Detection of Copy Number Alterations by Shallow Whole-Genome Sequencing of Formalin-Fixed, Paraffin-Embedded Tumor Tissue

Malaïka Van der Linden, MSc; Lennart Raman, MSc; Ansel Vander Trappen, BSc; Annelies Dheedene, PhD; Matthias De Smet, BSc; Tom Sante, PhD; David Creytens, MD, PhD; Yolande Lievens, MD, PhD; Björn Menten, PhD; Jo Van Dorpe, MD, PhD; Nadine Van Roy, PhD

Context.—In routine clinical practice, tumor tissue is stored in formalin-fixed, paraffin-embedded blocks. However, the use of formalin-fixed, paraffin-embedded tissue for genome analysis is challenged by poorer DNA quality and quantity. Although several studies have reported genome-wide massive parallel sequencing applied on formalin-fixed, paraffin-embedded samples for mutation analysis, copy number analysis is not yet commonly performed.

Objective.—To evaluate the use of formalin-fixed, paraffin-embedded tissue for copy number alteration detection using shallow whole-genome sequencing, more generally referred to as copy number variation sequencing.

Design.—We selected samples from 21 patients, covering a range of different tumor entities. The performance of copy number detection was compared across 3 setups: array comparative genomic hybridization in combination with fresh material; copy number variation sequencing on fresh material; and copy number variation sequencing on formalin-fixed, paraffin-embedded material.

Results.—Very similar copy number profiles between paired samples were obtained. Although formalin-fixed, paraffin-embedded profiles often displayed more noise, detected copy numbers seemed equally reliable if the tumor fraction was at least 20%.

Conclusions.—Copy number variation sequencing of formalin-fixed, paraffin-embedded material represents a trustworthy method. It is very likely that copy number variation sequencing of routinely obtained biopsy material will become important for individual patient care and research. Moreover, the basic technology needed for copy number variation sequencing is present in most molecular diagnostics laboratories.

Using aCGH, important discordances were revealed, with more CNAs detected in FFPE-derived DNA. This discordance was not seen by Hostetter et al., but FFPE-derived profiles showed more background signal. In the latter study, aCGH was performed on 6 paired FF/FFPE tissue samples. In order to counter the input quality requirements of aCGH, the OncoScan CNV Assay (Thermo Fisher, Waltham, Massachusetts) was designed to work with low-quality DNA from FFPE material and uses the molecular inversion probe technology, originally developed for single-nucleotide polymorphism genotyping. In brief, a molecular inversion probe panel binds regions of interest and gap filling leads to amplification. Using a tag sequence, unique for a region of interest, array hybridization is visualized and compared with reference intensity data. However, using arrays, sample throughput is relatively low whereas costs are rather high.

A more recent technique for CNA detection is shallow whole-genome sequencing, more generally referred to as copy number variation sequencing (CNVseq). This approach requires a genome coverage of only 0.1x to 1x and is based on counting the number of aligned reads within chromosomal windows or bins. As CNVseq makes use of short-read sequencing, it opens up the possibility of analyzing tissue samples that have been collected during routine clinical practice and hence have been fixed with formalin. However, CNVseq of FFPE-derived DNA is underexplored: only a few studies have reported on its feasibility. However, CNVseq of FFPE-derived DNA was underexplored: only a few studies have reported on its feasibility. Analysis of 356 FFPE tissue samples by Chin et al showed good-quality data. Moreover, others showed that CNAs were comparable between FF and FFPE tissue samples. However, these studies comprised only one to 7 paired cases.

In this study, we evaluated CNVseq for CNA detection performed on 21 paired fresh and FFPE tumor tissue DNA samples, covering a range of different tumor entities. Moreover, aCGH profiles from most fresh tumor samples were available for comparison.

**MATERIALS AND METHODS**

**Retrospective Cohort**

Patients qualified for inclusion if both FFPE (nondecalcified; fixation time between 6 and 48 hours) and fresh material were available for sequencing. Notwithstanding that selection was random, we aimed at combining a variety of disease entities, including tumor types routinely tested by clinical laboratories. In total, 21 patients were included, whereof 17 had available aCGH (60K array, Agilent Technologies, Santa Clara, California) profiles (generated in accordance with Kumps and colleagues) (Supplemental Tables 1 and 2; see supplemental digital content at www.archivesofpathology.org in the August 2020 table of contents, containing 3 tables and 18 figures). Nine more nontumoral FFPE archives (ureter, tonsil, skin, stomach, lung, and brain tissue).

**FPFE DNA Extraction**

Although DNA from fresh material was directly available for sequencing through routine clinical practice, 7 sections were taken from the paraffin blocks to initiate FFPE DNA extraction. The first and last section (1.5–2 μm) were stained with hematoxylin-eosin and served as references to guide dissection and to determine tumor cell percentage, whereas the others sections (10 μm) were dissected and deparaffinized by deparaffinization solution (Qiagen, Hilden, Germany). Extraction was executed using the QIAamp DNA FFPE tissue kit (Qiagen), according to the manufacturer’s instructions, followed by quantification using a Qubit fluorometer and the Qubit dsDNA BR assay kit (Life Technologies, Bleiswijk, The Netherlands).

**Fragment Analysis and Shearing**

Size distribution of the FFPE-derived DNA was determined using the Fragment Analyzer, the High Sensitivity Small DNA Fragment Analysis Kit, and PROSize data analysis software, according to the manufacturer’s instructions (Agilent).

DNA shearing to 200–base pair (bp) fragments was completed by Covaris’ adaptive focused acoustics shearing process and an M220 focused ultrasonicator (Covaris, Woburn, Massachusetts). Different shearing treatment times were evaluated. Using 200 ng/50 μL DNA as input, the following treatment parameters were set: peak incident power = 50 W; duty factor = 20%; cycles per burst = 200; treatment time = 160 seconds for fresh DNA, 100 seconds for FFPE-derived DNA; and temperature = 20°C. The fragments’ size distributions were evaluated using the Fragment Analyzer, the dsDNA 910 reagent kit, and PROSize data analysis software (Agilent).

**Library Construction and Sequencing**

DNA libraries were prepared using the NEXTflex rapid DNA-seq kit (following option 2, with bead-based size selection prior to polymerase chain reaction amplification) and NEXTflex DNA bar codes (PerkinElmer, Austin, Texas), starting from 128 ng of DNA. Libraries were examined using the Fragment Analyzer, the dsDNA 910 reagent kit (fresh DNA) or the High Sensitivity Small DNA Fragment Analysis Kit (FFPE DNA), and PROSize data analysis software, and by quantitative PCR using the KAPA library quantification kit (Roche, Pleasanton, California).

Libraries were finally equimolarly pooled (4–6 nM). Cluster generation was completed using a cBot 2 system (Illumina, San Diego, California). Sequencing was performed by Illumina’s HiSeq 3000 (Illumina) in a single-read 50-cycle run mode. The minimal number of reads per sample was set at 15 million (mean coverage of 0.25x).

**Copy Number Profiling**

Demultiplexing was executed by the bc2fastq software (v2.19.1, Illumina). To allow fair comparison, paired fresh and FFPE sample data were read normalized by randomly removing reads from the highest coverage sample until equal read depths were obtained. Raw reads were mapped by Bowtie 2 (v2.2.8) onto human reference genome GRCh37, using the fast-local flag, where the resulting BAM files were indexed by SAMtools (v1.5). Reads with a mapping quality lower than 40 were removed. To deduce normalized binwise log2 ratios between the observed and expected copy number, the reference-free QDNAseq (v1.18.0) was favored. A bin size of 50 kilobases (kb) was selected, resembling the resolution of the aCGH 60K arrays (Agilent). Profile segmentation was executed by circular binary segmentation, implemented in the R package DNAcopy (v1.56.0). At some loci, copy number is undeterminable by short-read shallow whole-genome sequencing (eg, at centromeres): QDNAseq masks these regions by applying a blacklist. Therefore, a unified blacklist, defined as the union of all obtained blacklists, was applied to all resulting profiles (including aCGH), enabling unbiased concordance analyses.

**FFPE Lower Tumor Fraction Detection Limit Estimation**

Assuming our negative control DNA had no structural aberrations, the variance across these flat FFPE-derived profiles illustrated the minimal required tumor fraction (TF) to significantly call aberrations. These profiles were sequentially divided into segments of length 100 kb (2 bins) until 10 Mb (200 bins), where their log2 ratio equaled the interpolated average of its bins’. For each segmental locus, the standard deviation (sd) and mean (x) of the...
Different DNA Quality Outcomes Between Fresh and FFPE Tumor Tissue

Extraction of FFPE DNA yielded on average 4.56 ± 0.71 µg DNA (Supplemental Table 2). Fragment analysis of this DNA showed a wide size distribution with both highly degraded and less-degraded fragments (Supplemental Figures 1 through 3). Four FFPE samples needed no further fragmentation (cases 5, 12, 14, and 17). After DNA shearing, fragmented FFPE-derived DNA showed an overall wider size distribution when compared with fragmented fresh tumor tissue DNA (Supplemental Figures 4 through 7). The DNA quality number, a quality metric provided by the PROSize data analysis software, gives an indication of the fraction of fragments with sizes above a specified threshold, set at 200 bp. The average DNA quality number was 5.9 ± 0.1 (Supplemental Table 3) for fresh DNA and 6.4 ± 0.1 (Supplemental Table 3) for FFPE-derived DNA, supporting different size distributions (paired t test; P < .001).

Library construction resulted in on average 39.90 ± 4.79–nM libraries, starting from fresh samples (Supplemental Table 1), and 9.18 ± 2.32–nM libraries starting from FFPE samples (Supplemental Table 2). Library validation showed different traces for FFPE-derived libraries compared with fresh DNA libraries. Fresh samples showed a peak around the expected library size of 300 bp, but had a right-skewed distribution. Formalin-fixed, paraffin-embedded samples displayed 2 peaks, one around 300 bp and one at higher molecular weight (Supplemental Figures 8 through 12).

**RESULTS**

**Different DNA Quality Outcomes Between Fresh and FFPE Tumor Tissue**

The TF was estimated by ichorCNA (Broad Institute, Cambridge, Massachusetts), a software that predicts the portion of tumor fragments based on observed structural aberrations. Although ichorCNA is benchmarked on cell-free DNA, the authors ensured its applicability to FF and FFPE material (Viktor Adalsteinsson, PhD, written communication, December 21, 2017). The tool generates elementary copy number profiles (bin size 500 kb), which were normalized using the provided guanine and cytosine content and mappability tracks. Five different starting values, representing the nontumoral contamination, were passed to the optimization algorithm: 10%, 25%, 50%, 75%, and 90%. The obtained tumor content corresponding to the highest likelihood was interpreted as the final TF.

Aberrations were called once the log2 ratio of profile segments surpassed a gain or loss cutoff valued as log2 \(\frac{\text{tumor fraction}}{\text{lib ratio}}\). If the TF did not exceed the previously established lower detection limit, no copy number states were called.

**Statistical Comparative Analyses**

All comparative analyses were executed by custom scripting in R (v3.5.1). The median segment variance (MSV), a measure that reflects the overall vertical spread of log2 ratios within segments, was adopted to perform profile noise comparison. Profile correlation was visualized by hierarchical clustering, using the unweighted pair group method with arithmetic mean as an agglomerative clustering procedure and the Pearson correlation as a concordance measure. The latter was applied to a predefined set of 1000 equally sized windows, obtained by merging segment-interpolated binwise log2 ratios—this to reduce computational complexity and to increase robustness.

Finally, performance was established in the form of a receiver operating characteristic analysis. The CNVseq/FFPE combination was compared with the CNVseq/fresh and aCGH/fresh setups, using aCGH/fresh and CNVseq/fresh as gold standards, respectively. Patients without aCGH test or with insufficient TF (according to the lower detection limit) were excluded. Specificity and sensitivity were measured while sliding the aberration cutoff from liberal (ratio of 1) to conservative (0.5 for deletions; 1.5 for gains). The mean area under the curve defined the overall performance.
CNVseq of FFPE Material Represents a Trustworthy Method

In this study, we aimed at detecting CNAs of at least 1 Mb. Assuming a coverage of 0.25x (the negative control samples had a mean coverage of 0.25x ± 0.01x, where coverage impacts noise) and 100 segments (the case samples had on average 96.13 segments, where the number of segments impacts Bonferroni correction), the minimal required TF to detect CNAs of at least 1 Mb was measured to be 20% (Figure 4, a). Three cases (cases 4, 7, and 9) had a TF below the lower detection limit (Supplemental Tables 1 and 2). This naturally explains why matched samples do not cluster for cases 7 and 9 (Figure 3).

Although the variance in the profiles of FFPE material was often higher, this had little impact on the performance of aberration calling (Figure 4, b and c). Receiver operating characteristic analysis using aCGH as a gold standard showed a mean area under the curve of 0.960 for CNVseq of fresh material and 0.949 for CNVseq of FFPE material (Figure 4, b). For one case (case 17), CNVseq showed poor concordance with aCGH. The aCGH profile (Supplemental Figure 17, q) displayed an abundance of, probably, false-positive bins.

Using CNVseq in combination with fresh material as the gold standard, the mean area under the curve for FFPE material amounted to 0.973 (Figure 4, c). In one case (case 12), a 6.2-Mb gain on the short arm of chromosome 2 (amp(2)(2p16.1p14)) could be detected in both fresh and FFPE-derived DNA samples, using aCGH or/ and CNVseq (Supplemental Figure 15, l), illustrating that small CNAs could be detected using FFPE tumor tissue. In this region, the REL gene is located; amplification of this gene is a recurrent event in B-cell lymphomas.

DISCUSSION

In this study, we demonstrate the feasibility of CNA detection by CNVseq using FFPE-derived DNA. In total, 21 patients were selected during a recent time frame (2016 for cases with aCGH data)—this to minimize the effects of fixation and storage time. Nondecalcified FFPE material was chosen as decalcification introduces additional DNA damage. We observed that the DNA integrity of FFPE tissue was less affected than expected. Like Hostetter et al7 and Munchel et al,12 but different from Mc Sherry et al,1 we saw a wide size distribution of FFPE-derived DNA, indicating only partial DNA degradation. Higher-molecular-weight DNA was more present than expected.

Very similar CNA profiles were obtained across paired fresh and FFPE samples. Differences possibly were caused by sampling errors or resulted from tumor heterogeneity, as previously cited by Little et al.23 Although we observed less degradation, poorer DNA quality from FFPE samples was obvious. The higher variance seen in our FFPE-derived copy number profiles is partly caused by the lower-quality FFPE-derived DNA or the higher error load in FFPE-derived sequencing data. This was reflected mostly during the mapping phase of sequencing. In contrast to Schweiger et al,10 who saw variations in read mappability independent of the DNA source, we observed that FFPE samples had a lower percentage of mapped reads; thus, the lower percentage of mapped reads was probably due to DNA quality.

Using nontumoral FFPE samples, we determined that the minimum TF necessary for accurate aberrations calling on FFPE material is 20%. This is similar to the sensitivity of aCGH, debated to be 20% as well,7 whereas with targeted sequencing, where coverage is regionally increased, a lower detection limit is possible.24

Besides detection of large chromosomal alterations, the determination of single-gene copy number status is also important in cancer diagnosis, prognosis, and treatment. We showed that detection of even small CNAs, such as the gain at 2p (containing the REL gene) that we could readily detect in our study, is possible using CNVseq on FFPE material. Classical techniques such as conventional karyotyping are limited to a resolution of 5 to 10 Mb because of chromosome banding.25
We showed that the higher variance seen in FFPE-derived profiles has little impact on the performance of statistically calling an aberration if TF is at least 20%.

Although several studies have reported massive parallel sequencing applied on FFPE samples, especially in the context of mutation detection, CNA detection on FFPE material is not yet commonly executed. Although our study is limited to recently archived samples, it shows the possibility of performing CNA detection on routinely collected FFPE material. The use of FFPE material for DNA analysis has
Figure 3. Clustering fresh (*) and formalin-fixed, paraffin-embedded (FFPE) (**) samples resulted in a paired pattern. Rows contain samples, whereas columns correspond to equally sized genomic windows. The dendrogram (left) represents clusters based on the Pearson distance between the samples, which clustered according to expectation, with the exception of cases 7 and 9 (Supplemental Figures 14, g, and 15, i), resembling flat profiles.

Figure 4. The minimum tumor fraction (TF) necessary for reliable copy number variation sequencing (CNVseq) of formalin-fixed, paraffin-embedded (FFPE) material was determined to be 20%. a, Assuming a coverage of 0.25x and 100 segments, FFPE-derived DNA from negative controls enabled the calculation of a TF lower detection limit in function of the aberration size. Dots indicate measures for discrete sizes, whereas lines represent LOESS fits. In this study, we aimed at capturing aberrations of at least 1 Mb, requiring tumor tissue with a purity of at least ~20%. b, Receiver operating characteristic analysis performed using array comparative genomic hybridization (aCGH) as a gold standard. Samples with insufficient TF were excluded (<20%; cases 4, 7, and 9). Thin dotted lines represent separate samples, whereas thick lines indicate the mean relation. The area under the curve (AUC) quantifies the latter. c, Similar to b, yet the gold standard is given by CNVseq in combination with fresh material.
several important advantages. Formalin fixation and paraffin embedding is the standard procedure for examination of surgical biopsies, as it provides optimal morphology. Handling and storage of FFPE material is much easier and cheaper than that of FF or unfixed material. In addition, selection of tissue blocks or dissection of areas with a maximal tumor cell percentage makes the process of tumor tissue selection for DNA analysis much more controlled. Compared with aCGH, CNVseq can be easily implemented as routine clinical practice, because a very similar workflow, as for noninvasive prenatal testing, is often already implemented in numerous genetic laboratories. Moreover, pooling enables CNVseq for multiple types of analyses at once (e.g., in combination with noninvasive prenatal testing), resulting in overall turnaround times of less than 10 days. In addition, using low-coverage sequencing, this test can be performed at a low cost relative to aCGH. When aiming at detecting smaller insertions or deletions, the sequencing depth can be easily increased, but will result in an increase of the price tag.

In contrast to targeted sequencing, shallow whole-genome sequencing cannot be used for single-nucleotide variant detection, as it is low coverage and does not need predefined targets. We believe that CNVseq for CNA detection is likely to become response prediction for a wide range of different solid tumor entities. Copy number variation sequencing can be used for detection of losses or gains at single loci, like HER2 amplification in breast cancer, but is especially interesting for tumor types with complex CNA profiles. For instance, bladder carcinoma and melanoma harbor tumor-specific CNAs. The current standard diagnostic procedure to detect these CNAs is via fluorescent in situ hybridization (FISH) using different probes. However, this technique is labor-intensive and highly dependent on the quality of the samples. The advantages of CNVseq over FISH are that it is less observer dependent and it is cheaper when multiple probes are needed. Even more important, CNVseq is genome wide, and thus can be used without prior knowledge of the target region and can be applied on every tumor type. A time-consuming FISH test that is currently offered by many molecular pathology centers detects for 1p/19q codeletion in gliomas. Codeletion of 1p and 19q is offered by many molecular pathology centers detects for 1p/tumor type. A time-consuming FISH test that is currently known of the target region and can be applied on every samples. The advantages of CNVseq over FISH are that it is lower cost relative to aCGH. When aiming at detecting smaller insertions or deletions, the sequencing depth can be easily increased, but will result in an increase of the price tag.

In contrast to targeted sequencing, shallow whole-genome sequencing cannot be used for single-nucleotide variant detection, as it is low coverage and does not need predefined targets. We believe that CNVseq for CNA detection is likely to become response prediction for a wide range of different solid tumor entities. Copy number variation sequencing can be used for detection of losses or gains at single loci, like HER2 amplification in breast cancer, but is especially interesting for tumor types with complex CNA profiles. For instance, bladder carcinoma and melanoma harbor tumor-specific CNAs. The current standard diagnostic procedure to detect these CNAs is via fluorescent in situ hybridization (FISH) using different probes. However, this technique is labor-intensive and highly dependent on the quality of the samples. The advantages of CNVseq over FISH are that it is less observer dependent and it is cheaper when multiple probes are needed. Even more important, CNVseq is genome wide, and thus can be used without prior knowledge of the target region and can be applied on every tumor type. A time-consuming FISH test that is currently offered by many molecular pathology centers detects for 1p/19q codeletion in gliomas. Codeletion of 1p and 19q is characteristic for oligodendroglioma, a glioma subtype associated with better prognosis and response to chemotherapy. Testing with FISH for oligodendroglioma requires evaluation of 4 signals in many tumor cells (at least 50–100). Moreover, interpretation of this test is difficult, as exemplified by the many different cutoffs and varying guidelines that are used by different laboratories and in clinical studies. Given the advantages of CNVseq, FISH may, in time, be replaced by CNVseq on FFPE tumor biopsies. Copy number variation sequencing on FFPE material has also great potential for the diagnosis of high-grade serous ovarian carcinoma (a tumor with an exceptionally high CNA burden), the evaluation of BRCA1/2 in breast carcinoma, the detection of CNAs in neuroblastoma, and the study of CNAs in many other benign and malignant tumor entities.

CONCLUSIONS

Formalin-fixed, paraffin-embedded tissue enables harvesting of tumor material in a more controlled manner. The higher variance seen in FFPE-derived copy number profiles has little impact on the performance of statistically calling an aberration if the TF exceeds the established lower detection limit of 20%. Importantly, archival FFPE tissue is available for many tumor entities in laboratories for surgical pathology and thus represents a large biobank for future clinical research.

We are thankful for the technical assistance of Lynn Supply, Tamara De Clercq, and the Fragment Analyzer and CNVseq team.

References


