Tumor Heterogeneity Predicts Metastatic Potential in Colorectal Cancer

Je-Gun Joung1, Bo Young Oh2, Hye Kyung Hong3, Hisham Al-Khalidi4, Faisal Al-Alem5, Hae-Ock Lee1, Joon Seol Bae1, Jinho Kim1, Hong-Ui Cha1, Maram Alotaibi6, Yong Beom Cho3,6, Mazen Hassanain5, Woong-Yang Park1,6,7, and Woo Yong Lee3

Abstract

Purpose: Tumors continuously evolve to maintain growth; secondary mutations facilitate this process, resulting in high tumor heterogeneity. In this study, we compared mutations in paired primary and metastatic colorectal cancer tumor samples to determine whether tumor heterogeneity can predict tumor metastasis.

Experimental Design: Somatic variations in 46 pairs of matched primary-liver metastatic tumors and 42 primary tumors without metastasis were analyzed by whole-exome sequencing. Tumor clonality was estimated from single-nucleotide and copy-number variations. The correlation between clinical parameters of patients and clonal heterogeneity in liver metastasis was evaluated.

Results: Tumor heterogeneity across colorectal cancer samples was highly variable; however, a high degree of tumor heterogeneity was associated with a worse disease-free survival. Highly heterogeneous primary colorectal cancer was correlated with a higher rate of liver metastasis. Recurrent somatic mutations in APC, TP53, and KRAS were frequently detected in highly heterogeneous colorectal cancer. The variant allele frequency of these mutations was high, while somatic mutations in other genes such as PIK3CA and NOTCH1 were low. The number and distribution of primary colorectal cancer subclones were preserved in metastatic tumors.

Conclusions: Heterogeneity of primary colorectal cancer tumors can predict the potential for liver metastasis and thus, clinical outcome of patients. Clin Cancer Res; 23(23). 7209-16. ©2017 AACR.

Introduction

Colorectal cancer is one of the most commonly diagnosed cancers worldwide and a leading cause of cancer-related death (1, 2). Early detection by endoscopic examination and more effective chemotherapy have led to increased survival of colorectal cancer patients, although the mortality rate remains high in cases of metastatic disease and/or recurrence (2). Synchronous metastases are detected in 20% to 25% of patients at the time of diagnosis; in addition, approximately 50% develop metastases during the disease course (3). Metastatic progression is a multistep process involving phenotypic changes in primary tumor cells as a result of genetic and/or epigenetic alterations that facilitate dissemination and tissue invasion of tumor cells (4, 5). Many proto-oncogenes and tumor suppressor genes associated with colorectal cancer have been identified; however, little is known regarding the molecular events responsible for metastatic progression of the disease.

Tumor heterogeneity, the concept that a single tumor consists of many tumor cell subclones, has become an important topic in cancer genomics (6). It is hypothesized to play a critical role in the progression of many cancer types and is a major obstacle to precision cancer therapy. During this process, subclones continuously arise via genomic mutation. The presence of subclones has been shown to adversely affect outcome in chronic lymphocytic leukemia (7), head and neck cancer (8), and lung adenocarcinoma (9). The so-called Big Bang model of tumor progression postulates that malignant potential is mostly determined early in tumor development and is nonselective (10). However, the full complement of factors that lead to tumor heterogeneity during colorectal cancer progression is unknown.

Next-generation sequencing technology has recently been used in clinical applications to identify mutations and copy-number variations in oncogenes. Specifically, whole-exome sequencing (WES) has revealed informative mutations within exome coding regions. On the basis of WES data, subclones of tumor cell populations can be inferred by computational methods such as PyClone (11), SciClone (12), and EXPANDS (13). Clonal heterogeneity analysis may be useful for predicting patient prognosis and response to therapy. Early clonal events are potential therapeutic targets as they represent the tumor cell population, whereas...

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Corresponding Authors: Woong-Yang Park, Samsung Medical Center, 811woon-ro, Gangnam-gu, Seoul 06351, Korea. Phone: 822-2148-9810; Fax: 822-2148-9819; E-mail: woongyang.park@samsung.com; Mazen Hassanain, King Saud University, Riyadh 11472, Saudi Arabia. Phone: 966-50-514-1090; E-mail: mhassanain@ksu.edu.sa; and Woo Yong Lee, wooyong123.lee@samsung.com

doi: 10.1158/1078-0432.CCR-17-0306
©2017 American Association for Cancer Research.
Translational Relevance

Understanding the genetic heterogeneity of tumors is important for predicting patient treatment responses and disease progression. In colorectal cancer, tumor heterogeneity may be associated with worse prognosis and response to therapy. In this study, we demonstrate that a high degree of tumor heterogeneity is correlated with poor clinical outcome in colorectal cancer patients. Highly heterogeneous primary colorectal cancer was more likely to lead to liver metastasis and was linked to vascular invasion and the occurrence of colorectal cancer–associated somatic mutations. Interestingly, the subclonal pattern of primary colorectal cancer was largely preserved in metastatic tumors. Our results suggest that colorectal cancer liver metastasis develops through the collective spread of multiple clonal subpopulations in parallel. Thus, tumor heterogeneity plays a critical role in the metastatic progression of colorectal cancer and may serve as a marker for predicting clinical outcome.

Materials and Methods

Patients

This study was approved by the Institutional Review Boards of Samsung Medical Center (SMC; IRB no. 2010-04-004) and King Saud University (KSU; IRB no. E-12-592). All procedures were conducted in accordance with the Declaration of Helsinki. Written, informed consent was obtained from all enrolled patients.

The study population included 88 colorectal cancer patients consisting of 42 stage I–II and 46 stage IV patients with synchronous liver metastasis. All had histologically confirmed primary colorectal adenocarcinoma and had undergone curative surgical resection for the primary tumor and metastatic lesions at the SMC and KSU from 2004 to 2014. Patients were excluded if they had recurrent disease, local excision, palliative surgery, or previous treatment before colonic surgery.

A total of 222 samples from 88 patients were obtained; the primary and metastatic tumor and blood (or normal colon) samples were obtained from metastatic patients, and primary tumor and blood or normal colon samples were obtained from nonmetastatic patients.

Whole-exome sequencing

Genomic DNA was extracted from fresh frozen and formalin-fixed, paraffin-embedded (FFPE) tissue using the QiAamp DNA mini and a DNA FFPE Tissue kits (Qiagen). The tissue was microdissected. Because paraffin fixation chemically modifies genomic (g)DNA and thereby reduces its quality, we applied the following quality control thresholds: (i) purity: 260/280 > 1.8 and 260/230 > 1.8; (ii) total amount > 250 ng; and (iii) degradation (FFPE): ΔC, value < 2.0 or DNA median size > 0.35 kb. The highest quality gDNA in each sample was sheared with an S220 ultrasonicator (Covaris) and used to construct a library with the SureSelect XT Human All Exon v5 and SureSelect XT Reagent Kit, HSQ (Agilent Technologies) according to the manufacturer's protocol. This kit is designed to enrich 335,756 exons of 21,058 genes, covering approximately 71 Mb of the human genome. Enriched exome libraries were multiplexed and sequenced on the HiSeq 2500 platform (Illumina). Briefly, a paired-end DNA sequencing library was prepared through gDNA shearing, end-repair, A-tailing, paired-end adaptor ligation, and amplification. After hybridizing the library with bait sequences for 16 hours, the captured library was purified and amplified with an indexing barcode tag and library quality and quantity were assessed. The exome library was sequenced using the 10-bp paired-end mode of the TruSeq Rapid PE Cluster and TruSeq Rapid SBS kits (Illumina).

Exome sequence data analysis

Sequencing reads were aligned to the University of California Santa Cruz hg19 reference genome (downloaded from http://genome.ucsc.edu) using the Burrows–Wheeler Aligner v. 0.6.2 (15) with default settings. PCR duplications were marked using Picard-tools-1.8 (http://picard.sourceforge.net/) and data cleanup was achieved using GATK-2.2.9 (16). Point mutations were identified with the MuTect tool (https://github.com/broadinstitute/mutect) in paired samples. Annovar was used to annotate variants. Signature analysis of mutational processes was carried out using the deconstructSigs tool (17). Hypermutated samples with more than 1,000 mutations were excluded as they introduced bias. Copy-number variations were detected using EXCA-VATOR software (18).

Clonality analysis

Subclones were obtained using clustering cancer cell fractions (CCF) by PyClone, which deconvolves tumor sequences into subclones based on a hierarchical Bayesian clustering model (11). Input data were generated from somatic single-nucleotide variants (SNV) detected by MutTect and corresponding copy number variations. SNVs were clustered according to similar CCF, after which each set of clustered mutations (a subclone) of cancer cells was identified. Step filtering of SNVs with large, credible intervals from the previous stage was applied to remove noninformative posterior distributions. Each tumor was classified into two types of clonality based on the number of subclones they possessed. We defined a tumor as an oligo-clone when there were one or two subclones; otherwise, the tumor was defined as a multiclonal. A clonal evolution model for metastatic colorectal cancers was generated using ClonEvol R (https://github.com/hdng/clonevol).

Measurement of heterogeneity

Heterogeneity was measured by mutant-allele tumor heterogeneity (MATH) analysis, which was initially developed to measure intratumor heterogeneity in head and neck cancer samples (8). Mutation allele frequencies (MAF) of mutated loci in each tumor were determined; the center and width of the distribution of MAFs among these loci were obtained, and the ratio of the
width to the center of the distribution was calculated as median absolute deviation/median MAF.

Assessment of clinical outcomes

Patients were categorized according to clonality and groups were compared with respect to clinicopathologic features. To investigate the clinical relevance of clonality, we analyzed disease-free survival (DFS) according to clonality. Recurrence was established by biopsy following colonoscopy or based on imaging findings typical of cancer recurrence, and DFS was defined as the period from surgery to recurrence. The primary endpoint of this study was the degree of clonality of each tumor, and the secondary endpoint was survival outcome based on clonality.

Statistical analysis

Statistical analyses were performed using R v.3.1.2 software (https://www.r-project.org/). Categorical variables were compared using the χ² test. Survival rates were analyzed with the Kaplan–Meier method and log-rank test. Multivariate analysis was carried out using a Cox proportional hazard model to identify prognostic factors. P values were derived from two-tailed tests and those less than 0.05 were considered statistically significant.

Results

Identification of somatic mutations in colorectal cancers

A total of 88 patients with colorectal cancer, including 46 with and 42 without liver metastasis (mCRC and nmCRC, respectively) were evaluated. We performed WES using 46 matched pairs [primary colon tumor, liver metastatic tumor, and blood (or normal colon) samples] from the mCRC group and 42 matched pairs [primary colon tumor and blood (or normal colon) samples] from the nmCRC group (Supplementary Table S1). Patient characteristics according to clonality are shown in Table 1; there were no differences between groups for most variables.

WES of the 222 samples yielded 33.7 billion reads. The average target depth for the analysis was 165.6 (s.d. 173.9) × (s.d. 5 ng/mL). Whole-exome coverage was at least 99.4% (Supplementary Table S1). The depth of tumors and normal samples was 173.9 (s.d. 49.9) and 152.8 × (s.d. 49.6), respectively.

We identified numerous somatic mutations, including missense, nonsense, and splicing mutations in primary colon tumors obtained from mCRC and nmCRC groups (Fig. 1). These included several somatic mutations that frequently occur in colorectal cancers, with the most common and well-described (19) identified at particularly high frequencies, including adenomatous polyposis coli (APC, 54.6%), KRAS (42.0%), and TP53 (55.7%), consistent with previous reports (20). We also investigated whether there was a concordance of mutations between primary colon tumors and matched liver metastatic tumors in the mCRC group (Fig. 1). Most mutations were detected in both primary tumors and metastatic lesions, with high rates of concordance in APC (77.8%, 21/27 concordant mutations in primary tumors), KRAS (88.2%, 15/17), and TP53 (87.5%, 21/24); this is consistent with previous studies of primary metastasis in colorectal cancer (21). Other colorectal cancer–associated genes such as phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA), mothers against decapentaplegic homolog 4 (SMAD4), NRAS, catenin β-1 (CTNNB1), AKT1, NOTCH1, and anaplastic lymphoma kinase (ALK) were less frequently mutated.

Table 1. Demographic and clinical landscape of 88 colorectal cancer patients

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Oligoclone (n = 27)</th>
<th>Multi-clone (n = 61)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (y)</td>
<td>54.8</td>
<td>57.1</td>
<td>0.454</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td>0.846</td>
</tr>
<tr>
<td>Male</td>
<td>14 (51.9%)</td>
<td>33 (54.1%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>13 (48.1%)</td>
<td>28 (45.9%)</td>
<td></td>
</tr>
<tr>
<td>Preoperative CEA level, n (%)&lt;5 ng/mL</td>
<td>9 (33.3%)</td>
<td>28 (45.9%)</td>
<td>0.125</td>
</tr>
<tr>
<td>≥5 ng/mL</td>
<td>9 (33.3%)</td>
<td>22 (36.1%)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>9 (33.3%)</td>
<td>11 (18.0%)</td>
<td></td>
</tr>
<tr>
<td>Location of primary tumor, n (%)</td>
<td></td>
<td></td>
<td>0.940</td>
</tr>
<tr>
<td>Colon</td>
<td>21 (77.8%)</td>
<td>47 (77.0%)</td>
<td></td>
</tr>
<tr>
<td>Rectum</td>
<td>6 (22.2%)</td>
<td>14 (23.0%)</td>
<td></td>
</tr>
<tr>
<td>Cell type, n (%)</td>
<td></td>
<td></td>
<td>0.893</td>
</tr>
<tr>
<td>WD/MC</td>
<td>20 (74.1%)</td>
<td>46 (75.4%)</td>
<td></td>
</tr>
<tr>
<td>PD/MUC/SRC</td>
<td>7 (25.9%)</td>
<td>15 (24.6%)</td>
<td></td>
</tr>
<tr>
<td>Stage at diagnosis, n (%)</td>
<td></td>
<td></td>
<td>0.046</td>
</tr>
<tr>
<td>I</td>
<td>1 (3.7%)</td>
<td>2 (3.3%)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>7 (25.9%)</td>
<td>14 (23.0%)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>10 (37.0%)</td>
<td>8 (13.1%)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>9 (33.3%)</td>
<td>37 (60.7%)</td>
<td></td>
</tr>
<tr>
<td>Vascular invasion, n (%)</td>
<td></td>
<td></td>
<td>0.015</td>
</tr>
<tr>
<td>Yes</td>
<td>2 (7.4%)</td>
<td>20 (32.8%)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>25 (92.6%)</td>
<td>41 (67.2%)</td>
<td></td>
</tr>
<tr>
<td>Lymphatic invasion, n (%)</td>
<td></td>
<td></td>
<td>0.930</td>
</tr>
<tr>
<td>Yes</td>
<td>10 (37.0%)</td>
<td>22 (36.1%)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>17 (63.0%)</td>
<td>39 (63.9%)</td>
<td></td>
</tr>
<tr>
<td>Perineural invasion, n (%)</td>
<td></td>
<td></td>
<td>0.351</td>
</tr>
<tr>
<td>Yes</td>
<td>7 (25.9%)</td>
<td>22 (36.1%)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>20 (74.1%)</td>
<td>39 (63.9%)</td>
<td></td>
</tr>
<tr>
<td>Adjuvant treatment, n (%)</td>
<td></td>
<td></td>
<td>0.520</td>
</tr>
<tr>
<td>Yes</td>
<td>24 (88.9%)</td>
<td>51 (83.6%)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>3 (11.1%)</td>
<td>10 (16.4%)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CEA, carcinoembryonic antigen; MD, moderately differentiated; MUC, mucinous carcinoma; PD, poorly differentiated; SRC, signet ring cell carcinoma; WD, well differentiated.

Correlation between clonality and clinical features

We investigated whether there was an association between clonality and primary tumor genomics or related clinical features. Clonality was determined on the basis of subclones in each tumor. The number of subclones was positively correlated with tumor heterogeneity, as determined by MATH analysis (Fig. 2A). This indicates that tumors with a larger number of subclones exhibit a higher degree of heterogeneity.

We next performed a survival analysis to evaluate the clinical relevance of clonality, and found a significant association between DFS and number of subclones (Fig. 2B). Colorectal cancer patients with multi-clone tumors showed worse clinical outcome relative to those with oligo-clone tumors (P = 0.04). Patients with a high tumor heterogeneity score showed a similar result (P = 0.009; Supplementary Fig. S1). Tumor recurrence rate also differed between oligo- and multiclonal groups (P < 0.012; χ² test; Fig. 2C). Multivariate analyses were carried out using a Cox proportional hazard model to identify prognostic factors for DFS; these included variables that were significant in the univariate analysis. The results showed that the presence of multiple clones was an independent prognostic factor for reduced DFS (Table 2).

Mutations in the three major colorectal cancer–associated genes, that is, APC, KRAS, and TP53, were more common; however, the P values are all >0.05 (Fig. 2D). Identified mutations are shown in Fig. 2E. Other clinical features, including preoperative carcinoembryonic antigen level, cell type, tumor stage at diagnosis, vascular invasion, lymphatic invasion, and perineural...
invasion were also compared with the number of subclones (1, 2, 3, 4, and $\geq$5). Among the assessed criteria, progression to an advanced stage and presence of vascular invasion were significantly associated with multiclonal tumors ($P = 0.046$ and 0.015, respectively; $\chi^2$ test; Table 1).

Characteristics of somatic mutations according to clonality

We categorized somatic mutations according to their order of CCFs to determine the relative timing of mutation acquisition. Mutations were classified into four categories based on CCF distribution (G1, 1st; G2, 2nd; G3, 3rd; and G4, 4th quartile; Fig. 3A). G1 included events that occurred earlier than those in other categories. Mutations in TP53, APC, KRAS, SMAD4, NRAS, CTNNB1, and F-box and WD repeat domain-containing 7 (FBXW7), which are considered as early events, belonged to the high-CCF group. Interestingly, these mutations occurred predominantly in multiclonal samples. For example, 71% of APC mutations belonged to the G1 and G2 groups (Fig. 3B). This supports the hypothesis that colorectal tumorigenesis is initiated by a mutation in APC. In multiclonal samples, 74.2% of APC mutations were detected in the G1 and G2 groups as compared with 57.1% in these groups in oligoclonal samples. This suggests that the occurrence of mutations decreased substantially during expansion into multiple subclonal populations. KRAS and TP53 showed similar patterns in the relative timing of mutation acquisition (Supplementary Fig. S2). In contrast, mutations in PIK3CA and NOTCH1 belonged to a low-CCF group, indicating that they were later events.

We then carried out a signature analysis of the mutation process. Overall mutation patterns were similar to those of Signature 1, which is observed in most cancer types (22). We compared the substitution spectra of oligo- and multiclonal (Fig. 3C) and found that there were no differences in the patterns of nucleotide changes between oligo- and multiclonal samples. However, when the signatures were decomposed, an additional one (Signature 19) was observed in a small number of oligoclonal samples (Supplementary Fig. S3), although its etiology remains unclear.

Changes in clonality from primary to metastatic tumors

We investigated quantitative changes (i.e., differences in the number of subclones) in clonality during the progression from primary tumor to liver metastatic lesion. There are four possible transitions of oligo- versus multiclones during this process (CRC-OO, oligo $\rightarrow$ oligo; CRC-OM, oligo $\rightarrow$ multi; CRC-MM, multi $\rightarrow$ multi; Fig. 4A; Supplementary Table S2). The CRC-MM transition (71.7%, 33 of 46 patients with metastatic colorectal cancer) was predominant, indicating that the number of subclones was highly consistent between multiclonal primary tumors and liver metastatic lesions. A combination of APC, KRAS, and TP53 mutations was associated with both multiclonal primary and metastatic samples. For example, mCRC-35 showed a CRC-MM transition type (Fig. 4B); the patient harbored mutations in KRAS, TP53, CTNNB1, and APC, which may be required for clonal expansion. Interestingly, only APC and TP53 mutations and not KRAS mutations were associated with CRC-OO transitions. Nonetheless, we speculate that KRAS mutations have an important role in clonal expansion.

Discussion

In this study, a comprehensive set of somatic mutations and subclones of individual primary and hepatic metastatic tumors from 88 colorectal cancer patients was identified by WES. We observed that tumor heterogeneity was significantly associated with survival outcome in colorectal cancer. Primary tumors with high heterogeneity were more likely to spread to the liver from the primary tumor site, suggesting an association with vascular invasion. It is highly likely that tumor cells in subclones spread together rather than individually to metastatic lesions. Six tumors in the mCRC group were hypermutated, but these did not always display high heterogeneity. For example, a Mutt homolog 6 (MSH6) mutation was oligoclonal in both primary and metastatic lesions (Fig. 4B).

Mutations frequently occurring in colorectal cancer such as those in APC, TP53, and KRAS were more common in colorectal
cancer patients with high as compared with low tumor heterogeneity and were considered as early clonal events in high-CCF groups in multiclonal samples. These mutations would likely accumulate before clonal expansion into multiple subclonal populations. On the other hand, mutations in \textit{PIK3CA} and \textit{NOTCH1} in low CCFs occurred later through continuous clonal expansion, which is consistent with the Big Bang model of tumor progression (10). Thus, a rapidly growing clone within small tumor may arise from several early key mutations, with subsequent mutations improving the fitness of the local clonal expansion of tumor cell subpopulations. Somatic mutations in \textit{PIK3CA} are late events during tumorigenesis (23, 24), and those in \textit{NOTCH1} are secondary events during T-cell acute lymphoblastic leukemia (25). The low allele frequencies in \textit{NOTCH1} observed in this study support the notion that these mutations occur later. However, the effects of \textit{NOTCH1} mutations in colorectal cancer and other cancer types remain to be determined.

There was no relationship between mutational signatures and the number of subclones, indicating that changes in mutational composition do not determine the number of subclones during tumor progression. Although most of our analyses were not affected by hypermutated samples, mutation spectrum could be highly affected by hypermutation. Therefore, hypermutators were excluded from the mutation spectrum analysis as the patterns could reflect hypermutations rather than tumor heterogeneity. Previous studies of lung cancer reported that the spectrum of mutations caused by transformations associated with \textit{apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC)} cytidine deaminase differed between the trunk (early events) and non-trunk (late events; ref. 26). Additional studies are needed to identify the changes in the mutation spectrum that occur during colorectal cancer progression.

This study has a limitation that stage was differently distributed between groups according to clonality. Multiclonal patients had...
more metastatic disease at diagnosis than oligoclonal patients (60.7% vs. 33.3%). This may affect our result that the risk of relapse was higher in multiclonal patients than oligoclonal patients. Thus, multivariate analysis using a Cox proportional hazard model was performed to overcome the confounding bias and we identified that multiclone was an independent poor

Table 2. Multivariate analysis for disease-free survival in 88 colorectal cancer patients

<table>
<thead>
<tr>
<th>Variables</th>
<th>Univariate</th>
<th>Multivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clonality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi vs. oligo</td>
<td>P = 0.041</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Age (years) ≥65 vs. &lt;65</td>
<td>0.585</td>
<td>3.085 (1.035–9.184)</td>
</tr>
<tr>
<td>Gender female vs. male</td>
<td>0.622</td>
<td></td>
</tr>
<tr>
<td>CEA level (ng/mL) ≥ 5 vs. &lt; 5</td>
<td>&lt;0.001</td>
<td>2.345 (0.858–6.393)</td>
</tr>
<tr>
<td>Location of tumor rectum vs. colon</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Stage III vs. I-II</td>
<td>0.404</td>
<td>1.639 (0.390–6.887)</td>
</tr>
<tr>
<td>Stage IV vs. I-II</td>
<td>&lt;0.001</td>
<td>6.006 (1.659–21.743)</td>
</tr>
<tr>
<td>Cell type PD/MUC/SRC vs. WD/MD</td>
<td>0.303</td>
<td></td>
</tr>
<tr>
<td>Vascular invasion yes vs. no</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>Lymphatic invasion yes vs. no</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>Perineural invasion yes vs. no</td>
<td>&lt;0.001</td>
<td>5.736 (2.443–13.468)</td>
</tr>
<tr>
<td>Adjuvant treatment yes vs. no</td>
<td>0.105</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.

Characteristics of mutations according to clonality. A, Heatmap of mutational frequency in four categories based on CCF distribution (G1, 1st; G2, 2nd; G3, 3rd; and G4, 4th quantile). G1 (or G4) corresponds to an early (or late) event. The size of each circle corresponds to the number of mutations. B, Circle diagram showing the number of APC mutations in each of the four categories (G1–G4). C, Spectrum of substitutions in oligo- and multiclonal. Each bar represents the proportion of each type of substitution.
prognostic factor for survival. Despite this limitation, our study had the strength of integrating tumor heterogeneity into clinical applications. Understanding tumor heterogeneity is essential for developing personalized treatment for patients with cancers, including colorectal cancer. Tumor heterogeneity resulting from the continuous accumulation of mutations during tumor progression determines tumor characteristics, influencing the clinical outcome of patients. Tumor heterogeneity, defined as the genetic diversity that exists within individual tumors, is one of the most challenging current issues in the field of cancer biology. Tumor heterogeneity resulting from continuous accumulation of mutations can aid in predicting response and resistance to drugs and dictate clinical outcome. In this study, there was a significant association between the rate of primary colorectal cancer recurrence and number of subclones, suggesting that recurrence is also related to high tumor heterogeneity. Our findings provide a basis for the development of personalized treatments for patients with colorectal cancer and other cancers.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: B.Y. Oh, M. Hassanain, W.-Y. Park, W.Y. Lee
Development of methodology: J.-G. Joung, B.Y. Oh

Figure 4.
Changes in clonality from primary to metastatic tumor. A, Four possible cases of clonality transition for oligo- versus multiclones and primary versus metastatic tumors (CRC-OO, oligo → oligo; CRC-OM, oligo → multi; CRC-MO, multi → oligo; and CRC-MM, multi → multi). Each triangle indicates one of APC, KRAS, and TP53 mutations, and indicates that none of them are found. Small arrow and triangle represented which mutations appeared between primary and metastatic tumor. (4) indicates the number of patients with the corresponding pattern. Large arrows indicate metastasis and their thickness are represented according to the number of patients with a transition type. B, Inferred clonal evolution in metastatic colorectal cancer. Each area indicates a subpopulation containing a set of mutations. Cancer-associated genes are shown. A tree depicts the evolution of tumors. P# and M# indicate a subclone in primary and metastatic tumors, respectively.

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H.K. Hong, H. Al-Khalidi, F. Al-Alem, Y.B. Cho, M. Hassanain
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J.-G. Joung, B.Y. Oh, J. Kim, H.-U. Cha, M. Hassanain
Writing, review, and/or revision of the manuscript: J.-G. Joung, B.Y. Oh, H. Al-Khalidi, F. Al-Alem, H.-O. Lee, M. Hassanain, W.-Y. Park
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H.K. Hong, H. Al-Khalidi, F. Al-Alem, J.S. Bae, Y.B. Cho, W.-Y. Park, W.Y. Lee
Study supervision: M. Alotaibi, Y.B. Cho, W.-Y. Park, W.Y. Lee

Acknowledgments
We thank Seok Hyung Kim and Kyoung-Mee Kim (Department of Pathology, Samsung Medical Center) for pathological analyses. Biospecimens from Samsung Medical Center were provided by SMC BioBank. All sequencing was performed by Samsung Genome Institute.

Grant support
The study was supported in part by the Korea Health Technology R&D Project (HI13C2096; to W.Y. Park) through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Korea.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 1, 2017; revised May 13, 2017; accepted September 18, 2017; published OnlineFirst September 22, 2017.
References
Tumor Heterogeneity Predicts Metastatic Potential in Colorectal Cancer

Je-Gun Joung, Bo Young Oh, Hye Kyung Hong, et al.


Updated version
Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-17-0306

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2017/09/22/1078-0432.CCR-17-0306.DC1

Cited articles
This article cites 25 articles, 5 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/23/23/7209.full#ref-list-1

Citing articles
This article has been cited by 4 HighWire-hosted articles. Access the articles at:
http://clincancerres.aacrjournals.org/content/23/23/7209.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link http://clincancerres.aacrjournals.org/content/23/23/7209.
Click on "Request Permissions" which will take you to the Copyright Clearance Center’s (CCC) Rightslink site.