

# Uncovering tumor heterogeneity in FFPE samples by laser capture microdissection and next-generation sequencing



## In this application note, we show:

- Laser capture microdissection (LCM) uncovers differences in allele and transcript abundance that are missed in whole tissue scrapes
- Ion AmpliSeq™ DNA and RNA sequencing panels empower the analysis of many sequences from 10 ng or less of FFPE tissue–extracted material
- Efficient cataloging of variants present in a tumor requires sampling different regions of the tumor

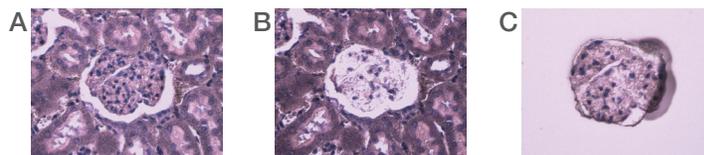
## Introduction

Current models of cancer progression postulate that tumors arise from a single mutated cell, followed by clonal expansion and acquisition of additional genetic and genomic alterations. The continual accumulation of new mutations can result in the appearance of different tumor subclones with a variety of phenotypic advantages. For example, these mutations could confer the ability to proliferate uncontrollably, leave the primary site where the original tumor mass occurred, and colonize different organs [1]. Intra-tumor heterogeneity, the presence of more than one clone of cancer cells within a given tumor mass, and inter-tumor heterogeneity, the presence of different genetic alterations in different metastatic tumors from a single source, have been identified in several tumor types [2-5]. A complete understanding of the biology, and ultimately, the design of treatments, requires efficiently identifying mutations present and transcript abundances in subsets of cells in a tumor.

Traditional analysis of a whole tumor mass in bulk could provide misleading results. First, the presence of normal or unrelated cells could overestimate the fraction of normal, or non-pathogenic alleles, that are in the tumorous parts of the specimen. Second, the abundance of unrelated alleles could

mask the ability to detect the pathogenic alleles. Finally, it is difficult to precisely isolate a relatively small region of interest by macrodissection of a specimen on a slide. Thus, although macrodissection-based tissue sampling is faster and less expensive than LCM, efficient analysis requires enrichment of mutant cells that is facilitated by precision laser microdissection.

The Applied Biosystems™ ArcturusXT™ LCM System offers the powerful combination of laser capture and laser cutting for microdissection applications (Figure 1). The solid-state infrared (IR) laser, exclusive to the ArcturusXT LCM System, delivers a gentle capture technique that helps preserve biomolecular integrity and is ideal for single cells and small numbers of cells. The solid-state UV laser permits superior speed and precision and is well suited for microdissecting dense tissue structures and for capturing large numbers of cells. Together, these two lasers provide the flexibility to capture individual cells and large regions from the same sample with minimal damage to the molecules contained within those regions.



**Figure 1. Example of the utility of laser capture microdissection. (A)** intact tissue, **(B)** tissue after removal of region of interest, and **(C)** cells contained within region of interest.

Microdissected tissue is captured onto Applied Biosystems™ CapSure™ LCM caps, allowing investigators to maintain sample custody throughout the entire LCM process. The microdissected areas can be inspected and verified by examining the LCM cap prior to downstream processing for genomic, transcriptomic, or proteomic analysis.

As scientific inquiries become more sophisticated, the need for extracting as much genetic information as possible from ever-smaller sample amounts has intensified. Moreover, clinical research specimens are often preserved by fixation in formalin or other chemical crosslinkers, and nucleic acids extracted from such samples are often degraded and difficult to work with. Fortunately, methods have been developed that can amplify target sequences to levels where molecular characterization becomes practical. Ion AmpliSeq™ technology delivers simple and fast targeted sequencing of specific genes and genomic regions. Based on ultrahigh-multiplex PCR, Ion AmpliSeq technology requires as little as 10 ng of input DNA or RNA to target sets of genes, making sequencing of formalin-fixed, paraffin-embedded (FFPE) samples routine on the Ion Personal Genome Machine™ (PGM™) System. Convenient, predesigned Ion AmpliSeq™ Ready-To-Use Panels offer extensive gene coverage for cancer and inherited diseases and allow researchers to focus on data generation and analysis, and not on the labor-intensive primer design and target selection steps. For maximum flexibility, the Ion AmpliSeq™ Designer allows custom panels to be designed using the simple, intuitive online tool. Since many different loci in a sample can be amplified and analyzed at the same time, the precious sample is most efficiently utilized. The Ion Torrent™ sequencing platform coupled with Ion AmpliSeq technology allows investigators to obtain the maximal amount of information on targeted loci from small amounts of sample or degraded samples.

In this application note, we describe an end-to-end workflow for extracting variant allele frequencies and transcript levels from custom targets using Ion AmpliSeq technology. We show that the ArcturusXT LCM System provides a rapid and reliable method to collect isolated cell populations from heterogenous tissue samples. We show that DNA and RNA extracted from FFPE samples can be analyzed using Ion AmpliSeq technology, allowing a targeted approach to mutation detection and gene expression analysis. Because many different loci can be targeted at one time, we show that multiplexed targeted sequencing can detect a large number of sequence variants from a single LCM sample. Finally, we show that tumors contain tremendous genetic heterogeneity, and that efficient cataloging of such heterogeneity requires the division, collection, and analysis of individual subregions of the tumor. The Thermo Fisher Scientific products described in the following workflow give investigators the tools that can facilitate understanding of the genetic basis of tumor formation and heterogeneity (Figure 2).

## Methods

### Tissue staining and LCM

Human lung FFPE tissue blocks were acquired from a commercial vendor (Asterand, Detroit, MI). The blocks were sectioned at 7 μm, mounted on slides, and stored at room temperature until use. Pathogenic cells were identified by staining with hematoxylin and eosin; adjacent sections were left unstained. Tumor cells in the hematoxylin and eosin-stained section were marked by a certified pathologist. Prior to laser capture, an unstained section was stained with the Applied Biosystems™ Arcturus™ Paradise™ PLUS FFPE LCM Staining Kit following the protocol indicated in the user manual. The stained slides and Applied Biosystems™ CapSure™ Macro LCM Caps were then loaded onto the ArcturusXT LCM System. Circles 2,000 μm in diameter were

defined in ArcturusXT™ LCM Software from the tumor regions and collected on the Applied Biosystems™ CapSure™ Macro LCM Caps. For comparison, whole tissue scrapes (WTS) from a slide, representing a mixed cell population sample, were processed. Genomic DNA was extracted from the caps by following the protocol in the Applied Biosystems™ PicoPure™ DNA Extraction Kit with one minor modification: instead of extracting in 50 μL as recommended in the protocol, we eluted overnight at 65°C in 20 μL of extraction buffer with periodic vortexing. This kept the extracted DNA at a high enough concentration such that further concentration was not necessary. Extracted DNA was quantified on an Invitrogen™ Qubit™ Fluorometer. Typically, yields were between 25–60 ng from the LCM-captured specimens and around 2 μg from the WTS.

### Ion AmpliSeq and OncoPrint panel library construction

We used an Ion AmpliSeq™ Colon and Lung Cancer Research Panel targeting 22 genes commonly mutated in lung and colon tumors. To help ensure complete coverage of these genes, the panel consists of 92 pairs of primers in a single pool, with an average amplicon length of 162 bp. We also used the OncoPrint™ Comprehensive Assay targeting 143 genes (2,531 amplicons) known to be commonly mutated in tumors. This larger panel allowed us to query many more genes and mutations, providing a more complete catalog of the different alleles that may be present. Ion AmpliSeq™ DNA libraries were constructed using 1–10 ng of DNA from each laser-captured isolate. The samples were processed using the Ion AmpliSeq™ Library Kit 2.0 and the Ion Library Equalizer™ Kit according to the recommended protocols. The template and enrichment steps were carried out on the Ion OneTouch™ 2 System. The samples were applied to an Ion 318™ Chip v2 platform and processed for sequence information on an Ion PGM System.

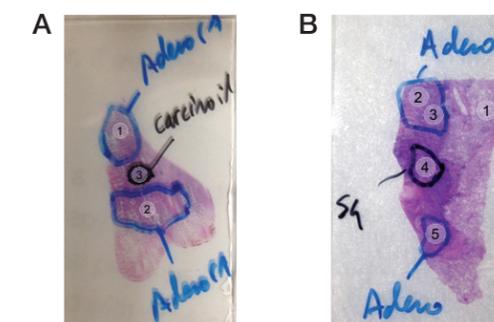
For RNA sequencing, a custom Ion AmpliSeq™ RNA Panel was designed and constructed consisting of 53 transcripts commonly mutated or differentially expressed in human cancers. Tissue samples were collected onto CapSure Macro LCM Caps from 2,000 μm diameter circles as previously described. RNA was extracted from the caps following the protocol in the Applied Biosystems™ Arcturus™ Paradise™ Plus RNA Extraction and Isolation Kit, and quantified on a Qubit Fluorometer. Total RNA (10 ng) was reverse transcribed using reagents in the Ion AmpliSeq™ RNA Library Kit as described in the manual. The targets were amplified by PCR for 22 cycles, and libraries were prepared for sequencing following the protocol described in the manual. Final library concentration was determined using the Ion Library TaqMan™

Quantification Kit and quality checked on a Bioanalyzer™ system (Agilent Technologies, Santa Clara, CA). Libraries were adjusted to a concentration of 40 pM.

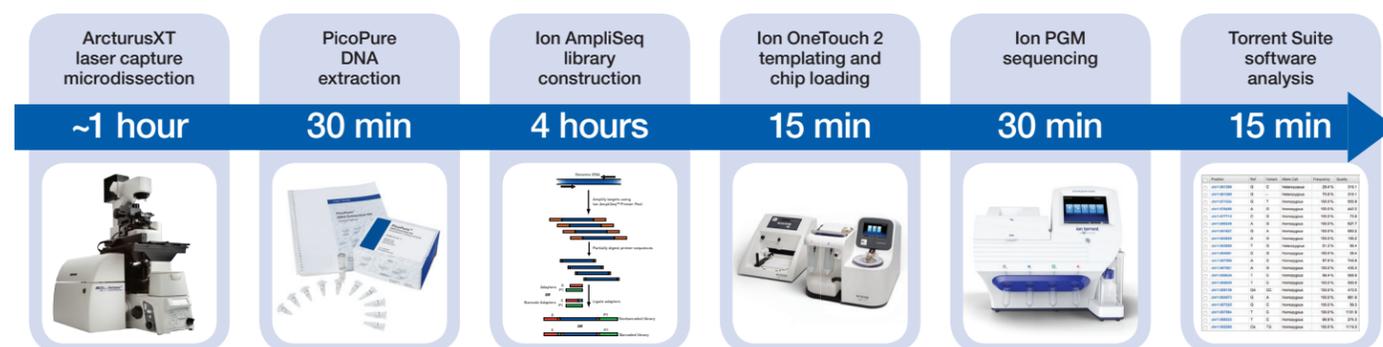
Sequencing templates were prepared and loaded onto chips from Ion 318™ Chip Kit v2 BC using an Ion Chef™ instrument and sequenced on an Ion PGM System. Mapped reads for each transcript were normalized to reads per million (RPM) reads using the Ion AmpliSeq™ plugin for Torrent Suite Software.

## Results

We focused on two specimens that showed visible diversity in cellular morphology across the tumors. One sample, designated 2182, was from a stage pT1b pN0 tumor. There were two types of tumorous cells distinguishable in this sample: two regions with adenoma-like cells (Figure 3A, regions 1 and 2) and one region with carcinoid cells (region 3). The second sample, designated 2162, was from a stage pT3 tumor (Figure 3B). It was divided into three different tumorous regions: two that had adenoma-like cells (regions 2, 3, and 5) and one with cells that were squamous in appearance (region 4). In addition, we included a region that was not marked as tumorous (region 1). Finally, for both samples, we collected a WTS that included all the microdissected regions.

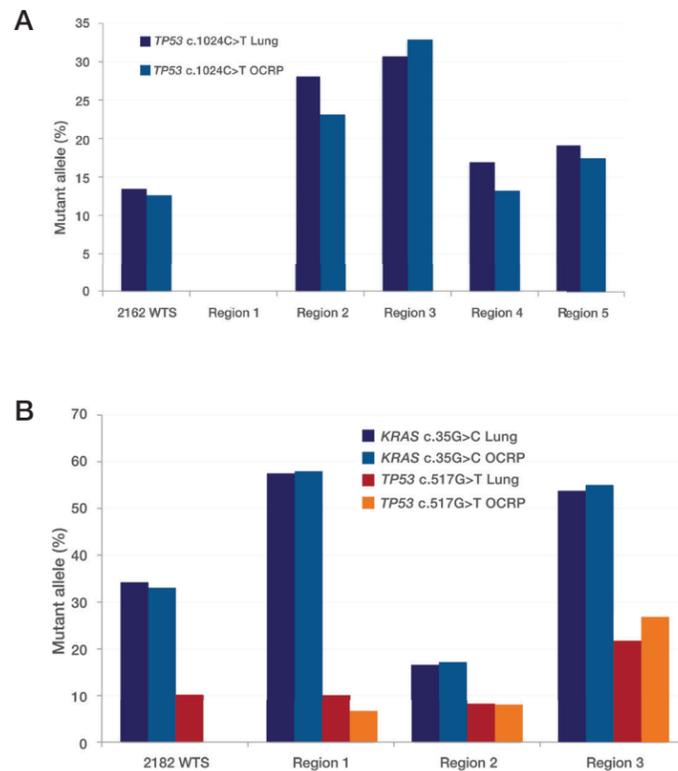


**Figure 3. Hematoxylin and eosin-stained sections and LCM-collected regions of lung cancer tumors used for this study. (A)** Sample 2182, a stage pT1b pN0 tumor. **(B)** Sample 2162, a stage pT3 tumor. Numbered circles are regions collected by LCM, and pen markings are pathologist-supplied descriptions of morphology. Adeno: adenoma-like; carcinoi: carcinoma-like; Sq: squamous.



**Figure 2. Overall workflow for obtaining targeted sequencing data from LCM of FFPE specimens.** The hands-on time (indicated above) is minimal; when the incubation and machine processing times are included, allelic information can be obtained in a total of three days. The workflow is identical for Ion AmpliSeq™ RNA experiments, except for an additional hour for cDNA synthesis after the extraction step.

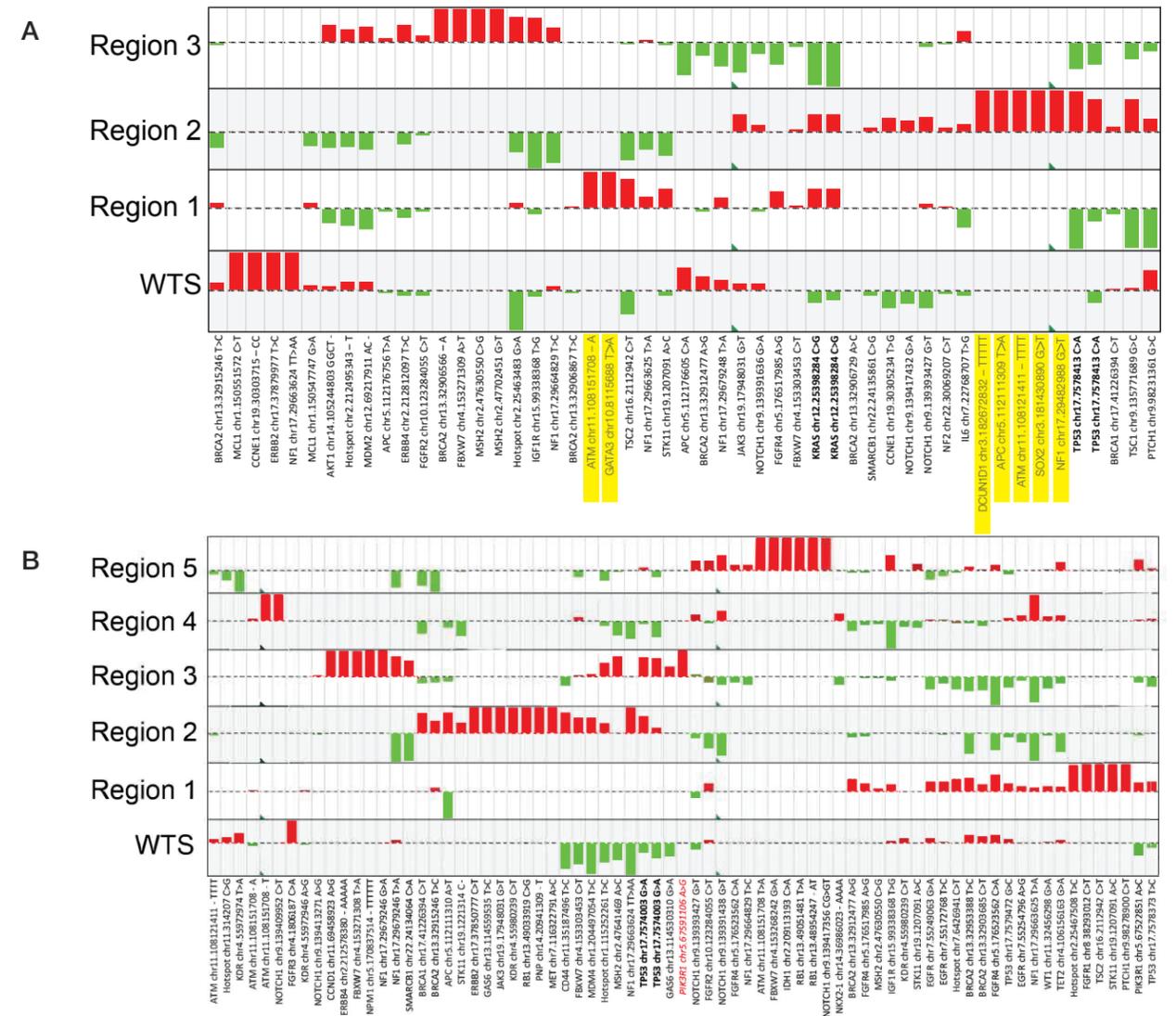
We first analyzed allele frequencies using the 22-gene Ion AmpliSeq Colon and Lung Cancer Research Panel. In the 2162 WTS (Figure 4A), a single pathogenic allele of *TP53* c.1024C>T (p.R342\*) was detected at a frequency of 13%. However, the frequency varied considerably across the different regions of the tumor, ranging from 17% to 30.7%. Notice that in the region not marked as tumorous (Figure 4A, region 1), no pathogenic alleles were detected with this panel. When the 2182 WTS was analyzed (Figure 4B), two different pathogenic alleles were detected: *KRAS* c.35G>C (p.G12A) was detectable at 34%, and *TP53* c.517G>T (p.V173L) at 10.2%. When smaller regions of the tumor were examined, the frequencies varied as described above. In most of the regions, the allele frequency was the same or greater than the WTS. However, the frequency of the *KRAS* allele was much reduced in region 3 of the tumor. These results show that LCM enriches for minority cells and can facilitate detection of pathogenic alleles in a complex mixture by increasing the frequency at which they are detected.



**Figure 4. Mutation analysis of lung cancer samples.** (A) Analysis of different subregions of sample 2162. Data from the Ion AmpliSeq™ custom lung cancer panel (dark blue bars) and OncoPrint Comprehensive Assay (red bars) are shown. WTS: whole tissue scrape. (B) Analysis of different subregions of sample 2182. Data from the Ion AmpliSeq custom lung cancer panel (dark blue and red bars) and OncoPrint Comprehensive Assay (OCR, light blue and orange bars) is shown. Note that the frequencies of alleles detected are very similar between the two panels.

To get a more complete understanding of the genetic heterogeneity of these tumors, we performed a similar analysis using the OncoPrint Comprehensive Assay. As previously described, many more loci were queried with this panel, and therefore using the OncoPrint Comprehensive Assay had the potential to detect many more variants in a sample. The OncoPrint Comprehensive Assay queries several loci that overlapped with the custom lung cancer panel, and therefore comparison of the allele frequencies detected by the two panels in the different regions was possible (Figure 4). There was an excellent correlation between the two panels, detecting almost the same frequencies of the mutants in the regions. One exception was that the *TP53* c.517G>T allele was not detected with the OncoPrint Comprehensive Assay in the 2182 WTS (Figure 4B). This might be because the frequency was low, and stochastic variation might have pushed the levels below the threshold of detection in that sample. Nevertheless, the overall correlation between the two panels in both samples was 98.3%, indicating that the two panels arrived at nearly the same answer for these alleles.

The OncoPrint Comprehensive Assay queries many more loci and thus revealed a much greater degree of tumor heterogeneity than the custom lung cancer panel. In sample 2182, there were a total of 49 variants found, and as might be expected, there were 72 variants detected in the later-staged 2162 tumor. However, the frequencies of each allele varied tremendously in the different regions of the tumor. To better illustrate the variation, the frequency in each sample was normalized to the average of the frequencies in each of the regions from that specimen (Figure 5). They were then grouped according to similar distribution patterns. For example, the five loci towards the rightmost portion of Figure 5A (*DCUN1D1* chr3.182672832 ins TTTTT to *NF1* chr17.29482988 C>T) were enriched in region 2, and in some cases detectable only in that region and in none of the others. Similarly, the *GATA3* chr10.8115688 T>A allele and *ATM* chr11.08151708 ins A allele detected were only in region 1 of sample 2182, and not in the others.



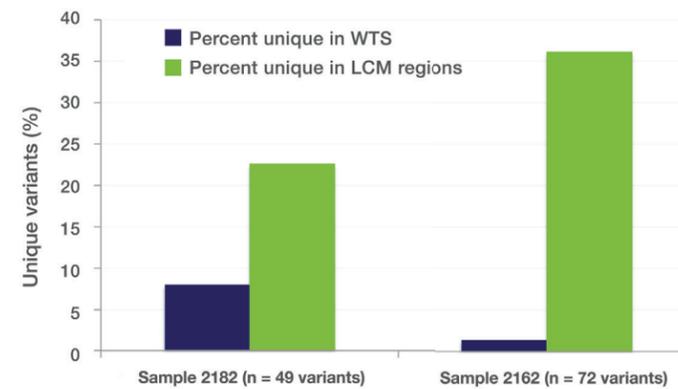
**Figure 5. Allele frequency analysis of lung cancer samples.** (A) Analysis of allele frequency heterogeneity in subregions of specimen 2182. Forty-nine allelic variants were detected in this tumor. Each allele frequency was normalized to the average of frequencies in the different regions, log<sub>2</sub> transformed, and plotted relative to the average. Red bars indicate frequencies greater than the mean, and green bars indicate frequencies lower than the mean for each regions. The dotted line represents the axis (equal to zero, or no difference from the mean) for each region and allele. For clarity, alleles that are unique to a region are shown as maximal red bars, and no green bars. Note that most of these variants are novel alleles and currently have no defined clinical relevance. **Alleles that overlap in the Ion AmpliSeq custom lung cancer panel and OncoPrint Comprehensive Assay.** Highlighted loci were detected in one region and not in others. (B) Analysis of allele frequency heterogeneity in subregions of sample 2162. Seventy-two different variant alleles were found in this tumor specimen. Frequencies are illustrated using the method described above. Red italicized text highlights a pathogenic allele detected only by LCM.

In the later-staged 2162 tumor, which presumably had more opportunity to accumulate new mutations, the heterogeneity was more pronounced. Twenty-five of the 72 variants detected (34.7%) were found exclusively in one of the regions, and not detected at all in the WTS. Interestingly, the heterogeneity was evident even in regions that had a similar overall morphology and were close to each other (Figure 5A, regions 2 and 3), suggesting independent clonal histories. The unique alleles also differed in their abundance. For example, the *ATM* chr11.08151708 ins T allele found in region 4 was present at a frequency of 25.2%, whereas

*NOTCH1* chr9.139409952 C>T in the same region was present at a frequency of only 6.5%. Although most of these variants are alleles that are novel and have as of yet no defined pathological relevance, one clinically significant allele of *PIK3R1* (chr5.67591106 A>G, p.K567E) was found only in region 3 (Figure 5B), and not in the WTS or any of the other regions. This allele has been identified in other tumor types, and it may confer additional growth advantages to cells harboring this mutation. This allele would have been missed if analysis was confined to macrodissected regions alone.

These data indicated that clinical research-relevant and actionable alleles may only be detected by examining subregions of a tumor. Although rare mutations were more easily identified in the LCM-captured subregions, alleles that were not as rare were also detected only in LCM specimens, and not in WTS specimens. A list of alleles found that were unique for each subregion is given in Table 1. To further illustrate the power of LCM in identifying new alleles, we determined the fraction of unique variants detected in the WTS, and the fraction of unique variants found in the microdissected regions (Figure 6). Of the 49 variants detected in the 2182 sample, 8% were found only in the WTS, whereas 22.5% were detectable only in the LCM regions. Of the 72 variants detected in the 2162 sample, less than 1.5% were found in the WTS, but 36% were found in the LCM regions. These results demonstrate that in order to have the best chance of uncovering mutant alleles in a complex tumor, smaller subregions of the tumor should be analyzed independently. Such subregions can easily be collected by LCM.

We also examined the heterogeneity in the tumor sections by analyzing levels of 53 transcripts in the different regions and WTS. To facilitate comparisons of transcript levels, the data were normalized to mapped reads of each transcript per million reads (RPM). In sample 2182, the transcript expression pattern of the WTS was close to the median for each gene across all the regions (Figure 7A). This is expected, as the whole tissue scrape should reflect the contribution of all cells contained in the scrape, including the regions. Interestingly, some expression heterogeneity is evident.



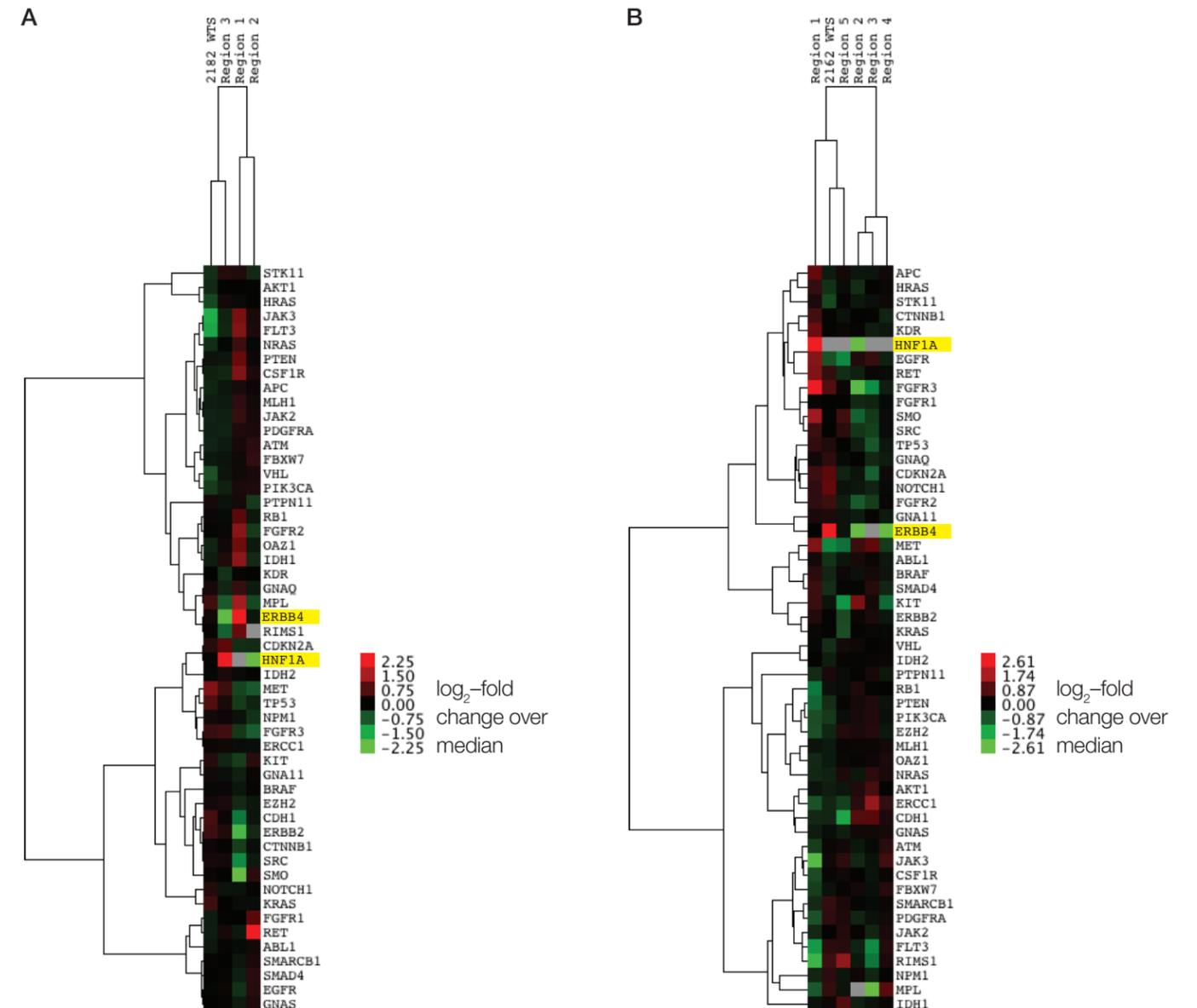
**Figure 6. Fraction of unique alleles detected in WTS specimens or LCM regions.** Analyzing individual LCM regions identified more unique alleles than the WTS. Notably, there were alleles identified in the WTS that were not detected in any of the regions. This might reflect the fact that not all of the tumor area was sampled by LCM.

**Table 1. Frequencies of alleles detected only in microdissected regions.**

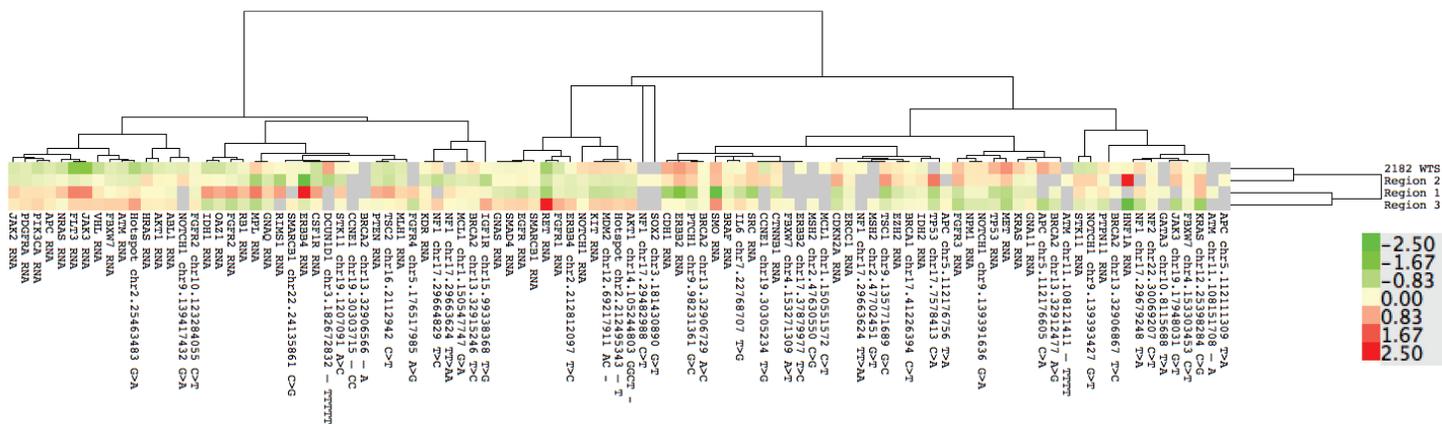
Sample	Region	Allele	Frequency (%)
2162	1	<i>TSC2</i> chr16.2112942 C>T	22.9
2162	1	<i>FGFR1</i> chr8.38293012 C>T	6.8
2162	1	<i>STK11</i> chr19.1207091 A>C	6.8
2162	1	Hotspot chr2.25467508 T>C	5.2
2162	2	<i>KDR</i> chr4.55980239 C>T	29.5
2162	2	<i>ERBB2</i> chr17.37850777 C>T	23.5
2162	2	<i>GAS6</i> chr13.114559535 T>C	9.8
2162	2	<i>MET</i> chr7.116322791 A>C	9.8
2162	2	<i>JAK3</i> chr19.17948031 G>T	8.7
2162	2	<i>PNP</i> chr14.20941309 ins T	7.3
2162	2	<i>RB1</i> chr13.49033919 C>G	6.1
2162	3	<i>PIK3R1</i> chr5.67591106 A>G	13.4
2162	3	<i>ERBB4</i> chr2.212578380 ins AAAAA	18.3
2162	3	<i>NPM1</i> chr5.170837514 ins TTTTTT	16.9
2162	3	<i>NF1</i> chr17.29679246 G>A	14.2
2162	3	<i>CCND1</i> chr11.69458923 A>G	7.0
2162	4	<i>ATM</i> chr11.108151708 ins T	31.6
2162	4	<i>NOTCH1</i> chr9.139409952 C>T	6.5
2162	5	<i>ATM</i> chr11.108151708 T>A	25.2
2162	5	<i>RB1</i> chr13.49051481 T>A	12.4
2162	5	<i>FBXW7</i> chr4.153268242 G>A	5.5
2162	5	<i>IDH1</i> chr2.209113193 C>A	5.4
2162	5	<i>NOTCH1</i> chr9.139417356 CG>GT	5.4
2162	5	<i>RB1</i> chr13.48954247 ins AT	5.2
2182	1	<i>ATM</i> chr11.108151708 ins A	24
2182	1	<i>GATA3</i> chr10.8115688 T>A	7.4
2182	2	<i>SOX2</i> chr3.181430890 G>T	11.5
2182	2	<i>NF1</i> chr17.29482988 C>T	9.7
2182	2	<i>APC</i> chr5.112111309 T>A	9.2
2182	3	<i>MSH2</i> chr2.47630550 C>G	29.4
2182	3	<i>BRCA2</i> chr13.32906566 ins A	10.9
2182	3	<i>FBXW7</i> chr4.153271309 A>T	7.4
2182	3	<i>MSH2</i> chr2.47702451 G>T	6.7

For example, *ERBB4* and *HNF1A* show marked differences in transcript abundance between regions 1 and 3, and *RET* is clearly more abundant in region 2 than in the other regions. Similarly, heterogeneity in transcript abundance is evident in sample 2162 (Figure 7B). However, there appear to be more differences in levels than sample 2182, perhaps reflecting the fact that this is a later-staged tumor with more opportunity for cellular divergence. In this sample, *HNF1A* expression was

not detected at all in the WTS or in regions 3, 4, or 5, but was present in regions 1 and 2. In addition, *ERBB4* demonstrated relatively higher levels than the other regions, suggesting there were cells present that expressed high levels of *ERBB4* that were not collected in the LCM regions. Thus, analysis of RNA expression patterns in different regions of the tumors reflects another layer of tumor heterogeneity.



**Figure 7. Gene expression analysis of lung cancer samples.** (A) Transcript abundances in the different regions of the 2182 sample. Although the WTS is a reflection of transcript abundances in all regions, some transcripts are differently expressed in the WTS when compared to the different regions. (B) Transcript abundances in the 2162 sample. In this sample, heterogeneity in the different regions is more pronounced. Grey = transcript not detected in that region.



**Figure 8. Co-analysis of variant allele frequencies and transcript abundances in the 2182 tissue sample.** Subsets of the variants and transcripts show similar patterns of distribution. This might reflect a common history of cells that contain those patterns. Grey = transcript or variant not detected in that region.

Finally, to obtain a complete snapshot of the heterogeneity of these tumors, we combined the variant allele frequencies and transcript abundances into a single analysis to search for patterns that tended to appear together. An example using the data from sample 2182 is shown (Figure 8). In spite of the fact that regions 1 and 2 had similar gross cellular morphologies, the pattern of allele frequencies and transcript levels suggest region 3 is more similar to region 1 than region 2. Additionally, certain subsets of variants and transcripts appear to be present in similar patterns, suggesting these cells might have a similar clonal history or epigenetic modifications. For example, the *MDM2* chr12.69217911 ins AC and *Hotspot* chr2.212495343 ins T variants as well as *NOTCH1* and *KIT* transcripts show similar overrepresentation in the WTS and region 3 relative to regions 1 and 2. Although the significance of these results is affected by the low number of samples analyzed in this study, these data illustrate that a tumor mass is more than a collection of identical cells. Further insights into the history and heterogeneity of these tumors could be made by examining more regions, more alleles, and more transcripts.

## Conclusions

In this application note, we showed that the ArcturusXT LCM System can reveal the presence of variant alleles that cannot be detected in whole tissue scrapes. In addition, we showed that the Ion AmpliSeq™ workflow for analyzing DNA variants and RNA transcript levels allows the efficient analysis of many sequences from small amounts of FFPE tissue-extracted starting material. Finally, we demonstrated that because tumors are heterogeneous, efficient cataloging of the different variants in a tumor requires sampling several different regions of the tumor.

These data suggest that a combination of mutation detection and RNA expression analysis could reveal clones of cells with similar histories or epigenetic modifications. Such clones may not be evident by confining analyses to sequences obtained from a preparation of a whole tumor mass. Some clones of cells might arise due to the activities of specific oncogenes or other targetable molecules affecting entire pathways. By identifying these clones from a heterogeneous mixture, it may be possible to design interventions targeted to specific cells. Therefore, understanding tumor heterogeneity is extremely important, as it has been shown to affect responses to molecularly targeted treatments of cancers [6].

## References

1. Fidler IJ, Kripke ML (1977) Metastasis results from preexisting variant cells within a malignant tumor. *Science* 197:893–895.
2. Katona TM, Jones TD, Wang M et al. (2007) Genetically heterogeneous and clonally unrelated metastases may arise in patients with cutaneous melanoma. *Am J Surg Pathol* 31:1029–1037.
3. Liegl B, Kepten I, Le C et al. (2008) Heterogeneity of kinase inhibitor resistance mechanisms in GIST. *J Pathol* 216:64–74.
4. Maley CC, Galipeau PC, Finley JC, et al. (2006) Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat Genet* 38:468–473.
5. Taniguchi K, Okami J, Kodama K et al. (2008) Intratumor heterogeneity of epidermal growth factor receptor mutations in lung cancer and its correlation to the response to gefitinib. *Cancer Sci* 99:929–935.
6. Yancovitz M, Litterman A, Yoon J et al. (2012) Intra- and inter-tumor heterogeneity of *BRAF*<sup>V600E</sup> mutations in primary and metastatic melanoma. *PLoS ONE* 7:e29336.

Find out more at [thermofisher.com/lcm](http://thermofisher.com/lcm)

**ThermoFisher**  
SCIENTIFIC