Complex Tumor Genomes Inferred from Single Circulating Tumor Cells by Array-CGH and Next-Generation Sequencing

Ellen Heitzer1, Martina Auer1, Christin Gasch6, Martin Pichler3, Peter Ulz1, Eva Maria Hoffmann1, Sigurd Lax5, Julie Waldispuehl-Geigl1, Oliver Mauermann6, Carolin Lackner2, Gerald Höfler3, Florian Eisner3, Heinz Sill3, Hellmut Samonigg3, Klaus Pante6, Sabine Riethdorf6, Thomas Bauernhofer3, Jochen B. Geigl1, and Michael R. Speicher1

Abstract

Circulating tumor cells (CTC) released into blood from primary cancers and metastases reflect the current status of tumor genotypes, which are prone to changes. Here, we conducted the first comprehensive genomic profiling of CTCs using array–comparative genomic hybridization (CGH) and next-generation sequencing. We used the U.S. Food and Drug Administration–cleared CellSearch system, which detected CTCs in 21 of 37 patients (range, 1–202/7.5 mL sample) with stage IV colorectal carcinoma. In total, we were able to isolate 37 intact CTCs from six patients and identified in those multiple colorectal cancer–associated copy number changes, many of which were also present in the respective primary tumor. We then used massive parallel sequencing of a panel of 68 colorectal cancer–associated genes to compare the mutation spectrum in the primary tumors, metastases, and the corresponding CTCs from two of these patients. Mutations in known driver genes [e.g., adenomatous polyposis coli (APC), KRAS, or PIK3CA] found in the primary tumor and metastasis were also detected in corresponding CTCs. However, we also observed mutations exclusively in CTCs. To address whether these mutations were derived from a small subclone in the primary tumor or represented new variants of metastatic cells, we conducted additional deep sequencing of the primary tumors and metastasis and applied a customized statistical algorithm for analysis. We found that most mutations initially found only in CTCs were also present at subclonal level in the primary tumors and metastases from the same patient. This study paves the way to use CTCs as a liquid biopsy in patients with cancer, providing more effective options to monitor tumor genomes that are prone to change during progression, treatment, and relapse. Cancer Res; 73(10); 1–11. ©2013 AACR.

Introduction

Circulating tumor cells (CTC) are rare cells found in the blood of patients with solid tumors. The isolation and characterization of CTCs has tremendous potential for new biologic insight with very real clinical applications, such as the identification of prognostic, predictive, and pharmacokinetic biomarkers (1, 2). For example, in colorectal cancer, the KRAS mutations in exon 2 (codons 12 and 13) represent a paradigm that has been established as a negative predictive marker for treatment with EGF receptor (EGFR) inhibitors (3). Thus, “real-time” longitudinal monitoring of CTC-derived genotypes may provide a noninvasive approach to identify drug sensitivity and resistance-associated markers, guiding therapeutic decisions. However, because CTCs constitute as few as 1 cell per 1 × 109 normal blood cells in patients with metastatic cancer, it is difficult to identify and isolate these cells with current detection methods (4). Multiple approaches to CTC isolation and characterization have been published (5, 6). However, the main application of the most currently available CTC detection systems consists of an enumeration of putative CTCs without further analyses.

A powerful example of the application of CTC genotyping was provided by a study of patients with non–small cell lung cancer (NSCLC). A subset of NSCLC has somatic activating mutations in EGFR, and these patients are likely to benefit from treatment with selective EGFR kinase inhibitors. Through monitoring of CTCs, the acquisition of the recurrent
T790M-EGFR drug resistance mutation has become evident in CTCs during the course of therapy and has coincided with the development of clinically refractory disease (7).

In fact, single-cell analyses have evolved to an active area, with various attempts to assess the genome of individual cells. For example, single-nucleus sequencing was shown to be capable of characterizing tumor evolution in breast cancer (8). However, this method was not suitable for assessing the genetic characteristics of single tumor cells at a single-nucleotide resolution because sequencing was conducted at merely approximately 0.2 times coverage. Therefore, only copy number variations in single tumor cells were analyzed (8). In other studies, individual cells from hematopoietic (9) and kidney (10) tumors were analyzed by exome amplification and subsequent sequencing. Obtaining a 30 times coverage, the data could be used to establish lineage relationships between the cells, based on the point mutations (9, 10). These studies also showed that amplification and sequencing errors are a concern for single-cell mutation analysis, as the false discovery ratio was 2–3 × 10^{-5} and the average allele dropout ratio for all single cells was approximately 11% (9, 10). To limit these problems, the authors defined a true mutation only if it occurred in a specified number of cancer cells. For example, somatic-mutant loci had to be present in at least 5 myeloproliferative tumor cells and be homozygous normal in the respective nonmalignant tissue used for comparison (9). More recently, a single-cell whole-genome analysis to measure the genomic diversity in one individual’s gamete genome was described to create personal recombination maps from whole-genome sperm data (11). Using a special microfluidic device on haploid cells, the authors reduced the false discovery rate to 4 × 10^{-9} with 5 times coverage.

These studies suggest that high-resolution single-cell analysis is feasible. However, compared with the aforementioned technologies, single-cell CTC analyses differ in 4 important aspects. First, somatic mutations in tumor cells include base substitutions, indels, structural rearrangements, and changes in the copy number of DNA segments (12). However, no single-cell approach that effectively combines the reliable mapping of copy number changes with next-generation sequencing approaches to detect nucleotide changes has been published so far. Second, CTCs are extremely rare events and, despite many efforts (2, 13, 14), there is still a lack of technologies capable of isolating them in sufficient numbers. As a consequence, usually only very limited numbers of cells are available for analysis. This excludes the aforementioned strategies, such as calling mutations only if they are observed in a specified number of cells. Third, isolation of CTCs out of millions of normal cells is a much more complicated procedure than isolation of cells from a primary tumor or a selection of sperm cells. Fourth, no suitable material for comparison of CTC results is available. This is because CTCs may recur years after initial diagnosis of the primary tumor and may have acquired multiple novel changes since then. Furthermore, CTCs may be released from various metastatic sites, and their origin can usually not be traced. In addition, CTCs have been reported to be heterogeneous (15–17) and may therefore exhibit a tremendous cell-to-cell variability.

For these reasons, the detailed characterization of CTCs is still in its infancy. To date, only a few examples exist in which high-dimensional analyses of individual CTCs were conducted, and these were analyses of CTC gene expression (18, 19). Here, we carried out a pilot study to investigate strategies for copy number analysis and next-generation sequencing of single CTCs. Our results suggest that complex tumor genomes can be reconstructed from the peripheral blood of patients with cancer, and this approach may pave the way for new disease monitoring strategies.

Materials and Methods

Patients

All patients had advanced-stage [Union Internationale Contre le Cancer (French) stage IV] progressive disease at the time of blood collection. Clinical characteristics are summarized in Supplementary Table S1. The study was approved by the local ethics committee, and written informed consent was obtained from all the patients. The patients were seen at the Division of Clinical Oncology, Department of Internal Medicine, Medical University of Graz (Graz, Austria).

Detection of CTCs and whole-genome amplification

Blood samples (7.5-mL each) were collected into CellSave tubes (Veridex, Raritan). The Epithelial Cell Kit (Veridex) was applied for CTC enrichment and enumeration with the CellSearch system as described previously (20). In brief, in a first step, CTCs were captured by antitumor cell adhesion molecule (EpCAM) antibody-bearing ferrofluid. Subsequent identification of CTCs was based on cytookeratin positivity and negativity for the leukocyte common antigen CD45. In addition, 4', 6-diamidino-2-phenylindole staining was done to evaluate the integrity of the nucleus. The further processing of CTCs is described in the Supplementary Materials and Methods.

Enrichment of genes and 454 Life Sciences/Roche Diagnostics

The enrichment of genes and subsequent sequencing is outlined in the Supplementary Materials and Methods.

Deep sequencing with the Illumina MiSeq

For validation of ”private CTC mutations”, i.e., mutations found in only one CTC, we employed a comparative ultradepth sequencing approach using the Illumina MiSeq as exemplified in the Supplementary Materials and Methods.

Results

Patients

In this study, we used the U.S. Food and Drug Administration (FDA)–approved CellSearch system for CTC detection (20) in 37 patients with stage IV colorectal cancer. In 5 (13.5%) cases, analysis was not possible due to technical reasons, for example clotting of the blood. In 11 (29.7%) cases, no CTCs were found, in 15 (40.5%) cases between 1 and 10 CTCs were found, and in 6 (16.2%) cases more than 10 CTCs were identified. In the latter 6 cases (i.e., patient #6, #9, #18, #22, #26, and #38), we isolated CTCs for further analyses. For these 6 patients, the interval between diagnosis of the primary tumor and blood collection to
obtain CTCs varied, with a range from 1 month (patient #9 and #18) to 3 years and 7 months (45 months, patient #38; mean, 15 months; median, 5 months). In addition, clinical characteristics of these patients are summarized in Supplementary Table S1.

Establishment of copy number changes in CTCs

In a next step, we wanted to test whether tumor-specific copy number changes are reflected in CTCs. To this end, we used our previously published whole-genome amplification (WGA) and array–comparative genomic hybridization (CGH) technologies for analysis of a single cell or a few cells (21–27). Altogether, we analyzed 37 CTCs and, as a control, we isolated leukocytes (CD45-positive cells). As expected, the CD45-positive cells showed balanced array-CGH profiles (data not shown).

Our analysis consisted of several steps, which we carried out for each CTC of each patient. We use patient #26 as an example to explain these steps. The primary tumor of patient #26 had initially been diagnosed and completely removed 34 months before our analysis. A liver metastasis was noted and resected 10 months after the initial diagnosis. Twenty-one months later, metastases were identified in the liver, spleen, and bone. We started our analyses by comparing the ratio of the profiles of the available tumor material, in this case, the primary tumor (Fig. 1A), with the liver metastasis (Fig. 1B). We observed several common copy number changes, including, for example, losses of chromosomes 4, 5q13.2–5q31.2, which harbors the adenomatous polyposis coli (APC) gene, 8p, 17, and 18, and gains of chromosomes 8q, 9, and 20. However, some differences were also observed, such as losses on chromosomes 10, 11, and 12 or gain of chromosome 7, which were only present in the metastasis but not in the primary tumor.

Figure 1. Analysis of the primary tumor, metastasis, and one circulating tumor cell (CTC05) from patient #26. A, ratio profile of the primary colorectal cancer tumor of patient #26. The single green and red bars summarize the regions that were gained or lost based on all iterative calculations of our algorithm (Supplementary Information). The black profile regions represent balanced regions, lost regions appear in red, and gained regions are shown in green. B, array-CGH of the liver metastasis of patient #26. C, ratio profile of one representative CTC, i.e., CTC05. D, heatmaps comparing the copy number changes in the primary tumor (PT), metastasis (Met.), and CTC05 (black, balanced regions; red, underrepresented regions; green, overrepresented regions). E, the bar chart displays the percentages of chromosomal regions that were commonly lost (red), balanced (black), or gained (green) in all 3 samples, i.e., primary tumor, metastasis, and CTC05, shared by metastasis and CTC05 only (blue), shared by primary tumor and CTC05 only (yellow), or unique to CTC05 (gray).
Accordingly, CTCs showed similar copy number changes compared with both primary tumor and metastasis but also some others, such as gains of chromosome 10 and 13 in CTC05 (Fig. 1C). For our array-CGH, we used a 60-K microarray platform consisting of 59,012 oligonucleotides for analysis, and we calculated for each of these oligonucleotides whether the ratio values were decreased, balanced, or increased and illustrated the results in heatmaps (Fig. 1D). Finally, we determined whether the copy number status for each oligonucleotide occurred only in the respective CTC or also in the primary tumor and/or metastasis (Fig. 1E). This detailed analysis revealed that the copy number status of 60.5% of the oligonucleotides on our array platform was identical in the primary tumor and metastasis as well as in CTC05 (i.e., of all 59,012 oligonucleotides on our array, 8.7% oligonucleotides showed a loss, 42.9% remained balanced, and 8.9% showed a gain in copy number status); 9.3% of the oligonucleotides were shared only by the primary tumor and CTC05, 12.6% were shared only by the metastasis and CTC05, and 17.6% were unique to CTC05 (Fig. 1E). These findings suggested that the observed changes in the CTC05 were tumor specific and consisted of changes also present in the primary tumor and the liver metastasis, and in addition some other changes.

To investigate the CTC population structure, we extended these analyses to 10 CTCs from patient #26 (Fig. 2A).

Figure 2. Analysis of 10 CTCs from patient #26. A, percentages of oligonucleotides that were commonly lost (red), balanced (black), or gained (green) in the primary tumor, metastasis, and respective CTC. Identical copy number changes occurring only in metastasis and respective CTCs are shown in blue and identical changes in both primary tumor and CTC are shown in yellow. Copy number changes, which were observed only in the CTCs but not in the primary tumor or metastasis, are displayed in gray. On average, for these 10 CTCs, 52.0% (median, 53.4%; range, 37.3%–61.5%) of the copy number status changes (i.e., gains, losses, and balanced regions) were present in all CTCs, primary tumors, and analyzed metastases; 8.3% (median, 8.2%; range, 6.7%–10.4%) were partially shared between all CTCs and the metastasis only; 14.2% (median, 14.2%; range, 12.0%–16.3%) were partially shared between the CTCs and primary tumor only; and 25.6% (median, 24.8%; range, 17.6%–41.9%) of all CTC copy number changes were not observed in the primary tumor or metastasis. B, integer copy number profile of CTC26. C, combined copy number profile representing an average profile from all individual CTCs. D, hierarchical cluster analysis to determine the CTC population substructure and the relationship of CTCs to the primary tumor and metastases. mCTCl, main CTC lineage, i.e., the average copy number profile from all CTCs (black, balanced regions; red, underrepresented regions; and green, overrepresented regions).
Furthermore, assuming that single cells would have distinct copy number states, we calculated integer copy number profiles (Supplementary Methods) for each CTC. This analysis revealed that all 10 CTCs were within the tetraploid range (e.g., CTC26, Fig. 2B; additional CTCs are depicted in Supplementary Fig. S1A–S1D). To establish the most frequently occurring copy number changes within the CTC population of patient #26, we constructed an average copy number profile from all individual CTCs (Supplementary Methods; Fig. 2C). The "main CTC lineage" differed from the primary tumor and its metastasis (Fig. 1A and B) by gains on chromosomes 1q, 6q, and 12p, a high-level gain on chromosome 13, and loss on chromosome 14. When we conducted hierarchical clustering from the copy number profiles (Fig. 2D), the integer copy number profile of one cell (i.e., CTC28; Supplementary Fig. S1D) deviated from that of the other cells, suggesting the presence of cells with a "private" pattern of alterations.

CTCs display multiple colorectal cancer-associated copy number changes

We extended these analyses to CTCs from the other patients. In patient #6, we noted extensive differences between the copy number profiles of the primary tumor and its metastasis (Supplementary Fig. S2A and S2B; details to patient’s #6 history are in Supplementary Information). From patient #6, we analyzed 9 cells. In 3 cells, we did not find the KRAS G12V mutation previously identified in the primary tumor, and in these cells we did not observe any tumor-related copy number changes. From the 6 other cells that contained the KRAS mutation, 2 cells were in the diploid range with few copy number changes, whereas 4 cells were tetraploid and had multiple copy number changes (Fig. 3A). Interestingly, the average CTC copy number profile (Fig. 3B) differed from the primary tumor (Supplementary Fig. S2A) by a gain of chromosome 8q, a numerical change that has frequently been described in colorectal cancer (www.progenetix.net). Our
analyses revealed that the CTCs had more chromosomal copy number changes in common with the primary tumor than with the cerebellar metastasis (Fig. 3C). Again, cluster analysis (Fig. 3D) identified a cell with a “private” copy number profile (i.e., CTC14; Supplementary Fig. S2C).

At the time of diagnosis, patient #9 already had metastases in liver, bone, and abdomen (peritoneal carcinomatosis). We obtained biopsies from the primary tumor (Supplementary Fig. S3A) and the peritoneal carcinomatosis (Supplementary Fig. S3B), which revealed almost identical array-CGH profiles from both sites. We conducted our CTC analyses 1 month after diagnosis and observed many similarities between the copy number profiles of the CTCs and those from the primary tumor and peritoneal carcinomatosis, respectively (Fig. 4).

We analyzed 6 epithelial and CD45-negative cells selected by the CellSearch system from patient #38 (Supplementary Fig. S4A and S4B; history in Supplementary Information). However, we were unable to identify the KRAS G12D mutation, identified in the primary tumor, in any of these cells. Furthermore, all of these cells had a balanced profile.

In 2 further cases (#18, #22) only small biopsies were taken at the time of diagnosis, so that insufficient material was available to analyze the primary tumor for comparison. However, in their CTCs, we again identified multiple copy number changes associated with colorectal cancer (as documented at www.progenetix.net). For example, in patient #22, one CTC revealed losses on chromosomes 3, 4, 5, 8p, and 18 and gains on chromosomes 7p, 17q, and 20 (Supplementary Fig. S5; CTC03).

Sequencing of the ultraconserved region UCR41 in CTCs

To establish appropriate CTC sequencing strategies, we selected CTCs from patients #6 and #26 because we had the largest CTC number for these 2 patients. However, sequencing of single-cell amplification products can be prone to experimental and technical errors (9–11). To address this issue, we exploited ultraconserved regions (UCR), which represent genomic regions that do not accumulate true mutations (28). Thus,
all sequence variations in UCRs should represent bona fide errors, making UCRs an ideal control for DNA amplification and sequencing errors (28). Therefore, we sequenced the ultraconserved region UCR41 (ref. 29; Supplementary Fig. S6, green bar), its adjacent extremely conserved region (Supplementary, Fig. S6, gray bar), and nonconserved flanking regions (28) and analyzed 5 cells each from healthy male and female controls. The only sequence variant in single cells from healthy controls was at single-nucleotide polymorphism (SNP) rs17701179, where the male donor had A/G alleles, which we also identified in each of his 5 single cells (Supplementary Fig. S6, blue arrows).

Next, we analyzed the constitutional DNA, primary tumor, metastasis, and all the CTCs from patients #6 (n = 4) and #26 (n = 10), using the same procedure. Again, we found no sequence variation in the ultraconserved or extremely conserved regions. However, we identified 2 variations in the nonconserved flanking regions. One was in the metastasis sample from patient #6 (Supplementary Fig. S6, right red arrow), in which the sample had not been subjected to WGA but only to the amplification steps necessary for the targeted gene enrichment. The second variation was in one of the 10 CTCs from patient #26 (CTC21; Supplementary Fig. S6, left red arrow). Both of these variants were within a region that seems to have increased mutability compared with the average mutability (28). Furthermore, patient #6 had constitutional A/G alleles for SNP rs3910657; however, we identified A/A in CTC7 and G/G in CTC14. Our results suggest that, despite the amplification steps involved in our protocol, no sequencing artifacts occurred in this ultraconserved region.

**Mutation spectrum in primary tumors, metastases, and CTCs**

We then aimed at establishing a mutation spectrum for a panel of 68 colorectal cancer-related genes (Supplementary Table S2) that are frequently (> 3%) mutated in colorectal cancer according to the COSMIC database. We analyzed these 68 genes in the different samples, including constitutional DNA, the primary tumors, and their metastases in patients #6 and #26. In addition, we sequenced 3 CTCs (7, 13, and 14) from patient #6 and 5 CTCs (5, 21, 22, 24, and 28) from patient #26.

A total of 959,790 and 2,070,381 sequences were generated for the formalin-fixed paraffin-embedded (FFPE) pool (tumor and metastasis of patient #6 and #26) and the 2 CTC pools (pool 1: CTC 7, 13, 14 from patient #6 and CTC 5 from patient #26; pool 2: CTC 21, 22, 24, 28 from patient #26), respectively. Read length was remarkably consistent throughout the experiment. The average read length was 232 bp for the FFPE material and 329 bp for the CTCs, and the total yield from the sequencing run was 223 Mb for the 4 FFPE samples and 683 Mb for the 8 CTCs. The average coverage of tumors and metastases (×24) was lower as compared with the CTCs (×41). Sequencing of enriched CTCs achieved more on-target reads (55.6%) than did sequencing of the FFPE material (41.5%).

With the exception of an exonic deletion in ERCC6L in patient #6 and a 1-bp insertion in MLH1 in patient #26 (in CTC28), all identified insertions and deletions were found only in introns. We focused our analysis on exonic, nonsynonymous variants not included in the National Center for Biotechnology Information dbSNP build 132 database. In the initial screen, we found 25 mutations in 17 genes in at least one of the analyzed samples of patients #6 and 22 mutations in 10 genes in patient #26. Sanger sequencing confirmed 16 and 15 mutations in patients #6 (Fig. 5A) and #26 (Fig. 5B), respectively (Supplementary Table S3). In patient #6, 4 of these mutations were constitutional and therefore found in all the samples. This patient also had 3 somatic mutations in known colorectal cancer driver genes (i.e., in APC, KRAS, and PIK3CA), which were present in all the tumor samples including CTCs. Interestingly, we found 2 additional mutations in the cerebellar metastasis in NF1 and TP53, which we also detected in the 3 CTCs. The other mutations found were present only in single CTCs examined (Fig. 5A).

In patient #26, we identified one constitutional mutation in all analyzed samples. However, within our panel of genes, we did not find any other mutation present in all samples (Fig. 5B). However, array-CGH had identified a deletion on chromosome 5 harboring the APC gene in all tumor samples, and therefore in this case copy number changes may indicate genomic regions, which had been involved in tumor initiation and/or progression. In patient #26, we again identified “private mutations,” defined as mutations observed in only one CTC but not in any other tumor material. Because of the high experimental error rate of single-cell analysis, we carried out additional experiments to address the nature of these mutations as outlined below.

**Most private CTC mutations are present at a subclonal level in the primary tumor**

We reasoned that mutations found in only one CTC could be amplification/sequencing artifacts or mutations already present in the primary tumor or metastasis at subclonal level and therefore could have been missed by our initial sequencing efforts. To address this question, we conducted ultradepth sequencing of the primary tumors, metastases, and normal tissues to test whether private CTC mutations might reflect subclonal mutations. Subclonal single-nucleotide variants were detected and quantified with the deep single-nucleotide variant (SNV) algorithm, an approach capable of detecting variants with frequencies as low as one per 10,000 alleles (30). These analyses revealed that from 20 analyzable private CTC mutations (one variant was not amplifiable), 17 (85%) were present in at least one other corresponding sample (i.e., primary tumor or metastasis) with mutation frequencies ranging from 0.02 to 0.42 (Fig. 5; Table 1). For the remaining mutations, it is impossible to decide whether these are true mutations occurring with allele frequencies below the resolution limits of current detection methods or if they are amplification or sequencing artifacts.

**Correlation with clinical parameters and potential clinical applications**

Although this was a pilot study, we then attempted to investigate whether the identified set of genetic aberrations could potentially be informative for identifying rational therapies currently available or in clinical trials. Mutations
of biologic or predictive relevance are listed in Fig. 5A and B. For example, 2 mutations in patient #6 are of clinical relevance, i.e., \textit{KRAS} p.G12V (resistant to EGFR monoclonal antibodies) and \textit{PIK3CA} p.E542K [an exon 9 activating mutation; possible predictor for EGFR antibodies and/or insulin-like growth factor (IGF)-1R inhibitors; ref. 31]. Although the CTC analysis confirmed that these mutations were still present 8 months after diagnosis, it did not reveal novel targets for therapeutic interventions as compared with the analysis of the primary tumor. In contrast, in patient #26, we identified cyclin-dependent kinase 8 (\textit{CDK8}) amplification on chromosome 13q12.13 in 9 of 10 CTCs that was not present in the parts of the primary tumor or metastasis that we analyzed. This amplification may represent a viable target for CDK inhibitors, which are currently in clinical trials (32–34).

**Discussion**

A major goal of cancer medicine is to move from fixed treatment regimens to therapies tailored to a patient’s individual tumor. In this study, we addressed whether complex tumor genomes can be inferred noninvasively from CTCs of patients with cancer. Detailed characterization of CTCs may provide the opportunity for clinical impact in the treatment of selected cancers in which appropriate targeting of tumor-associated genetic lesions is critical. Such an approach would allow monitoring of tumor genomes with an unprecedented resolution. However, the science of CTCs is still in its infancy. Little is known about the biology of these cells, and only recently technologic advances have provided a window into the composition of this rare cell population, allowing molecularly designed analyses (18, 19).
The importance of a molecular characterization of these cells is underlined by the fact that we had 9 epithelial and CD45-negative cells selected by the CellSearch system (3 from patient #6 and 6 from patient #9), in which we did not find the KRAS mutation previously identified in the respective primary tumors and which each had a balanced copy number profile. Formally, we cannot exclude that these cells are CTCs; however, alternatively these cells could be circulating epithelial cells, which have recently been described in patients with benign colon diseases (35). Thus, a simple enumeration of epithelial, CD45-negative cells without molecular characterization for prognostic purposes may in some cases be erroneous.

Although the CellSearch system used for CTC enrichment has been cleared by the FDA, it has shortcomings, as the enrichment is based on EpCAM expression of CTCs and EpCAM-negative CTCs are missed. Increasing evidence suggests that EpCAM-negative CTCs might have undergone an epithelial-to-mesenchymal transition, a process linked to the stemness of cancer cells and increased chemoresistance (36). Other emerging CTC isolation techniques, including non-EpCAM or size-based enrichment approaches (5), have the advantage of assessing the relevant genomic changes in a broader range of CTC subsets. However, these new CTC assays have not been validated to a comparable level as the CellSearch system with regard to their specificity, reproducibility, and clinical relevance so far. Moreover, recent reports indicate that carcinoma cells have to express epithelial markers (such as EpCAM) to colonize distant organs and to form overt metastases (36).

We applied our single-cell technologies (21–27) to a detailed analysis of CTCs and showed for the first time that mutation analysis of multiple genes by next-generation sequencing in combination with establishing the copy number status is feasible from CTCs. Our strategy differs from other recently published single-cell approaches as these have shortcomings for the analysis of rare cell events. For example, single-cell sequencing (8) allows only assessment of copy number changes.

### Table 1. Results of ultradeep sequencing with DNA from primary tumor and metastases of patients #6 and #26, respectively

<table>
<thead>
<tr>
<th>Gene</th>
<th>Amino acid change</th>
<th>Sampleb and patient</th>
<th>Primary tumor % Mutatedc Total reads</th>
<th>Metastasis % Mutatedc Total reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCA1</td>
<td>p.P1209S</td>
<td>CTC7,#6</td>
<td>0.05 950,656</td>
<td>0.08 480,081</td>
</tr>
<tr>
<td>C10orf137</td>
<td>p.A990E</td>
<td>CTC14,#6</td>
<td>0.02 938,780</td>
<td>not significantd</td>
</tr>
<tr>
<td>CACNA2D3</td>
<td>p.G84S</td>
<td>CTC14,#6</td>
<td>0.07 842,856</td>
<td>0.07 996,436</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>p.A149T</td>
<td>CTC14,#6</td>
<td>0.08 547,498</td>
<td>0.42 432,851</td>
</tr>
<tr>
<td>GNAS</td>
<td>p.G869D</td>
<td>CTC14,#6</td>
<td>not significantd</td>
<td>0.10 109,096</td>
</tr>
<tr>
<td>GUCY1A2</td>
<td>p.H439Y</td>
<td>CTC14,#6</td>
<td>0.06 299,872</td>
<td>0.02 299,872</td>
</tr>
<tr>
<td>LAMA1</td>
<td>p.P1414S</td>
<td>CTC7,#6</td>
<td>0.16 433,621</td>
<td>0.11 607,163</td>
</tr>
<tr>
<td>LAMA1</td>
<td>p.H1002Y</td>
<td>CTC7,#6</td>
<td>not significantd</td>
<td>not significantd</td>
</tr>
<tr>
<td>ABCA1</td>
<td>p.A1714T</td>
<td>CTC21,#26</td>
<td>0.04 315,684</td>
<td>0.11 153,447</td>
</tr>
<tr>
<td>ADAMTSL3</td>
<td>p.Q756X</td>
<td>CTC24,#26</td>
<td>0.20 258,482</td>
<td>0.02 242,142</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>p.L160I</td>
<td>CTC5,#26</td>
<td>0.07 546,086</td>
<td>0.06 433,461</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>p.G277D</td>
<td>CTC5,#26</td>
<td>0.18 374,324</td>
<td>0.03 118,076</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>p.C429Y</td>
<td>CTC24,#26</td>
<td>0.10 155,213</td>
<td>0.05 108,917</td>
</tr>
<tr>
<td>CSMD3</td>
<td>p.V1723I</td>
<td>CTC22,#26</td>
<td>not significantd</td>
<td>0.21 247,608</td>
</tr>
<tr>
<td>GNAS</td>
<td>p.G869D</td>
<td>CTC5,#26</td>
<td>not significantd</td>
<td>not significantd</td>
</tr>
<tr>
<td>KIAA1409</td>
<td>p.S1344P</td>
<td>CTC21,#26</td>
<td>not significantd</td>
<td>not significantd</td>
</tr>
<tr>
<td>LAMA1</td>
<td>p.W818X</td>
<td>CTC5,#26</td>
<td>not significantd</td>
<td>not significantd</td>
</tr>
<tr>
<td>MLH1</td>
<td>p.R497Pfs*6</td>
<td>CTC28,#26</td>
<td>not significantd</td>
<td>not significantd</td>
</tr>
<tr>
<td>NAV3</td>
<td>p.G1921S</td>
<td>CTC21,#26</td>
<td>0.08 359,151</td>
<td>0.03 190,238</td>
</tr>
<tr>
<td>NAV3</td>
<td>p.S594F</td>
<td>CTC5,#26</td>
<td>0.08 125,637</td>
<td>not significantd</td>
</tr>
<tr>
<td>OR51E1</td>
<td>p.A312V</td>
<td>CTC21,#26</td>
<td>not analyzablee</td>
<td>not analyzablee</td>
</tr>
</tbody>
</table>

**NOTE:** The columns "Gene" and "Amino acid change" list all private CTC mutations, i.e., mutations found in only one CTC. The percentages of mutated DNA fragments (% Mutated) and the total number of reads (Total reads) are given for the respective samples of patients #6 and #26, respectively.

*Human Genome Variation Society mutation nomenclature.

*CTCs of patient 6 (#6) or patient 26 (#26).

%d% Mutated, percentage of total number of mutated reads (Phred quality of >15) calculated with the deepSNV algorithm.

Not significant, p value from deepSNV algorithm for the likelihood calculation whether an observed variant is true or a sequencing error was not significant after Benjamini–Hochberg correction.

*Not analyzable, PCR amplification failed.
but not of changes at the nucleotide level. The single-cell exome sequencing methods (9, 10) defined true mutations only if they occurred in a specified number of cancer cells. This is applicable for the analysis of cells from a primary tumor, in which a virtually unlimited number of cells is available for analysis, but not for rare events, such as CTCs, in which frequently less than 5 cells with sufficient quality are available. Furthermore, this approach does not allow a genome-wide measurement of copy number changes. The single-cell whole-genome analysis presented by Wang and colleagues (11) seems to be especially attractive as it may significantly reduce the false-discovery rate. However, it is currently unknown whether the microfluidic device used by these authors is applicable to the identification and isolation of CTCs. Very recently, a new amplification method offering high uniformity across the genome and also allowing both the genome-wide detection of single-nucleotide and copy-number variations of a single human cell has been reported (37, 38). Although this new amplification method has not been applied to CTCs yet, it seems to have great potential to advance CTC research further.

In any case, the aforementioned studies have documented that single-cell analysis is prone to artifacts, which may be introduced either during amplification or sequencing (9–11, 37, 38). To address this issue, we made use of a comparative sequencing strategy. This strategy involved reanalysis of the primary tumor and metastasis by deep sequencing. The same genomic region was compared between a heterogeneous test sample (i.e., tumor and metastasis) and a homogeneous control sample (i.e., normal tissue). Sequencing results were evaluated with deepSNV (30). With these sequencing efforts and statistical evaluations, variants occurring with allele frequencies ranging from 0.0002 to 0.34 can be identified in a heterogeneous tissue (30). Interestingly, these analyses suggested that the majority of private CTC mutations had already been present at a subclonal level in the primary tumor. Therefore the majority of these variants were likely true mutations and not artifacts.

Our study suggests that we can elucidate relevant changes in the tumor genome that had either not been present or not been observed at the time of initial diagnosis. Thus, an important application may be the detailed reevaluation of patients enrolled in clinical trials for novel substances. For example, analysis of the genomes of the primary tumor and metastasis of patient #26 did not reveal changes that would have made her eligible for enrollment in a trial with a CDK inhibitor such as flavopiridol (32–34). Yet, the CTCs obtained 34 and 24 months after diagnosis of the primary tumor and liver metastasis, respectively, revealed a high level of amplification of CDK8, which had not been noted in the previous analyses. CDK8 has been described as a positive regulator of catenin signaling in 15% to 20% of colorectal cancers (39). Furthermore, CDK8 inhibition suppresses damage-induced tumor-promoting paracrine activities of tumor cells, and it has recently been reported that CDK8 inhibitors offer a promising approach to increase the efficacy of cancer chemotherapy (40). Therefore this amplification may represent a viable target for CDK inhibitors, which are currently in clinical trials (32–34). Although we cannot yet provide treatment outcomes, patient #26 may serve as an example of the impact of reevaluation of the tumor genome status on treatment decisions.

The next challenge will be to test such approaches for the serial monitoring of patients with cancer and to evaluate their potential for the early-stage analysis of peripheral blood. However, our proof-of-concept study suggests that minimal invasive access to tumor material immediately before specific antitumor treatment initiation in advanced cases may become possible for reevaluation of the cancer genome and possibly tailored management of treatment choices.

Disclosure of Potential Conflicts of Interest
H. Sill has commercial research grants from Gerot and Roche, and K. Pantel is a consultant/advisory board member of Veridex. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions
Conception and design: E. Heitzer, M. Pichler, K. Pantel, T. Bauernhofer, J.B. Geigl, M.R. Speicher
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): M. Pichler, S. Lax, J. Waldispuehl-Geigl, C. Lackner, Gerald Höller, F. Eisner, H. Samonigg, K. Pantel, S. Riethdorf, T. Bauernhofer, J.B. Geigl
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): E. Heitzer, M. Auer, M. Pichler, P. Ulz, S. Lax, H. Sill, K. Pantel, S. Riethdorf, M.R. Speicher
Writing, review, and/or revision of the manuscript: E. Heitzer, M. Auer, M. Pichler, S. Lax, Gerald Höller, H. Sill, K. Pantel, S. Riethdorf, T. Bauernhofer, J.B. Geigl, M.R. Speicher
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): E.M. Hoffmann, F. Eisner, H. Samonigg, T. Bauernhofer, J.B. Geigl

Acknowledgments
The authors thank Mag. Maria Langer-Winter for editing the article.

Grant Support
This work was funded by the European Commission (GENINCA, contract no. 202230) to K. Pantel and M.R. Speicher), the Austrian Science Fund (FWF; grant#: P20338 and W 1226-B18, DKplus Metabolic and Cardiovascular Disease to M.R. Speicher) and by the COMET Center ONCOTYROL (M.R. Speicher), E.M. Hoffmann was supported by the Molecular Medicine PhD Program of the Medical University of Graz.

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 November 7, 2012; revised January 15, 2013; accepted February 18, 2013; published OnlineFirst March 7, 2013.

References


Complex Tumor Genomes Inferred from Single Circulating Tumor Cells by Array-CGH and Next-Generation Sequencing

Ellen Heitzer, Martina Auer, Christin Gasch, et al.

Cancer Res  Published OnlineFirst March 7, 2013.

Updated version  Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-12-4140

Supplementary Material  Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2013/03/07/0008-5472.CAN-12-4140.DC1

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, contact the AACR Publications Department at permissions@aacr.org.