4A are significantly mutated only in HCC.7–17 The distinct mutational landscape of HCC is likely a result of the unique biology of hepatocyte differentiation and liver functions. Of note, the frequently altered APOB, ALB, and HNF4A are not targeted by most commercial assays. In the noncoding regions, recent commercially available panels include TERT promoter mutation hotspot (c.-124C>T). However, long noncoding RNA (lncRNA) genes frequently mutated in HCC, such as MALAT1 and NEAT1,16 have yet to be included in commercial panels or in exome capture panels. Recent whole-genome studies have also uncovered mutation clusters in promoter regions of genes such as MED16, WDR74, and TFPI216,18 that are not covered in commercial panels.

In this study, we designed a high-throughput and cost-effective amplicon-based sequencing panel specifically to screen for somatic mutations and copy number alterations (CNAs) in HCC. Our panel includes genes and regions frequently altered in HCC, including those not currently covered by commercial panels. We tested the sequencing panel by using fresh-frozen and FFPE materials with low-input DNA to evaluate the feasibility of this panel in routine diagnostics.

Materials and Methods

Targeted Panel Design and Generation

A custom targeted sequencing panel that focused on the most frequently altered genes in HCC7–18 was designed by using Ion Ampliseq Designer (Thermo Fisher Scientific). The panel (hereafter the HCC panel) covers all exons of 33 protein-coding genes; recurrently mutated lncRNA genes MALAT1 and NEAT1; and the recurrently mutated promoter regions of TERT, WDR74, MED16, and TFPI2 (Figure 1A and Supplemental Table S1).7–18 Nine genes frequently altered by CNAs and mutation hotspots in seven cancer genes are also covered (Figure 1A and Supplemental Table S1).7–18 The HCC panel was designed by using the FFPE option for smaller amplicon size. The nine genes for CNA profiling were designed to be covered by at least 10 non-overlapping amplicons evenly distributed across the length of the genes. The designed panel was further inspected by
the white glove service (Thermo Fisher Scientific) for primer specificity in a multiplex PCR reaction. The HCC panel consists of 2120 amplicons split into two primer pools and covers genomic regions of approximately 203 kb.

Tissue Samples

Human tissues were obtained from patients undergoing diagnostic liver biopsy at the University Hospital Basel, Basel, Switzerland. Written informed consent was obtained from all included patients. Ultrasound-guided needle biopsies were obtained from tumor lesion(s) and adjacent nontumoral liver tissue (Figure 1B). The study was approved by the ethics committee of the northwestern part of Switzerland (protocol EKNZ 2014-099). For all patients except cases 2, 6, 7, and 9, a single tumor biopsy was included (Supplemental Table S2). For cases 6 and 7, two tumor biopsies were included, and for cases 2 and 9, three tumor biopsies were included. A portion of each biopsy was FFPE for clinical purposes, and the remaining portion of each biopsy was snap-frozen and stored at −80°C for research purposes. For this study, 45 fresh-frozen tumor biopsies and 39 fresh-frozen nontumor biopsies from 39 patients were included. FFPE tissue samples that remained after diagnostic routine (36 tumor biopsies and 31 nontumor biopsies from 36 patients) were included. Pathologic assessment of tumor content was performed by two expert hepatopathologists (M.S.M. and L.M.T.) with the use of diagnostic hematoxylin and eosin slides.

DNA Extraction

DNA from fresh-frozen biopsies was extracted by using the ZR-Duet DNA/RNA MiniPrep Plus kit (Zymo Research, Irvine, CA) according to the manufacturer’s instructions. Before extraction, tissue samples were crushed in liquid nitrogen to facilitate lysis. For DNA extraction from FFPE samples, one 5-μm-thick slide was cut directly in the tube, and DNA was extracted with the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. DNA was quantified with the use of diagnostic hematoxylin and eosin slides.

Library Preparation and Deep Sequencing Using the HCC Panel

Library preparation for the HCC panel was performed by using the Ion AmpliSeq library kit version 2.0 (Thermo Fisher Scientific) according to the manufacturer’s guidelines. For cases 2, 6, 7, and 9, DNA extracted from multiple fresh-frozen tumor biopsies was pooled equimolar before library preparation (Supplemental Table S2). In total, 20 fresh-frozen samples (10 tumor samples and 10 nontumoral counterparts) and 67 FFPE samples (36 tumor biopsies and 31 nontumoral counterparts) were sequenced by using the HCC panel.

The HCC panel consists of two pools of amplification primers. Ten nanograms of DNA per sample was used for library preparation for each pool. Amplification was performed according to the manufacturer’s guidelines. The amplicons from the two pools were combined and treated to digest the primers and to phosphorylate the amplicons. The amplicons were then ligated to Ion Adapters (Thermo Fisher Scientific) by using DNA ligase. Finally, cleaning and purification of the generated libraries were performed with Agencourt AMPure XP (Beckman Coulter, Brea, CA) according to the manufacturer’s guidelines. Quantification and quality control were performed with the Ion Library TaqMan Quantitation Kit (Thermo Fisher Scientific). Samples were diluted to reach the concentration of 40 pmol and then were pooled for sequencing. Twenty-five μL of the pooled libraries was loaded on Ion 530 Chip (Thermo Fisher Scientific) and processed in Ion Chef Instrument (Thermo Fisher Scientific). Sequencing was performed on Ion S5 XL system (Thermo Fisher Scientific).

Sequence Data Analysis for the HCC Panel

Sequence reads were aligned to the human reference genome hg19 by using TMAP within the Torrent Suite Software version 5.4 (Thermo Fisher Scientific; https://github.com/ontorrent/TS) for the Ion S5XL system. Coverage analysis was performed by using Picard’s CollectTargetedPerMetrics tool version 2.4.1 (http://broadinstitute.github.io/picard) (Supplemental Table S3). Uniformity of sequencing was defined as the proportion of target bases covered at >20% of mean amplicon coverage for a given sample. Comparison of the coverage for the two primer pools was performed by using paired Wilcoxon test.

Somatic mutations were identified with Torrent Variant Caller version 5.0.3 (Thermo Fisher Scientific; https://github.com/ontorrent/TS). For fresh-frozen samples, the corresponding fresh-frozen nontumoral samples were used as the germline control. For FFPE samples, FFPE nontumoral samples were used as the matched germline sample when available. When FFPE nontumoral samples were not available, the corresponding fresh-frozen nontumoral samples were used as germline control. Mutations at hotspot residues were whitelisted.21,22 Mutations supported by <8 reads, and/or those covered by <10 reads in the tumor or <10 reads in the matched nontumoral counterpart were filtered out. Only those for which the tumor variant allele fraction (VAF) was >10 times that of the matched nontumoral VAF were retained to ensure the somatic nature of the variants. Because of the repetitive nature and the high GC content of the TERT promoter region, TERT mutation hotspots (chr5:1295228 and chr5:1295250) were additionally screened. TERT promoter mutations were considered present if supported by at least five reads or VAF of at least 5%. All mutations were manually inspected by using the Integrative Genomics Viewer version 2.3.69 (https://software.broadinstitute.org/software/igv).

CNAs were defined as follows. For each sample, end-to-end sequence reads were extracted separately for the two
amplicon pools. A copy number reference for each pool was generated by using all nontumoral samples to estimate overall read depth, log2 ratio, and variability by using the reference function from CNVkit version 0.9.0 (https://github.com/etal/cnvkit). Amplicons with <100 read depth, absolute log2 ratio >1.5, or spread >1 were removed from copy number analysis. Protein-coding genes for which the complete coding region was included in the panel or for which amplicons were specifically designed for copy number analysis were included. Samples with excessive residual copy number log2 ratio (segment interquartile range >0.8) were excluded, as previously described.

For each tumor/nontumor pair, log2 ratio was computed for each amplicon, separately for the two amplicon pools by using Varscan2 version 2.4.3 (https://github.com/dkoboldt/varscan). Log2 ratios for the two pools were separately centered then merged for segmentation by using circular binary segmentation. CNAs were determined, adopting a previously described approach. In brief, SD of the log2 ratios of the 40% of the central positions ordered by their log2 ratios was computed. Copy number gains and amplifications/high gains were defined as +2 SDs and +6 SDs, respectively. Copy number losses and deep deletions were defined as −2.5 SDs and −7 SDs, respectively. All gene amplifications and deep deletions were visually inspected by using log2 ratio plots.

To evaluate the impact of tumor purity on CNA analysis, an in silico simulation was performed on 12 cases (six frozen and six FFPE, selected on the basis of the presence of gene amplification/high gain or deep deletion), by replacing tumor reads with reads sampled from the normal samples to simulate tumor content 5%, 10%, 20% up to the actual tumor content for the samples. CNA analysis was performed as described above.

**WES**

WES was performed for DNA extracted from the 45 tumor biopsies and 39 nontumoral counterparts from the 39...
patients (Supplemental Table S2). Whole-exome capture was performed by using the SureSelectXT Clinical Research Exome (Agilent, Santa Clara, CA) platform according to the manufacturer’s guidelines. Sequencing (2 × 101 bp) was performed at the Genomics Facility of ETH Zurich Department of Biosystems Science and Engineering (Basel, Switzerland) by using Illumina HiSeq 2500 (Illumina, San Diego, CA) according to the manufacturer’s guidelines. Sequence reads were aligned to the reference human genome GRCh37 by using Burrows-Wheeler Aligner-MEM version 0.7.12 (http://bio-bwa.sourceforge.net).28 Local realignment, duplicate removal, and base quality adjustment were performed by using the Genome Analysis Toolkit version 3.6 (https://software.broadinstitute.org/gatk)29 and Picard version 2.4.1 (http://broadinstitute.github.io/picard).

For WES samples, sequence reads overlapping with the target regions of the HCC panel were extracted for further comparative analyses. Sequencing statistics were evaluated for the overlap of the target regions of the WES and the HCC panel. For cases 2, 6, 7, and 9, for which DNA from multiple fresh-frozen tumor biopsies was pooled before sequencing by using the HCC panel, WES reads from the multiple biopsies were merged to facilitate downstream comparisons. For all four cases, the number of reads obtained from WES of individual biopsies was comparable (Supplemental Table S3).

Somatic single nucleotide variants and small insertions and deletions (indels) were detected by using MuTect version 1.1.4 (https://software.broadinstitute.org/cancer/cga/mutect)30 and Strelka version 1.0.15 (https://github.com/Illumina/strelka),31 respectively. Single nucleotide variants and small indels outside of the target regions, those with VAF of <1%, and/or those supported by <3 reads were filtered out. Only variants for which the tumor VAF was >5 times that of the matched nontumoral VAF were retained. Further, variants identified in at least two of a panel of 123 nontumoral liver tissue samples, using the artifact detection mode of MuTect2 implemented in Genome Analysis Toolkit version 3.6 were excluded,32 where the panel of 123 nontumoral liver tissue samples included the 39 nontumoral samples in the present study and were captured and sequenced with the same protocols. All indels were manually inspected by using the Integrative Genomics Viewer.33 Copy number analysis was performed with FACETS version 0.5.13 (https://github.com/mskcc/facets),32 and genes targeted by amplifications or deep deletions were defined by using the same thresholds as above.

Pairwise Comparisons between Mutations Identified by WES, Fresh-Frozen and FFPE Tissues

Pairwise comparisons of the somatic mutations identified by WES and by the HCC panel were performed, according to the originating biopsies (Supplemental Table S2). Discordant variants were reevaluated and interrogated for their presence by supplying Torrent Variant Caller version 5.0.3 with their positions as the hotspot list (for Ion Torrent sequencing) or by Genome Analysis Toolkit version 3.6 Unified Genotyper by using the GENOTYPE_GIVE-N_ALLELES mode (for WES).

Sanger Sequencing

To validate the discordant variants, Sanger sequencing was performed on both DNA from the fresh-frozen and the corresponding FFPE tumor biopsies. PCR amplification of 5 ng of genomic DNA was performed with the AmpliTaq 360 Master Mix Kit (Thermo Fisher Scientific) on a Veriti Thermal Cycler (Thermo Fisher Scientific) as previously described (Supplemental Table S4).30 PCR fragments were purified with ExoSAP-IT (Thermo Fisher Scientific). Sequencing reactions were performed on a 3500 Series Genetic Analyzer instrument by using the ABI BigDye Terminator chemistry version 3.1 (Thermo Fisher Scientific) according to the manufacturer’s instructions. All analyses were performed in duplicate. Sequences of the forward and reverse strands were analyzed with MacVector software version 15.1.3 (MacVector, Inc., Apex, NC).20

Analysis of TCGA Data

To determine the frequencies of high-level copy number gains/focal amplifications and deep deletions/focal homozygous deletions in HCC, the GISTIC 2.0 copy number calls for The Cancer Genome Atlas (TCGA) HCC cohort from the cBioPortal were obtained.33 High-level gains and deep deletions were defined as those with GISTIC copy number state 2 and −2, respectively. Focal amplifications and focal homozygous deletions were defined as high-level gains and deep deletions that affected <25% of a given chromosome arm. For the 37 genes included in the copy number analysis, the frequencies of high-level gains/deep deletions and of focal amplifications/focal homozygous deletions were computed.

Statistical Analysis

Correlation analyses were performed with Pearson’s r and r². Statistical analyses were performed in R version 3.4.2 (The R Foundation, Vienna, Austria).

Results

HCC-Specific Custom Targeted Sequencing Panel Design and Quality Assessment

An HCC sequencing panel was designed to specifically target genes and genomic regions frequently altered in HCC (Figure 1A and Supplemental Table S1). The HCC panel consisted of complete coding regions of 33 genes involved in several pathways implicated in HCC pathogenesis, including the WNT pathway (CTNNB1, AXIN1), chromatin remodeling (ARID1A, ARID2, and BAP1), cell cycle regulation (CDKN1A, CDKN2A, CDKN2B, CCND1, RPS6KA3, RB1, and TP53),
inflammatory response (IL6R, IL6ST), and hepatocyte differentiation (ALB, APOB, HNF1A, and HNF4A). In addition, the HCC panel also targeted recurrently mutated lncRNA genes MALAT1 and NEAT1 and recurrently mutated promoter regions of TERT, WDR74, MED16, and TFFP2. Genes frequently altered by CNAs (eg, CCNE1, VEGFA, TERT) and mutation hotspots in BRAF, EEF1A1, HRAS, IL6ST, KRAS, NRAS, and PIK3CA were also targeted. To enable the efficient profiling of DNA samples derived from potentially degraded FFPE materials, the panel was designed by using the FFPE option for smaller amplicon size, with a mean amplicon size of 118 bp (range, 63 to 252 bp) (Figure 2A). The HCC panel was tested on the DNA extracted from 20 fresh-frozen samples (10 from tumor biopsies and 10 from nontumoral counterparts) and 67 FFPE samples (36 from tumor biopsies and 31 from nontumoral counterparts) obtained from 39 patients (Figure 1B and Supplemental Table S2).

A coverage analysis of the HCC panel was performed with the 10 fresh-frozen and 31 FFPE nontumoral DNA samples. In the fresh-frozen and FFPE nontumoral DNA samples, a mean coverage of 1478× (range, 925× to 2420×) and 580× (range, 263× to 1300×), respectively, were achieved (Figure 2B and Supplemental Table S3). No difference was found between the depth of coverage of the two pools of amplicons (P = 0.9879, paired Wilcoxon test) (Supplemental Figure S1A). At least 96.8% and 91.1% of the amplicons were covered at >30× and at least 98.7% and 95.6% of the amplicons were covered at >10× in the fresh-frozen and FFPE nontumor samples, respectively (Figure 2C and Supplemental Figure S1B). Median uniformity (defined as the proportion of target bases covered at >20% of the mean amplicon coverage of a given sample) was 89.9% (range, 86.8% to 91.5%) in the fresh-frozen samples and 89.0% (range, 73.3% to 92.3%) in the FFPE samples (Figure 2D). As expected, depth of sequencing of the amplicons was associated with GC content, with reduced depth at extreme GC content (Figure 2E).

HCC Panel Captures Somatic Mutations Concordant with WES and Identifies Additional Mutations

Next, the somatic mutations identified in the 10 fresh-frozen tumor/nontumoral pairs sequenced with the HCC panel were evaluated. A median sequencing depth of 1495× (range, 1026× to 1855×) in the tumor samples was achieved (Figure 2B and Supplemental Table S3). A median of 2.5 somatic mutations (range, 0 to 74 somatic mutations) were identified, including a median of 2 mutations (range, 0 to 52 mutation) in protein-coding genes (Figure 3A and Supplemental Table S4). No somatic mutations were identified for 2 of 10 cases (cases 3 and 12), although both cases had ≥50% tumor cell content (Supplemental Table S2). One case (case 9) exhibited a hypermutator phenotype with 74 somatic mutations identified.

To evaluate the somatic mutations defined with the HCC panel, the somatic mutations derived from WES, generated on the orthogonal Illumina technology, of the same DNA aliquots from the fresh-frozen tumors and matched nontumor samples were used as a benchmark (Figure 1B). By considering only the coding regions covered by the HCC panel, the median depths of WES were 114× (range, 92× to 345×) and 51× (range, 45× to 84×) in the fresh-frozen tumors and matched nontumor samples, respectively (Supplemental Table S3). WES analysis confirmed that no mutations were present within the targeted protein-coding regions in cases 3 and 12 and that case 9 was hypermutated (Figure 3B). Of the 62 mutations in the coding region identified from WES analysis, 61 (98%) were also called by the HCC panel analysis (Figure 3B). One NRAS Q61K hotspot mutation (case 6) was missed by using the HCC panel analysis. Manual review of this position revealed that the mutation had VAF of 2.5% by WES and 2.0% by the HCC panel (Supplemental Figure S2 and Supplemental Table S4). Note, however, that 2% is close to the detection limit of the current sequencing technologies.

Compared with the WES analysis, the HCC panel analysis revealed an additional six mutations in the coding regions, including five in case 9 and one in case 11 (Figure 3B). Manual review of the WES data showed that all six mutations were in fact supported by at least one read in WES, but those positions were covered at reduced depth, with 4 of 6 covered by ≤40 reads (including three in LRPIB) and 5 of 6 ≤80 reads (Supplemental Figure S2C and Supplemental Table S4). This suggested that the increased sensitivity in the HCC panel analysis was likely due to the increased depth achieved.

Additional to the mutations in the protein-coding regions, the HCC panel also targeted the lncRNA genes MALAT1 and NEAT1 and the promoter regions of TERT, WDR74, MED16, and TFFP2 (Figure 1A). Within these noncoding regions, an additional 32 mutations were identified across the 10 cases, representing a 48% gain of information compared with sequencing the protein-coding genes alone (Figure 3B). TERT promoter mutations were found in 60% (6 of 10) of cases and 16 somatic mutations in the lncRNA gene NEAT1 were identified in 40% (4 of 10) of cases (Figure 3B and Supplemental Table S4).

Taken together, for the protein-coding genes frequently mutated in HCC, the HCC panel analysis produced highly reliable results compared with WES. Given the increased sequencing depth achieved by using the HCC panel, somatic mutations that were missed by WES were identified. Of importance, the HCC panel analysis enabled us to identify somatic mutations in promoter regions and frequently mutated lncRNA genes.

HCC Panel Analysis Identifies Somatic Mutations in FFPE Diagnostic Biopsies with Low-Input DNA

Nucleic acids from diagnostic specimens are frequently derived from small FFPE samples. Therefore, it would be important to determine whether the HCC panel could also be used for somatic mutational screening on low-input DNA.
The Journal of Molecular Diagnostics  ■  jmd.amjpathol.org  ■  2023  ■  Vol. 25, No. 5  ■  843

(20 ng) extracted from FFPE samples. The DNA extracted from 36 diagnostic FFPE tumor biopsies was subjected to HCC panel sequencing to a median depth of 530× (range, 192× to 1257×) (Figures 1A and 2, B and C, and Supplemental Table S3). The median tumor content for these 36 cases was 90% (range, 5% to 100%) (Supplemental Table S2), thus representative of the distribution of tumor content in diagnostic samples in clinical practice. A median of three mutations (range, 0 to 76 mutations) per sample, including a median of two mutations (range, 0 to 53 mutations) in the coding regions was identified (Figure 4, Supplemental Figure S3, and Supplemental Table S4). No somatic mutations were identified for 8% (3 of 36) of cases (cases 7, 12, and 37), indicating that at least one somatic mutation could be detected in 92% of HCC diagnostic samples. Of note, although somatic mutations in the one biopsy with 5% tumor content could not be detected, somatic alterations in samples with 30% to 40% tumor content were detected.

The mutations identified in protein-coding genes from these 36 FFPE diagnostic biopsies were compared with those identified by WES of the DNA from the corresponding fresh-frozen biopsies. All 104 mutations identified from WES analysis were also called based on the HCC panel analysis (Figure 4, Supplemental Figure S3, and Supplemental Table S2), thus representative of the distribution of tumor content in diagnostic samples in clinical practice. A median of three mutations (range, 0 to 76 mutations) per sample, including a median of two mutations (range, 0 to 53 mutations) in the coding regions was identified (Figure 4, Supplemental Figure S3, and Supplemental Table S4). No somatic mutations were identified for 8% (3 of 36) of cases (cases 7, 12, and 37), indicating that at least one somatic mutation could be detected in 92% of HCC diagnostic samples. Of note, although somatic mutations in the one biopsy with 5% tumor content could not be detected, somatic alterations in samples with 30% to 40% tumor content were detected.

The mutations identified in protein-coding genes from these 36 FFPE diagnostic biopsies were compared with those identified by WES of the DNA from the corresponding fresh-frozen biopsies. All 104 mutations identified from WES analysis were also called based on the HCC panel analysis (Figure 4, Supplemental Figure S3, and Supplemental Table S2), with 21 of 36 cases (58%) harboring CTNNB1 mutations, a higher proportion than the TCGA and other HCC cohorts that was likely due to the higher percentage of alcohol-associated HCC (Supplemental Tables S1 and S2).15 In addition, analysis of the HCC panel identified 18 mutations in the coding regions that were not found in the WES analysis in 11 cases. Of these 18 mutations, 13 were evident in WES but were not identified as mutations in the WES analysis, predominantly because of low sequencing depth (Supplemental Figures S2D and S3). The remaining five mutations were verified to be present in the corresponding FFPE samples but absent in the fresh-frozen samples by Sanger sequencing (Supplemental Figure S4 and Supplemental Table S4), indicating that they were genuine discordances between the fresh-frozen and FFPE DNA and not false positive calls from the HCC panel assay. Of note, two of five mutations validated to be absent from the fresh-frozen DNA affected mutation hotspots in CTNNB1 (D32N and S45A) (Figure 4 and Supplemental Figure S4). The increased number of detected mutations by the HCC panel analysis was likely due to a combination of intratumor heterogeneity and the higher sequencing depth achieved.

Considering the 36 FFPE diagnostic biopsies, the HCC panel identified 70 somatic mutations in lncRNA genes and promoter regions, including 22 TERT promoter mutations (Figure 4 and Supplemental Table S4). Somatic mutations in lncRNA genes and promoter regions accounted for 37% of the total number of somatic mutations identified in the FFPE samples.

Compared with the high correlation of VAF between the sequencing platforms used in the fresh-frozen samples (r = 0.89, r² = 0.79, Pearson correlation), the correlation between WES from fresh-frozen samples and HCC panel by using FFPE samples was more modest (r = 0.67, r² = 0.45, Pearson correlation) (Supplemental Figure S2, A and B). Mutations with large deviations in VAFs between the sequencing platforms used in the fresh-frozen samples tended to be covered at reduced depths on either platform (Supplemental Figure S2C). Similar observations could be made between VAFs of exome (fresh-frozen) and HCC panel (FFPE) (Supplemental Figure S2D). The deviations in the latter may be more noticeable by the overall lower depth achieved in the FFPE samples than in the HCC panel sequencing of the fresh-frozen samples. Intratumor heterogeneity between the fresh-frozen and FFPE aliquots likely contributed to the reduced correlation.

Taken together these results suggested that the HCC panel analysis has high specificity and sensitivity in somatic mutation detection. Furthermore, somatic mutations in promoter regions (TERT promoter) and lncRNA genes (MALAT1 and NEAT1) highly mutated in HCC could also be detected.

Copy Number Analysis of the HCC Panel Reveals High Concordance with WES

To determine whether the HCC panel could also be used to detect CNAs, 42 genes whose coding regions were entirely covered or were tiled across the lengths of the genes for CNA detection were evaluated (Figure 1A and Supplemental Table S1). Using the 41 nontumoral samples, the variability of the depth of coverage in the amplicons targeting the 42 genes was assessed (Materials and Methods). After removing amplicons with low depth of coverage or high variability, 1483 amplicons were used for CNA profiling. To assess the ability to detect per-gene CNA, each nontumoral sample was further paired with two other randomly selected, sex-matched nontumoral samples. The copy number log₂ ratio of five genes, namely LRP1B, ALB, BRD7, ACVR2A, and IRF2, was variable (SD > 0.3); therefore, these genes were excluded from further CNA analyses. Thirty-seven genes were included in the CNA analysis.

Figure 3  Comparison of somatic mutations defined by whole-exome sequencing (WES) and hepatocellular carcinoma (HCC) panel in fresh-frozen tissues. A: Number of coding and noncoding mutations per case identified in 10 fresh-frozen biopsies by using the HCC panel. B: Comparison of somatic coding and noncoding mutations found by WES and the HCC panel in the fresh-frozen samples. Heatmaps indicate the variant allele fractions of the somatic mutations (blue, see color key) or their absence (gray) in the eight cases in which at least one somatic mutation was identified. Mutation types are indicated as colored dots according to the color key. Mutations that were not called by mutation caller but were supported by at least one sequencing read are indicated by asterisks.
The copy number profiles of matched fresh-frozen tumor/nontumor pairs and those derived from WES were compared. Of the 10 fresh-frozen pairs sequenced by using the HCC panel, one was excluded for excessive residual copy number log2 ratio (segment interquartile range, >0.8). For the nine evaluable samples, a correlation of \( r = 0.80 \) (\( r^2 = 0.64 \)) was found between the copy number log2 ratio of the two platforms (Figure 5A). When the copy number profiles of the 34 evaluable FFPE tumors were compared with the matched profiles from WES, a correlation of \( r = 0.73 \) (\( r^2 = 0.54 \)) was observed between the copy number log2 ratios (Figure 5A). Overall, 86% of the evaluable genes had concordant copy number states (Figure 5B).

It has previously been reported that tumor purity had an impact on the ability to make CNA calls. The impact of tumor purity on CNA analysis was therefore evaluated by using an in silico simulation on 12 cases (six fresh-frozen and six FFPE, selected on the basis of the presence of gene amplification/high gain or deep deletion), by replacing tumor reads with reads sampled from the normal samples to simulate tumor content 5%, 10%, 20% up to the actual tumor content for the samples. It was observed that amplifications/high gains were readily detected at 5% tumor content in many cases and at 20% in all cases (Supplemental Figure S5). In this cohort, deep deletions could not be detected at tumor content <40%.

Taken together, these results demonstrated that, despite profiling only a small number of genes, the HCC panel was able to detect CNAs in genes frequently gained or lost in HCC in both fresh-frozen and FFPE tumor samples with low-input DNA.

**Discussion**

HCC has a distinct mutational landscape compared with the major tumor entities. Numerous genes have been found to be mutated frequently in HCC but rarely in other tumors, such as those important for hepatocyte differentiation (\( ALB \), \( APOB \), \( HNF1A \), \( HNF4A \)) and inflammatory response (\( IL6R \), \( IL6ST \)). Given the relative rarity of HCC, these genes are currently not targeted or are only partially targeted in commercial panels [eg, Oncomine Comprehensive Panel version 3 (Thermo Fisher Scientific)] and in panels used by sequencing services [eg, FoundationOne assay (Foundation Medicine, Cambridge, MA)] (Supplemental Table S1). Thus, the currently available commercial assays for genomic profiling have suboptimal utility for HCC, and a targeted sequencing panel specifically designed for HCC is warranted.

In this study, we designed a custom Ion Torrent AmpliSeq sequencing panel, targeting all exons of 33 protein-coding genes, two lncRNA genes, promoter regions of four...
genes previously found to be recurrently mutated in HCC, nine genes frequently affected by CNAs, and mutation hotspots in seven cancer genes.\(^7\) Of importance, a number of the genes targeted by using the HCC panel are not currently on these two commercial panels. Of the 39 cases profiled with the HCC panel (including both fresh-frozen and FFPE samples), at least one somatic mutation was detected in 90% (35 of 39) of the cases. Of the mutations in coding genes found using this panel, 22% (42 of 189) would have been missed by both Oncomine Comprehensive Panel version 3 and the FoundationOne assay. In addition, recent whole-genome studies of HCC have revealed frequent mutations in IncRNA genes \(\text{NEAT1}\) and \(\text{MALAT1}\), both of which are not currently targeted by commercial panels. In fact, it was found that approximately one-third of the mutations on the HCC panel were within the promoter and IncRNA regions.

Mutation screening and copy number profiling results from the HCC panel were benchmarked against those obtained from WES by the orthogonal Illumina sequencing technology. All but one mutation identified from WES were detected by using the HCC panel. An additional 10% to 15% of mutations within the coding regions were identified. Most of these additional mutations were in fact supported by few reads by WES; thus, the increased sensitivity was likely a direct result of the increased sequencing depth of both the tumor and the matched normal samples achieved. Crucially, however, evidence of intratumor genetic heterogeneity between adjacent fresh-frozen and FFPE biopsies, including two \(\text{CTNNB1}\) mutations, was found, suggesting that in these cases the \(\text{CTNNB1}\) mutations were not trunk mutations.

Although CNA detection using capture-based methods has been successful for targeted sequencing panel of several hundred genes,\(^{35}\) CNA detection using amplicon-based targeted sequencing has proven more difficult. A recent study investigated the use of an amplicon-based sequencing strategy that targeted all exons of 113 genes related to DNA repair.\(^{25}\) The researchers demonstrated that, with an appropriate analysis strategy and quality control, amplicon-based sequencing strategy is feasible and cost-effective for CNA profiling in FFPE samples.\(^{25}\) In the present study, the strategy of computing and centering the log\(_2\) ratios for the absolute difference in copy number states proved to be an effective strategy in resolving issues associated with variable amplification efficiencies, with 86% of the genes showing concordant copy number states. Considering few studies have investigated the use of small targeted sequencing panel for CNA profiling, further benchmarking studies comparing analysis strategies and including larger sample size will likely improve the accuracies.

In the clinical setting, the quality, type, and amount of input materials for genomic profiling are crucial considerations, particularly in light of the smaller tumors being detected in screening programs. Here, we demonstrated that the HCC panel could be used for genomic screening with high sensitivity and specificity with low-input DNA (20 ng) derived from FFPE samples without compromising the results. Although based on an analysis of the TCGA HCC cases, 92% and 85% of the cases would have exhibited at least one nonsynonymous mutation by using the FoundationOne and the Oncomine assays, respectively, the HCC panel holds the advantage of much lower input requirement.
than that required for commercial panels (eg, >40-μm tissue samples for the FoundationOne assay) and for capture-based targeted sequencing strategies. In addition, somatic genetic alterations (somatic mutations and amplifications) could be detected from tumor samples with as low as 30% tumor content. Considering that mutations in the one sample with 5% tumor content could not be detected, 30% may be the lower limit of successful genomic profiling. Although lower limits (approximately 20%) have also been reported, samples were not available to verify this. The samples included in this study are de facto samples obtained from routine diagnostic practice, and it was demonstrated that the low-input DNA requirement facilitates genomic profiling from small biopsies.

Driver genetic alterations have not yet become a tangible tool in clinical decision making for the treatment of HCC; thus, the immediate clinical application of our panel may be limited. However, recent studies have described the association of TERT promoter and CTNNB1 exon 3 mutations with increased risk of malignant transformation of hepatocellular adenomas, more frequent HNF1A and IL6ST mutations in hepatocellular adenomas than HCCs, as well as TP53 mutation as a poor prognostic indicator in HCC. These associations suggest a potential utility of genomic profiling in prognostication for hepatocellular adenomas and HCCs, in tissues or even in cell-free DNA. In terms of potential targetable alterations, three somatic mutations identified in our cohort of HCC are molecular targets in other cancer types according to OncoKB. These include ATM loss of function mutation using olaparib in prostate cancer (level 4; biological evidence), NRAS hotspot mutation with binimetinib or in combination with ribociclib in melanoma (level 3; clinical evidence), and TSC2 mutation with everolimus in central nervous system cancer (level 2; standard of care). Application of our panel in clinical decision may become feasible in the future.

This study has several limitations. First, the targeted nature of the HCC panel means that copy number profiling is not genome-wide and is restricted to the genes included on the panel. Clinically, focal amplifications, compared with gains of chromosome arm, are more likely to be true driver genetic event and may be considered drug targets. The targeted nature of the HCC panel makes it difficult to distinguish the two scenarios. However, a re-analysis of the TCGA data suggests that high-level gains of chr1q13.3 (encompassing CCND1, FGF19, FGF3, FGF4) are almost always focal amplifications (>93%), whereas 50% to 70% of high-level gains of TERT and VEGFA are focal amplifications (Supplemental Table S5). By contrast, high-level gains of chr1q (SETDB1 and IL6R) and chr8q (NCOA2, MYC, and PTK2) are frequently nonfocal (<10%), consistent with the frequent high-level gain of entire arms of chr1q and chr8q. For deletions, most deep deletions are focal deletions, including all deletions (100%) in ARID2, AXIN1, CDKN2A/B, PTEN, and TSC1/2. These results suggest that CNAs affecting some of the most promising drug targets on the HCC panel are frequently true focal CNAs. Second, given that a median of two to three mutations per tumor were identified, tumor mutational burden, a putative biomarker for response to immune therapy, may not be accurately defined. Third, the HCC panel does not include unique molecular identifiers, which would be useful to assess library complexity, particularly for samples with low-input DNA. We envisage that the addition of unique molecular identifiers would be particularly beneficial for the study of cell-free DNA from HCC patients. Fourth, we designed the panel specific for HCC. Recent studies have revealed that mixed HCC/cholangiocarcinoma and cholangiocarcinoma have recurrent mutations in genes such as IDH1/2, whereas FRK mutations decrease in frequency from hepatocellular adenoma to HCC. These genes are not covered by the HCC panel. However, as an amplicon-based sequencing panel, adding amplicons to include genes that may assist in the differential diagnosis of HCC is straightforward.

Conclusion

This study demonstrated that the HCC panel is a cost-effective strategy for mutation screening and copy number profiling for routine diagnostic HCC samples with low-input DNA.

Acknowledgments

S.P., C.K.Y.N., and L.M.T. conceived and supervised the study; L.Q., M.S.M., S.P., C.K.Y.N., and L.M.T. performed literature search and designed the sequencing panel; S.W. and M.H.H. provided the samples and the whole-exome sequencing data; V.Pa., N.T., M.L., V.Pe., and S.P. performed DNA extraction and sequencing and prepared the pipeline for mutation calling; V.Pa., A.G., S.P., C.K.Y.N., and L.M.T. analyzed the results and wrote the manuscript.

Supplemental Data

Supplemental material for this article can be found at https://doi.org/10.1016/j.jmoldx.2018.07.003.

References


8. Leiserson MD, Greenman CD, Pritchard CC, Wendl MC,一等奖


