

# Journal of Visualized Experiments

## Detecting Somatic Genetic Alterations in Tumor Specimens by Exon Capture and Massively Parallel Sequencing --Manuscript Draft--

Manuscript Number:	JoVE50710R1
Article Type:	Invited Methods Article - JoVE Produced Video
Corresponding Author:	Michael F Berger Memorial Sloan-Kettering Cancer Center New York, NY UNITED STATES
First Author:	Helen H Won
Order of Authors:	Helen H Won
	Sasinya N Scott
	A. Rose Brannon
	Ronak Shah
	Michael F Berger



**Michael Berger, PhD**  
Assistant Attending, Department of Pathology

March 22, 2013

Larissa Jarzylo  
Science Editor  
Journal of Visualized Experiments  
17 Sellers St.  
Cambridge MA 02139

Dear Larissa,

We are pleased to submit a revised version of our manuscript entitled: "Detecting Somatic Genetic Alterations in Tumor Specimens by Exon Capture and Massively Parallel Sequencing" to JoVE. This manuscript describes our method to use hybridization capture and next-generation sequencing to profile tumor samples for clinically relevant mutations in several hundred cancer-associated genes.

We have carefully considered all editorial and reviewer comments and have substantially improved our manuscript. The revised manuscript and a point-by-point response to reviewers are included in this submission.

Our submission also includes 4 figures and 3 tables. We have also provided our signed article and video license agreement for open access.

All authors participated in the methodology presented in this article. We look forward to hearing from you, and we thank you again for considering our manuscript.

Sincerely,

A handwritten signature in black ink, appearing to read "Mich 2 B", with a long horizontal flourish extending to the right.

Michael F. Berger, Ph.D.  
Assistant Attending, Department of Pathology  
Associate Member, Human Oncology & Pathogenesis Program  
Memorial Sloan-Kettering Cancer Center

# **Detecting Somatic Genetic Alterations in Tumor Specimens by Exon Capture and Massively Parallel Sequencing**

## **Authors:**

Helen H. Won, Sasinya N. Scott, A. Rose Brannon, Ronak Shah, Michael F. Berger

## **Authors: institutions(s)/affiliation(s) for each author:**

Helen H. Won

Department of Pathology

Memorial Sloan-Kettering Cancer Center

[wonh@mskcc.org](mailto:wonh@mskcc.org)

Sasinya N. Scott

Department of Pathology

Memorial Sloan-Kettering Cancer Center

[scottsl@mskcc.org](mailto:scottsl@mskcc.org)

A. Rose Brannon

Department of Pathology

Memorial Sloan-Kettering Cancer Center

[brannona@mskcc.org](mailto:brannona@mskcc.org)

Ronak Shah

Department of Pathology

Memorial Sloan-Kettering Cancer Center

[shahr2@mskcc.org](mailto:shahr2@mskcc.org)

Michael F. Berger

Department of Pathology & Human Oncology and Pathogenesis Program

Memorial Sloan-Kettering Cancer Center

[bergerm1@mskcc.org](mailto:bergerm1@mskcc.org)

## **Corresponding author:**

Michael F. Berger

## **Keywords:**

Massively parallel sequencing, targeted exon sequencing, hybridization capture, cancer, FFPE, DNA mutations

## **Short Abstract:**

We describe the preparation of barcoded DNA libraries and subsequent hybridization-based exon capture for detection of key cancer-associated mutations in clinical tumor specimens by massively parallel “next generation” sequencing. Targeted exon sequencing offers the benefits of high throughput, low cost, and deep sequence coverage, thus yielding high sensitivity for detecting low frequency mutations.

**Abstract:**

Efforts to detect and investigate key oncogenic mutations have proven valuable to facilitate the appropriate treatment for cancer patients. The establishment of high-throughput, massively parallel “next-generation” sequencing has aided the discovery of many such mutations. To enhance the clinical and translational utility of this technology, platforms must be high-throughput, cost-effective, and compatible with formalin-fixed paraffin embedded (FFPE) tissue samples that may yield small amounts of degraded or damaged DNA. Here, we describe the preparation of barcoded and multiplexed DNA libraries followed by hybridization-based capture of targeted exons for the detection of cancer-associated mutations in fresh frozen and FFPE tumors by massively parallel sequencing. This method enables the identification of sequence mutations, copy number alterations, and select structural rearrangements involving all targeted genes. Targeted exon sequencing offers the benefits of high throughput, low cost, and deep sequence coverage, thus conferring high sensitivity for detecting low frequency mutations.

**Introduction:**

The identification of “driver” tumor genetic events in key oncogenes and tumor suppressor genes plays an essential role in the diagnosis and treatment of many cancers [1]. Large-scale research efforts utilizing massively parallel “next-generation” sequencing have enabled the identification of many such cancer-associated genes in recent years [2]. However, these sequencing platforms typically require large quantities of DNA isolated from fresh frozen tissues, thus posing a major limitation in characterizing and analyzing DNA mutations from preserved tissues, such as formalin-fixed paraffin embedded (FFPE) tumor samples. Improved efforts to efficiently and reliably characterize “actionable” genomic information from FFPE tumor samples will enable the retrospective analysis of previously-banked specimens and further encourage individualized approaches to cancer management.

Traditionally, molecular diagnostic laboratories have relied on time-consuming, low-throughput methodologies such as Sanger sequencing and real-time PCR for DNA mutation profiling. More recently, higher-throughput methods utilizing multiplexed PCR or mass spectrometric genotyping have been developed to investigate recurrent somatic mutations in key cancer genes [3-5]. These approaches, however, are limited in that only pre-designated “hotspot” mutations are assayed, making them unsuitable for detecting inactivating mutations in tumor suppressor genes. Massively parallel sequencing offers several advantages over these strategies including the ability to interrogate entire exons for both common and rare mutations, the ability to reveal additional classes of genomic alterations such as copy number gains and losses, and greater detection sensitivity in heterogeneous samples [6, 7]. Whole genome sequencing represents the most comprehensive approach for mutation discovery, though it is relatively expensive and incurs large computational demands for data analysis and storage.

For clinical applications, where only a small fraction of the genome may be of clinical interest, two particular innovations in sequencing technology have been transformative. First, through hybridization-based exon capture, one can isolate DNA corresponding to key cancer-associated genes for targeted mutation profiling [8, 9]. Second, through ligation of molecular barcodes (i.e. DNA sequences 6-8 nucleotides in length), one can pool hundreds of samples per sequencing run and fully take advantage of the ever-increasing capacity of massively parallel sequencing instruments [10]. When combined, these innovations enable tumors to be profiled for

lower cost and at higher throughput, with smaller computational requirements [11]. Further, by redistributing sequence coverage to only those genes most critical to the particular application, one can achieve greater sequencing depth for higher detection sensitivity for low allele frequency events.

Here we describe our IMPACT assay (Integrated Mutation Profiling of Actionable Cancer Targets), which utilizes exon capture on barcoded sequence library pools by hybridization using custom oligonucleotides to capture all protein-coding exons and select introns of 279 key cancer-associated genes (Table 1). This strategy enables the identification of mutations, indels, copy number alterations, and select structural rearrangements involving these 279 genes. Our method is compatible with DNA isolated from both fresh frozen and FFPE tissue as well as fine needle aspirates and other cytology specimens.

### **Protocol Text:**

#### **1.) DNA and Reagent Preparation**

Note: This protocol describes the simultaneous processing and analysis of 24 samples (e.g., 12 tumor/normal pairs) but can be adapted for smaller and larger batches. DNA samples may derive from FFPE or fresh frozen tissue, cytological specimens, or blood. Typically, both tumor and normal tissue from the same patient will be profiled together in order to distinguish somatic mutations from inherited polymorphisms. The protocol begins immediately following DNA extraction.

1.1) Aliquot 50 - 250 ng (250 ng recommended) of extracted DNA per sample, diluted in 1X Tris-EDTA (pH 8.0) buffer to a final volume of 50 µl, into separate Covaris round bottom tubes.

1.2) Shear the DNA on the Covaris E220 instrument at the following shear settings: 360 seconds at 10 duty factor, 175 PIP, and 200 cycles per burst.

1.3) Thaw AMPure XP Beads (Table 2) to room temperature. Thaw all buffers and enzymes on ice prior to the appropriate step.

#### **2.) Library Preparation – End Repair Module**

2.1) Prepare the end repair master mix in a sterile microcentrifuge tube using the following volumes per sample: 10 µl of 10X NEBNext End Repair Reaction Buffer (Table 2), 5 µl NEBNext End Repair Enzyme Mix (Table 2), 35 µl sterile nuclease free H<sub>2</sub>O.

2.2) Aliquot the 50 µl of each sheared DNA into separate wells of a 96-well plate. Add 50 µl of the prepared end repair master mix to each reaction. Incubate in a thermocycler for 30 min at 20°C.

#### **3.) Post End Repair Clean-up**

3.1) Add 2X volume (200 µl) of AMPure XP Beads (Table 2) to each sample and gently pipette the entire volume up and down 10 times using a multichannel pipettor.

3.2) Incubate at room temperature for 15 minutes.

3.3) Place on the magnetic stand (Table 2) at room temperature for 15 minutes until the sample appears clear.

3.4) Gently remove and discard clear supernatant taking care not to disturb beads. Some liquid may remain in wells.

3.5) With tubes on the stand, gently add 200 µl of freshly prepared 80% ethanol to each sample and incubate plate at room temperature for 30 seconds. Carefully, remove ethanol by pipette.

3.6) Repeat step 5, for a total of 2 ethanol washes. Ensure all ethanol has been removed. Remove from the magnetic stand and let dry at room temperature for 5 minutes.

3.7) Resuspend dried beads with 44.5 µl sterile H<sub>2</sub>O. Gently, pipette entire volume up and down 10 times mixing thoroughly. Ensure beads are no longer attached to the side of the well.

3.8) Incubate resuspended beads at room temperature for 2 minutes. Place on magnetic stand for 5 minutes until the sample appears clear.

3.9) Gently transfer 42 µl of clear supernatant containing the sample to a new well.  
*The procedure may be safely stopped at this step and samples stored at -20°C for up to 7 days. To restart, thaw frozen samples on ice before proceeding.*

#### **4.) Library Preparation – dA-Tailing Module**

4.1) Prepare the dA-Tailing master mix in a sterile microcentrifuge tube using the following volumes per sample: 5 µl of 10X NEBNext dA-Tailing Reaction Buffer (Table 2), 3 µl (3' → 5' exo-) Klenow Fragment (Table 2).

4.2) Add 8 µl of the prepared dA-Tailing master mix to each well containing 42 µl of end-repaired DNA. Incubate in a thermocycler for 30 minutes at 37°C.

4.3) Perform a post dA-Tailing clean-up: Repeat steps 3.1-3.10 using 2X volume (100 µl) of beads, but resuspend in sterile H<sub>2</sub>O to a final volume of 33.75 µl.  
*The procedure may be safely stopped at this step and samples stored at -20°C for up to 7 days. To restart, thaw frozen samples on ice before proceeding.*

#### **5.) Library Preparation – Adapter Ligation Module**

5.1) Prepare the Ligation master mix in a sterile microcentrifuge tube using the following volumes per sample: 10 µl of 5X NEBNext Quick Ligation Reaction Buffer (New England Biolabs, Table 2), 5 µl Quick T4 DNA Ligase (Table 2).

5.2) Add 15 µl of the Ligation master mix to each well containing the dA-Tailed DNA.

5.3) Add 1.25 µl of the appropriate 25 µM NEXTflex barcoded adapter (Table 2) to each well. Each sample should receive a separate barcoded adapter. Incubate in a thermocycler for 15 minutes at 20°C.

5.4) Perform a post adapter ligation clean-up: Repeat steps 3.1-3.10 using 1X volume (50 µl) of beads, but resuspend in sterile H<sub>2</sub>O to a final volume of 50 µl.

5.5) Repeat steps 3.1-3.10 for a second 1X volume (50 µl) clean-up, but resuspend in sterile H<sub>2</sub>O to a final volume of 23 µl.

*The procedure may be safely stopped at this step and samples stored at -20°C for up to 7 days. To restart, thaw frozen samples on ice before proceeding.*

## **6.) Library Amplification**

6.1) Prepare the KAPA HiFi master mix in a sterile microcentrifuge tube using the following volumes per sample: 25 µl 2X KAPA HiFi HS RM (Table 2), 1 µl each of 100 µM Primer 1 and Primer 2 (Table 3).

6.2) Add 27 µl of the master mix to each well containing the ligation product. Set pipette to 50 µl and gently pipette the entire volume up and down 10 times.

6.3) Place in the thermocycler for the following PCR cycles: initial denaturation of 45 sec at 98°C, 10 cycles of 15 sec at 98°C, 30 sec at 60°C, 30 sec at 72°C, and final extension of 1 min at 72°C.

6.4) Perform a post amplification clean-up: Repeat steps 3.1-3.10 using 1X volume (50 µl) of AMPure XP Beads, and resuspend in sterile water to a final volume of 30 µl.

6.5) Quantify the concentration of each library using Qubit Broad Range Assay (Life Technologies, Table 2) according to the manufacturer's instructions.

## **7.) Roche Nimblegen SeqCAP EZ Library Hybridization, Wash, and Amplification**

7.1) Thaw universal and index oligonucleotide blockers (Table 3), SeqCap EZ Library, and Nimblegen capture components (COT DNA, 2X hybridization buffer, and hybridization component A, Table 2) on ice. All oligonucleotide blockers are stored in working stocks of 1 mM. Our SeqCap EZ Library is a custom-designed pool of capture probes encompassing all protein-coding exons of 279 cancer-associated genes, though other custom and catalog designs may be used.

7.2) Prepare a 1 mM pool of index oligonucleotide blockers corresponding to the specific barcoded adaptor sequences used in library preparation. Combine 1 µl of each blocker, and vortex.

7.3) In a new sterile 1.5 mL microcentrifuge tube, add 5 µl of 1mg/mL COT DNA, 2 µl of 1 mM universal blocker and 2 µl of the 1 mM pool of index oligonucleotide blockers.

7.4) Pool 24 barcoded libraries into a single reaction. Add a total of 1-3 µg of pooled, barcoded sequence library (prepared above) to the capture mixture. For 24 samples, 100 ng per sample is recommended. When pooling tumors and matched normal samples, we recommended adding more tumor library (e.g. 100 ng per tumor and 50 ng per normal) so that they are sequenced to greater depth of coverage.

7.5) Close the tube's cap and make 15-20 holes in the top of the cap with an 18-20 gauge or smaller needle. Speed Vac the amplified sample library/COT DNA/Oligos in a DNA vacuum concentrator on high heat (60°C).

7.6) Once completely dried down, rehydrate with 7.5 µl of 2X hybridization buffer and 3 µl of hybridization component A. The tube should now contain the following: 5 µg COT DNA, variable µg amplified sample library, 1000 pmol of Universal and Index Oligos, 7.5 µl 2X hybridization buffer, 3 µl hybridization component A in a total volume of 10.5 µl.

7.7) Cover the holes in the cap with a small piece of laboratory tape. Vortex the sample for 10 seconds and centrifuge at maximum speed for 10 seconds.

7.8) Incubate the sample at 95 °C for 10 minutes to denature the DNA, followed by centrifugation at maximum speed for 10 seconds at room temperature.

7.9) In a new 0.2mL PCR strip tube, prepare 2.25 µl aliquot of SeqCap EZ Library capture probes and 2.25 µl of sterile nuclease free H<sub>2</sub>O and add the contents of step 7.7. Gently pipette the entire volume up and down 10 times.

7.10) Incubate in a thermal cycler at 47°C for 48-96 hrs. Set and maintain the thermal cycler's lid temperature to 57°C.

7.11) Follow steps in Chapter 6 of the NimbleGen SeqCap EX Library SR User's Guide v3.0 (Table 2) for washing and recovery of the captured DNA.

7.12) Follows steps in Chapter 7 of the Roche NimbleGen protocol for amplification of the captured DNA with the following modifications:

- Use 2X KAPA HiFi Polymerase (Table 2) instead of Phusion High-Fidelity PCR Master Mix
- Use 100 µM Primer 1 and Primer 2 listed in Table 3.
- Run the following PCR conditions: initial denaturation at 98°C for 30 seconds, 11 cycles of 98°C for 10 seconds, 60°C for 30 seconds, 72°C for 30 seconds, and final extension at 72°C for 5 minutes. Hold at 4°C.

7.13) Quantify the concentration of the amplified captured DNA using Qubit High Sensitivity Assay (Table 2) according to the manufacturer's instructions.

7.14) Analyze the captured DNA quality using Agilent Bioanalyzer DNA HS chip (Table 2) according to the manufacturer's instructions.

- The PCR yield should be at least 100 ng (range 100 – 1000 ng).
- The average fragment length should be between 150-400 base pairs.

## 8.) Illumina Hi-Seq Sequencing

8.1) Dilute the amplified captured DNA to 2 pM, and load the sample on a single lane of an Illumina HiSeq 2000 flow cell. Sequence paired-end 75 base pair reads according to the



manufacturer's instructions. In our experience, one lane is sufficient to produce 500-700x unique sequence coverage per sample across 279 genes for a pool of 24 samples.

8.2) Illumina software (Real Time Analysis) will convert high-resolution images to cluster intensities to base calls and quality scores. Use CASAVA to demultiplex data according to barcode identity and generate individual FASTQ files for each sample.

## 9.) Data Analysis

9.1) Trim adapter sequences from the 3' end of sequence reads in FASTQ files using Cutadapt (<http://code.google.com/p/cutadapt/>). The minimum overlap length (-O) is 3 with an error rate (-e) of 1%. The minimum base length for retention of paired reads after trimming is 25.

9.2) Align sequence reads to the reference human genome hg19 using the Burrows-Wheeler Aligner tool (BWA) [12]. Perform post-processing steps including duplicate marking, local multiple sequence re-alignment, and base quality score recalibration using the Genome Analysis Toolkit (GATK) according to their standard best practices (<http://www.broadinstitute.org/gatk/guide/topic?name=best-practices>) [13]. When profiling multiple samples from the same patient (e.g., matched tumor and normal), local multiple sequence re-alignment should be performed jointly. The output of these post-processing steps is a separate BAM file for each sample, which contains all sequence, quality, and alignment information and can be loaded directly into the Integrative Genomics Viewer (IGV) for visualization of reads and sequence variants [14].

9.3) Calculate sequence performance metrics using Picard tools (<http://picard.sourceforge.net/>). Informative metrics include alignment rate, fragment size distribution, GC-content associated coverage bias, on-target capture specificity, PCR duplicate rate, library complexity, and mean target coverage. Representative values are displayed in Figure 1. Allele frequencies at common polymorphism sites calculated using DepthOfCoverage from GATK are used to monitor contamination from unrelated DNA and to ensure that matched samples came from the same patients.

*The following computational analyses depend on the presence of matched tumor and normal samples for the detection of somatic genetic alterations. Unmatched tumors may also be analyzed using a separate normal control sample, but most somatic mutations cannot easily be distinguished from inherited sequence variants.*

9.4) Call somatic single nucleotide variants (SNVs) for each tumor-normal pair using a somatic mutation caller, such as muTect [15], SomaticSniper [16], or Strelka [17]. We use muTect with modified filtering criteria, allowing for variants with low (non-zero) allele counts in the normal sample as long as the allele frequency is at least 5 times greater in the tumor. Variants recurrently called in a panel of normal DNAs are removed as likely systematic artifacts of sequencing or alignment. Each SNV passing filters is then carefully reviewed using IGV.

9.5) Call somatic indels for each tumor-normal pair using algorithms such as SomaticIndelDetector [13], Dindel [18], or SOAPindel [19]. We use SomaticIndelDetector with the following criteria: at least 10% allele frequency with at least 3 supporting reads, or at least

2% allele frequency with at least 10 supporting reads. We allow for indels with low (non-zero) allele counts in the normal sample as long as the allele frequency is at least 5 times greater in the tumor. Indels called in a panel of normal DNAs are removed, as above. Each indel passing filters is then carefully reviewed using IGV.

9.6) Extrapolate copy number status from the sequence coverage at target exons. For each individual sample, perform a Loess normalization of mean sequence coverage for all target exons, regression over GC percentage. Adjust the unnormalized coverage values by subtracting the Loess fit and adding the sample-wide median coverage. Calculate the ratio in normalized sequence coverage across all target exons for tumor-normal pairs. In the absence of a matched normal sample and/or to decrease noise, other gender-matched diploid normal controls lacking germline copy number variants may be used. Determine somatic copy number gains and losses based on increases or decreases in the coverage ratio.

9.7) Call somatic rearrangements involving cancer genes where at least one genomic breakpoint falls within or near a targeted interval of the genome. Suggested algorithms include GASV [20], Breakdancer [21], CREST [22], and dRanger [23]. Intrachromosomal inversions, deletions, and tandem duplications may be distinguished based on the relative position and orientation of supporting reads in discordant pairs. All candidate rearrangements should be manually reviewed using IGV. We recommend experimental validation by PCR and Sanger sequencing across the putative rearrangement breakpoints.

## **Representative Results**

One pool of 24 barcoded sequence libraries (12 tumor-normal pairs) was captured using probes corresponding to all protein-coding exons of 279 cancer genes and sequenced as 2x75 bp reads on a single lane of a HiSeq 2000 flow cell. Tumor and normal libraries were pooled in a 2:1 ratio. Sample performance metrics for a pool of frozen tumor DNA samples are shown in Figure 1, including alignment rate, fragment size distribution, on-target capture specificity, and mean target coverage. Example somatic mutations, insertions, deletions, and copy number alterations are shown in Figures 2-4.

**Tables and Figures** 4 figures and 3 tables in separate documents.

Figure 1: Representative figures of the sequence performance metrics. a) cluster density and alignment rate, b) mean target coverage, c) insert size distribution, and d) capture specificity.

Figure 2: Integrative Genomics Viewer (IGV) image showing specificity of sequence coverage at exons of EGFR in a lung cancer (top) and matched normal tissue (bottom). Gray bars represent unique sequence reads. At right, a heterozygous T>G mutation is present in 24% of reads in the tumor and 0% of reads in the normal, indicating that it is a somatic L858R amino acid substitution.

Figure 3: Examples of indels in a colorectal cancer tumor-normal pair: a) Somatic frameshift insertion in APC. b) Somatic frameshift deletion of 7 base pairs in TP53.

Figure 4: Copy number alterations in a tumor-normal pair. Each data point represents a single exon from 279 target genes. Copy number gains and losses are inferred from increases and decreases in tumor sequence coverage.

Table 1: Probe Design Features for Target Capture

Table 2 Reagents and Kits

Table 3: Index Oligo Blockers & Primers Sequences

## Discussion

Our IMPACT assay produces a high alignment rate, high on-target rate, high target coverage, and high sensitivity for detecting mutations, indels, and copy number alterations. We have demonstrated the capability of our IMPACT assay to sequence DNA from both fresh frozen and archived FFPE samples of low DNA input. By performing targeted exon sequencing of key cancer-associated genes, one can achieve very deep sequence coverage for the exons of these most critical genes thereby maximizing the ability to detect low frequency mutations. The use of barcoding and multiplexing enables higher throughput, lower cost per sample, and lower input amounts of DNA.

A benefit of the targeted approach is that it is feasible to optimize capture probes such that all exons are covered uniformly and sufficiently for mutation detection. As shown in Table 1, our design encompasses 4,535 exons in 279 genes. Following iterative improvements to our design, a typical sequencing experiment produces no more than 2% of exons at less than 20% of the median coverage for the entire sample. Thus, for a tumor sequenced to 500x sequence coverage (a conservative estimate as displayed in Figure 1), >98% of exons will be covered by >100x depth, guaranteeing high sensitivity for identifying low frequency mutations. This uniformity of coverage can be largely attributed to the flexibility of the Nimblegen SeqCap system, where probes may be synthesized to different lengths in regions of different nucleotide composition. We have also successfully used individually synthesized biotinylated oligonucleotides from Integrated DNA Technologies (IDT) to add new content to our panel and to boost the coverage of regions that are difficult to capture and/or sequence. By also capturing select introns of recurrently rearranged genes, we are able to identify structural rearrangements producing fusion genes even when only one fusion partner is captured.

This targeted sequencing approach does possess two main limitations in the ability to detect “driver” genetic alterations in tumors. First, only a fraction of the genome is interrogated. Targeted sequencing is, by nature, limited. As new cancer genes emerge from systematic genomics efforts, they may be rapidly added to a capture panel, but functionally important mutations may be missed that would be detected by whole genome or whole exome sequencing. Second, it is limited to DNA-based alterations. The detection of aberrantly expressed transcripts, novel splice isoforms, and gene fusions requires RNA-Seq or similar techniques [24]. Epigenetic alterations such as chromatin modifications and DNA methylation may also play key roles in tumorigenesis and tumor progression. As sequencing technologies continue to improve, one can envision integrated strategies employing whole genome sequencing, RNA-Seq and complementary approaches to comprehensively characterize the genetic and epigenetic make-up

of individual tumors on a routine basis [25]. As complete integrated strategies are currently cost- and computationally-prohibitive, targeted sequencing presents a practical alternative to capture the most salient genomic features for clinical and translational applications.

Our protocol assumes that tumors are sequenced alongside matched normal tissue taken from the same patient. The purpose of this normal control is to determine whether sequence variants detected in the tumor are inherited or somatically acquired. In the absence of a matched normal, one can infer that variants occurring at mutational hotspots (i.e. sites of recurrent somatic mutations in other cancers) are likely to be somatic. Further, analysis of copy number gains and losses can be performed with a non-genetically matched diploid normal sample. For all other novel genetic alterations, the simultaneous analysis of matched normal tissue greatly simplifies the interpretation of the resulting mutation data and is highly recommended.

The IMPACT protocol described here demonstrates consistency and high quality performance for frozen samples. We have observed occasional variability when working with FFPE samples, though. Depending on the age and quality of the FFPE sample, the DNA may be severely degraded or chemically-modified, and as a result may render sequencing and analysis difficult [26]. That being said, we have found these experimental conditions to be reliable for the large majority of FFPE samples. We have also successfully applied this protocol to characterize other clinically relevant specimens such as biopsies, fine needle aspirates, and cytologic specimens. It is for this reason—the flexibility to accommodate different types of specimens of variable quality, quantity, and heterogeneity—that this method stands to **impact** the diagnosis and treatment of cancer patients in the clinical arena.

#### **Acknowledgments:**

We thank Dr. Agnes Viale and the MSKCC Genomics Core Laboratory for technical assistance. This protocol was developed with support from the Geoffrey Beene Cancer Research Center and the Farmer Family Foundation.

#### **Disclosures:**

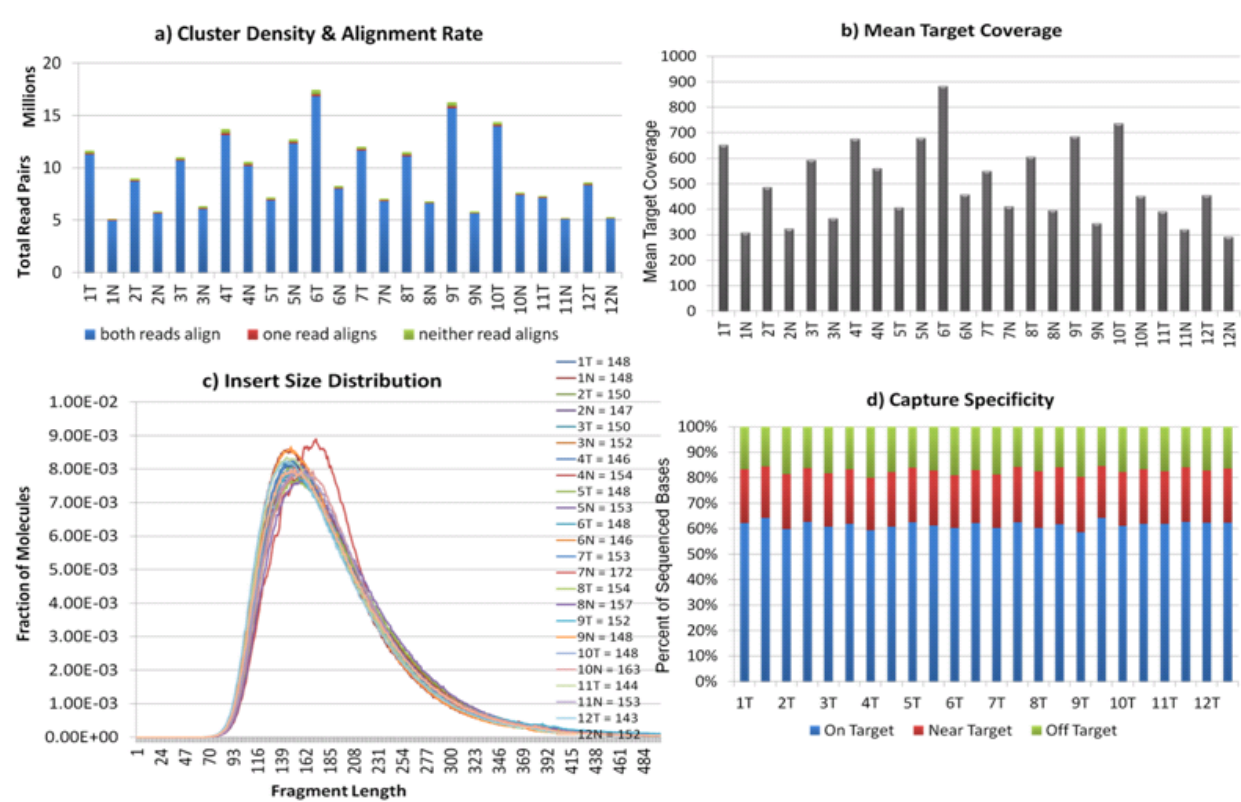
Dr. Berger has received consulting fees and research funding from Foundation Medicine, a cancer diagnostics company.

#### **References:**

1. Stratton MR, Campbell PJ, Futreal PA: The cancer genome. *Nature* 458(7239), 719-724 (2009).
2. Meyerson M, Gabriel S, Getz G: Advances in understanding cancer genomes through second-generation sequencing. *Nat Rev Genet* 11(10), 685-696 (2010).
3. Thomas RK, Baker AC, Debiasi RM *et al.*: High-throughput oncogene mutation profiling in human cancer. *Nature genetics* 39(3), 347-351 (2007).
4. Macconail LE, Campbell CD, Kehoe SM *et al.*: Profiling critical cancer gene mutations in clinical tumor samples. *PLoS One* 4(11), e7887 (2009).
5. Dias-Santagata D, Akhavanfard S, David SS *et al.*: Rapid targeted mutational analysis of human tumours: a clinical platform to guide personalized cancer medicine. *EMBO Mol Med* 2(5), 146-158 (2010).
6. Macconail LE, Van Hummelen P, Meyerson M, Hahn WC: Clinical implementation of comprehensive strategies to characterize cancer genomes: opportunities and challenges. *Cancer Discov* 1(4), 297-311 (2011).

7. Taylor BS, Ladanyi M: Clinical cancer genomics: how soon is now? *J Pathol* 223(2), 318-326 (2011).
8. Gnirke A, Melnikov A, Maguire J *et al.*: Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature biotechnology* 27(2), 182-189 (2009).
9. Mamanova L, Coffey AJ, Scott CE *et al.*: Target-enrichment strategies for next-generation sequencing. *Nat Methods* 7(2), 111-118 (2010).
10. Craig DW, Pearson JV, Szelinger S *et al.*: Identification of genetic variants using bar-coded multiplexed sequencing. *Nat Methods* 5(10), 887-893 (2008).
11. Wagle N, Berger MF, Davis MJ *et al.*: High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer Discov* 2(1), 82-93 (2012).
12. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14), 1754-1760 (2009).
13. Depristo MA, Banks E, Poplin R *et al.*: A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics* 43(5), 491-498 (2011).
14. Robinson JT, Thorvaldsdottir H, Winckler W *et al.*: Integrative genomics viewer. *Nature biotechnology* 29(1), 24-26 (2011).
15. Cibulskis K, Lawrence MS, Carter SL *et al.*: Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature biotechnology* 31(3), 213-219 (2013).
16. Larson DE, Harris CC, Chen K *et al.*: SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* 28(3), 311-317 (2012).
17. Saunders CT, Wong WS, Swamy S, Becq J, Murray LJ, Cheetham RK: Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* 28(14), 1811-1817 (2012).
18. Albers CA, Lunter G, Macarthur DG, Mcvean G, Ouwehand WH, Durbin R: Dindel: accurate indel calls from short-read data. *Genome Res* 21(6), 961-973 (2011).
19. Li S, Li R, Li H *et al.*: SOAPindel: Efficient identification of indels from short paired reads. *Genome Res* 23(1), 195-200 (2013).
20. Sindi S, Helman E, Bashir A, Raphael BJ: A geometric approach for classification and comparison of structural variants. *Bioinformatics* 25(12), i222-230 (2009).
21. Chen K, Wallis JW, McLellan MD *et al.*: BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* 6(9), 677-681 (2009).
22. Wang J, Mullighan CG, Easton J *et al.*: CREST maps somatic structural variation in cancer genomes with base-pair resolution. *Nature methods* 8(8), 652-654 (2011).
23. Drier Y, Lawrence MS, Carter SL *et al.*: Somatic rearrangements across cancer reveal classes of samples with distinct patterns of DNA breakage and rearrangement-induced hypermutability. *Genome Res*, (2012).
24. Wang Z, Gerstein M, Snyder M: RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10(1), 57-63 (2009).
25. Roychowdhury S, Iyer MK, Robinson DR *et al.*: Personalized oncology through integrative high-throughput sequencing: a pilot study. *Sci Transl Med* 3(111), 111ra121 (2011).
26. Kerick M, Isau M, Timmermann B *et al.*: Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. *BMC Med Genomics* 4, 68 (2011).

Figure 1



**Figure 2**

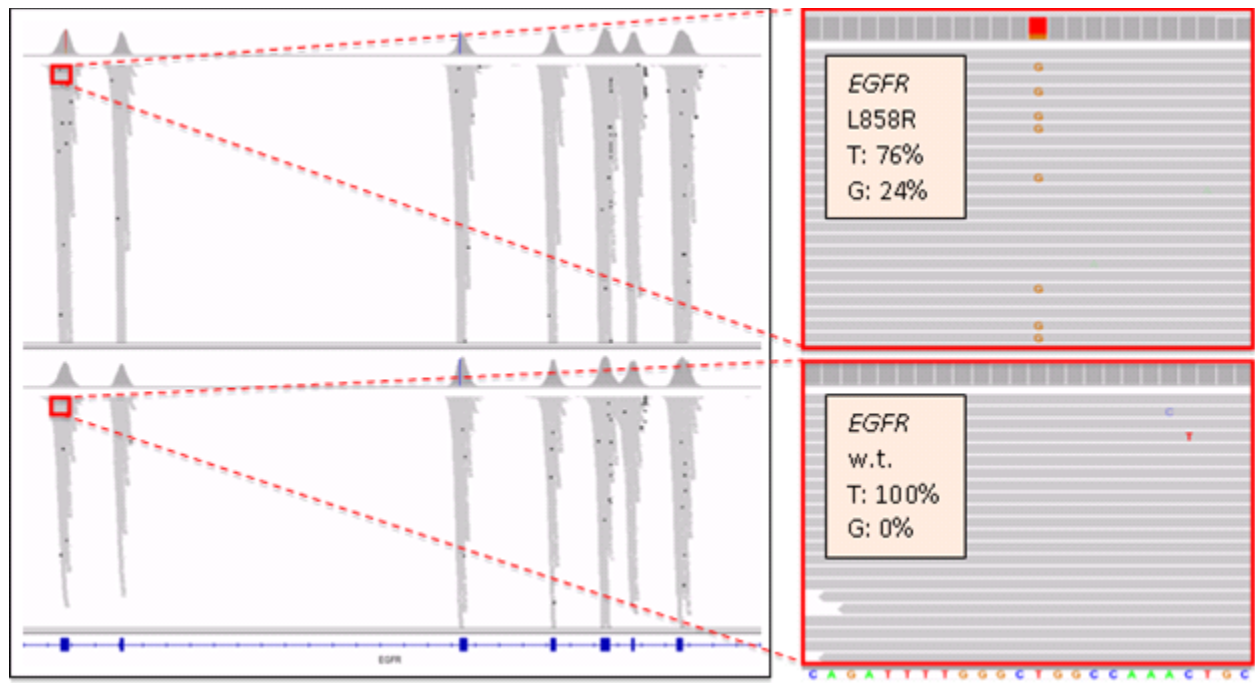
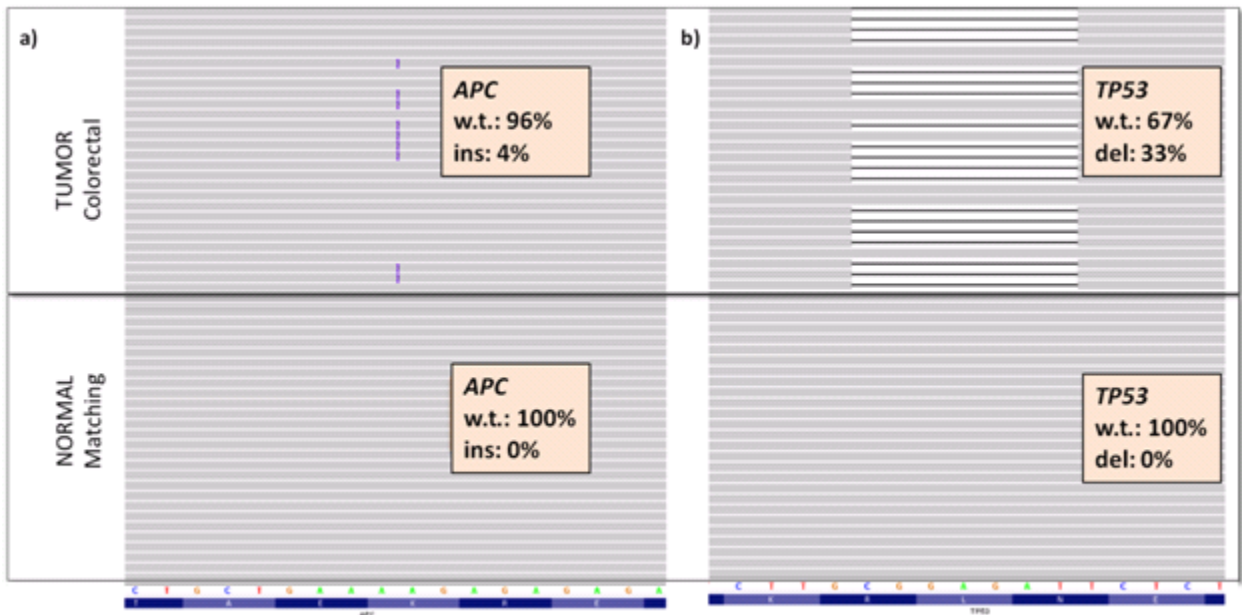
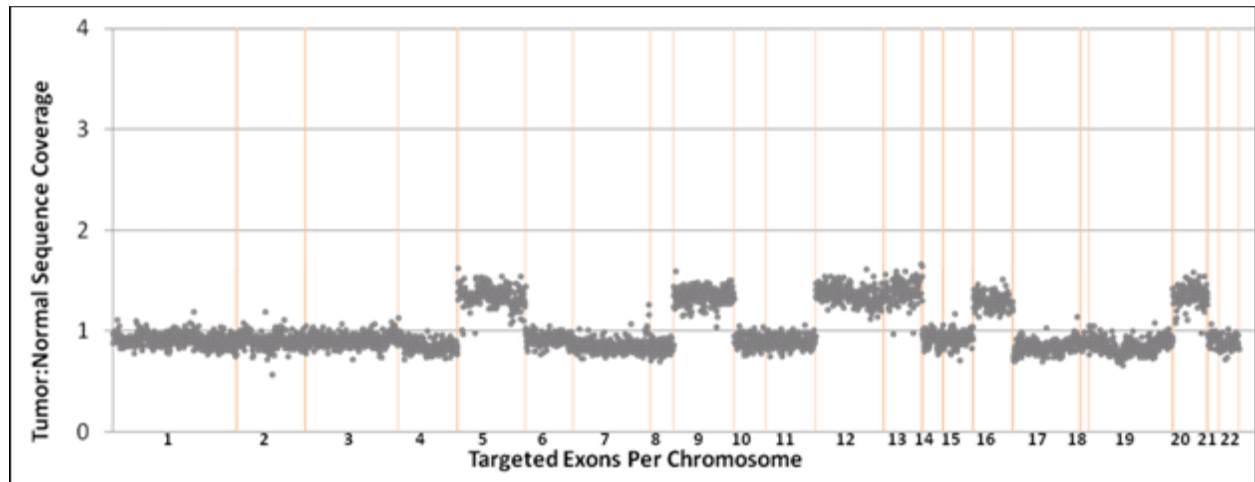


Figure 3





**Figure 4**



**Table 1: Target Capture Design**

Total genes	279
Total exons	4535
Total introns	14
SNPs*	1042
Total target territory	879,966 basepairs
Total probe territory	1,400,415 basepairs

\*Probes for SNPs with high minor allele frequency are tiled across all chromosome arms to aid with fingerprint genotyping, assessment of contamination, and estimation of copy number alterations.

**Table 2**

<b>Name of Reagent/Material</b>	<b>Company</b>	<b>Catalog Number</b>	<b>Comments</b>
NEBNext End Repair Module	New England Biolabs	E6050L	
NEBNext dA-Tailing Module	New England Biolabs	E6053L	
NEBNext Quick Ligation Module	New England Biolabs	E6056L	
Agencourt AMPure XP	Beckman Coulter Genomics		
NEXTflex PCR-Free Barcodes – 24	Bioo Scientific	514103	
HiFi Library Amplification Kit	KAPA Biosystems	KK2612	
COT Human DNA, Fluorometric Grade	Roche Diagnostics	05 480 647 001	
NimbleGen SeqCap EZ Hybridization and Wash kit	Roche NimbleGen	05 634 261 001	
SeqCap EZ Library Baits	Roche NimbleGen		
QIAquick PCR Purification Kit	Qiagen	28104	
Qubit dsDNA Broad Range (BR) Assay Kit	Life Technologies	Q32850	
Qubit dsDNA High Sensitivity (HS) Assay Kit	Life Technologies	Q32851	
Agilent DNA HS Kit	Agilent Technologies	5067-4626, 4627	
Agilent 2100 Bioanalyzer	Agilent Technologies		
Covaris E220	Covaris		
Magnetic Stand-96	Ambion	AM10027	
Illumina Hi-Seq 2000	Illumina		

**Table 3: Index Oligo Blockers & Primers Sequences**

TS-HE Universal Blocker	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
TS-INV-HE Index 1 Blocker	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 2 Blocker	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 3 Blocker	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 4 Blocker	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 5 Blocker	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 6 Blocker	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 7 Blocker	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 8 Blocker	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 9 Blocker	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 10 Blocker	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 11 Blocker	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 12 Blocker	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 13 Blocker	CAAGCAGAAGACGGCATACGAGATTTGACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 14 Blocker	CAAGCAGAAGACGGCATACGAGATGGAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 15 Blocker	CAAGCAGAAGACGGCATACGAGATTGACATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 16 Blocker	CAAGCAGAAGACGGCATACGAGATGGACGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 18 Blocker	CAAGCAGAAGACGGCATACGAGATGCGGACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 19 Blocker	CAAGCAGAAGACGGCATACGAGATTTTCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 20 Blocker	CAAGCAGAAGACGGCATACGAGATGGCCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

TS-INV-HE Index 21 Blocker	CAAGCAGAAGACGGCATAACGAGATCGAAACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 22 Blocker	CAAGCAGAAGACGGCATAACGAGATCGTACGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 23 Blocker	CAAGCAGAAGACGGCATAACGAGATCCACTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 25 Blocker	CAAGCAGAAGACGGCATAACGAGATATCAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TS-INV-HE Index 27 Blocker	CAAGCAGAAGACGGCATAACGAGATAGGAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
<b>Primer 1</b>	<b>5'-AATGATACGGCGACCACCGAGATCT-3'</b>
<b>Primer 2</b>	<b>5'-CAAGCAGAAGACGGCATAACGAGAT-3'</b>



17 Sellers Street  
Cambridge, MA 02139  
tel: +1 617 935 9051  
www.JoVE.com

# ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article: Detecting Somatic Genetic Alterations in Tumor Specimens by Exon *Capture and*  
 Author(s): Helen H. Won, Sasinya N. Scott, A. Rose Brannon, Michael F. Berger *Massively Parallel Sequencing*

Item 1 (check one box): The Author elects to have the Materials be made available (as described at <http://www.jove.com/publish>) via: ☐ Standard Access ☒ Open Access

Item 2 (check one box):

- ☒ The Author is NOT a United States government employee.
- ☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.
- ☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

## ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "Agreement" means this Article and Video License Agreement; "Article" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "Author" means the author who is a signatory to this Agreement; "Collective Work" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "CRC License" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: <http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode>; "Derivative Work" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "Institution" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "JoVE" means MyJoVE Corporation, a Massachusetts corporation and the publisher of *The Journal of Visualized Experiments*; "Materials" means the Article and / or the Video; "Parties" means the Author and JoVE; "Video" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to Sections 4 and 7 below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in Item 1 above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in Section 3 above, the



## ARTICLE AND VIDEO LICENSE AGREEMENT

Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

**5. Grant of Rights in Video – Standard Access.** This Section 5 applies if the "Standard Access" box has been checked in Item 1 above or if no box has been checked in Item 1 above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to Section 7 below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

**6. Grant of Rights in Video – Open Access.** This Section 6 applies only if the "Open Access" box has been checked in Item 1 above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to Section 7 below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this Section 6 is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

**7. Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in Item 2 above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict

shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

**8. Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

**9. Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

**10. JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including,



## ARTICLE AND VIDEO LICENSE AGREEMENT

without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

11. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or

damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

12. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US\$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

13. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to be one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement required per submission.

### AUTHOR:

Name:

MICHAEL F. BERGER

Department:

DEPARTMENT OF PATHOLOGY & HUMAN ONCOLOGY AND PATHOGENESIS PROGRAM

Institution:

MEMORIAL SLOAN-KETTERING CANCER CENTER

Article Title:

DETECTING SOMATIC GENETIC ALTERATIONS IN TUMOR SPECIMENS BY EXON CAPTURE AND MASSIVELY PARALLEL SEQUENCING

Signature:



Date:

1/16/13

Please submit a signed and dated copy of this license by one of the following three methods:

- 1) Upload a scanned copy as a PDF to the JoVE submission site upon manuscript submission (preferred);
- 2) Fax the document to +1.866.381.2236; or
- 3) Mail the document to JoVE / Attn: JoVE Editorial / 17 Sellers St / Cambridge, MA 02139

For questions, please email editorial@jove.com or call +1.617.945.9051.

MS # (internal use):



## **Response to Reviewers**

**JoVE 50710**

**Won et al., Detecting Somatic Genetic Alterations in Tumor Specimens by Exon Capture and Massively Parallel Sequencing**

### **Editorial comments:**

Protocol text: Please highlight less than 3 pages of text to identify which portions of the protocol are most important to film; i.e. which steps should be visualized to best supplement the written section of the protocol.

*We have reduced the highlighted section to less than 3 pages of text, as requested.*

Please remove all trademark symbols (ie uploaded Materials table).

*We have removed all trademark symbols, as requested.*

Please use the attached version of your manuscript for future revisions.

*We have made several revisions, with changes tracked, to the version of the manuscript supplied with the reviewers' comments. The specific revisions are discussed below in our point-by-point response to the reviewers.*

### **Reviewer #1:**

A number of software tools were also listed in Data Analysis method section. However, much more detailed description of the implementation and integration of different analysis methods would be very beneficial.

*We have significantly modified Section 9 ("Data Analysis"), adding considerable details for each step in our analysis pipeline. Specific parameters and filtering criteria are listed for the algorithms that we use for adapter trimming, mutation calling, and indel calling. Additional details are provided for our process for determining copy number alterations.*

More detailed description about the capture library design would be much appreciated, especially for the capture of translocations.

*We have created a new table (Table 1 in the revised manuscript) describing the overall features of our custom capture design. In summary, our panel contains all protein coding exons of 279 cancer-associated genes (4,535 total exons). We also include 14 introns*

*from 7 genes that are recurrently rearranged in solid tumors, as well as >1000 SNPs for fingerprint genotyping and quality control. The total target territory is approximately 880 kilobases of genomic sequence.*

Basic technical performance was presented. But more in-depth analysis of the platform would be important to evaluate the platform for potential clinical applications. The analysis could address coverage and uniformity of sequencing as well as capture and sequencing bias. The variant analysis could address variant detection sensitivity, specificity and reproducibility for different classes of variants, including single nucleotide variants, small indels, large indels, translocations and copy number changes.

*We have added a new paragraph to the Discussion addressing these points. Please see the revised manuscript for the full paragraph. In summary, we stress the following points: (1) A benefit of the targeted approach is that it is feasible to optimize capture probes such that all exons are covered uniformly and sufficiently for mutation detection. (2) Following iterative improvements to our design, a typical sequencing experiment produces no more than 2% of exons at less than 20% of the median coverage for the entire sample. Thus, for a tumor sequenced to 500x sequence coverage (a conservative estimate as displayed in Figure 1), >98% of exons will be covered by >100x depth, guaranteeing high sensitivity for identifying low frequency mutations. (3) By also capturing select introns of recurrently rearranged genes, we are able to identify structural rearrangements producing fusion genes even when only one fusion partner is captured.*

## **Reviewer #2:**

Targeted sequencing is a well-known, well-used method. This article does not illuminate anything new or little-known in the method. The video would involve mostly footage of workers pipetting.

*We agree that experimentalists familiar with targeted sequencing methods may not gain as much from this video and accompanying protocol as someone who is new to next-generation sequencing. However, we believe that a detailed exposition of our method, especially in its simplicity, will be of significant value to laboratories beginning to invest resources and personnel into next-generation sequencing workflows. Our experience is that labs that have traditionally relied on specialized core facilities to handle all of their sample prep and sequencing are now beginning to prepare their own samples given the wider accessibility of sequencing platforms and options. We hope that with the availability of protocols like ours, this process will be less intimidating for labs new to the field. Further, it is our understanding that the Journal of Visualized Experiments has not yet published a protocol or produced a video describing the preparation of libraries for targeted sequencing, and we believe that our method will serve as a valuable resource.*

The area of great interest, the actual analysis of the data, is only touched on in a manner that it would be impossible to re-capitulate the methods in another lab.

*This sentiment was shared by Reviewer #1. As described above, we have significantly modified Section 9 (“Data Analysis”), adding considerable details for each step in our analysis pipeline. Specific parameters and filtering criteria are listed for the algorithms that we use for adapter trimming, mutation calling, and indel calling. Additional details are provided for our process for determining copy number alterations.*

**Reviewer #3:**

If the authors have a preference for which custom capture probes to use, i.e. Agilent versus Roche versus IDT, might be worth mentioning or discussing further.

*As mentioned in our response to Reviewer #1, we have added a new paragraph to the Discussion describing our capture performance and the uniformity of sequence coverage we have achieved through iterative improvements to our probe design. In the revised text, we state that the uniformity of sequence coverage can be largely attributed to the flexibility of the Nimblegen SeqCap system, in which probes are synthesized to different lengths in regions of different nucleotide composition so that they may have similar melting temperatures. We also describe how we have successfully spiked in individually synthesized biotinylated oligonucleotides from IDT to add new content to our panel and to boost the coverage of regions that are difficult to capture and/or sequence.*