RESEARCH BRIEF

Genomic Landscape of Cell-Free DNA in Patients with Colorectal Cancer

John H. Strickler[1], Jonathan M. Loree[2], Leanne G. Ahronian[3], Aparna R. Parikh[3], Donna Niedzwiecki[1], Allan Andresson Lima Pereira[2], Matthew McKinney[1], W. Michael Korn[4,5], Chloe E. Atreya[4], Kimberly C. Banks[6], Rebecca J. Nagy[6], Funda Meric-Bernstam[2], Richard B. Lanman[6], AmirAli Talasaz[6], Igor F. Tsigelny[7,8], Ryan B. Corcoran[3], and Scott Kopetz[2]

ABSTRACT

"Liquid biopsy" approaches analyzing cell-free DNA (cfDNA) from the blood of patients with cancer are increasingly utilized in clinical practice. However, it is not yet known whether cfDNA sequencing from large cohorts of patients with cancer can detect genomic alterations at frequencies similar to those observed by direct tumor sequencing, and whether this approach can generate novel insights. Here, we report next-generation sequencing data from cfDNA of 1,397 patients with colorectal cancer. Overall, frequencies of genomic alterations detected in cfDNA were comparable to those observed in three independent tissue-based colorectal cancer sequencing compendia. Our analysis also identified a novel cluster of extracellular domain (ECD) mutations in *EGFR*, mediating resistance by blocking binding of anti-EGFR antibodies. Patients with *EGFR* ECD mutations displayed striking tumor heterogeneity, with 91% harboring multiple distinct resistance alterations (range, 1–13; median, 4). These results suggest that cfDNA profiling can effectively define the genomic landscape of cancer and yield important biological insights.

SIGNIFICANCE: This study provides one of the first examples of how large-scale genomic profiling of cfDNA from patients with colorectal cancer can detect genomic alterations at frequencies comparable to those observed by direct tumor sequencing. Sequencing of cfDNA also generated insights into tumor heterogeneity and therapeutic resistance and identified novel *EGFR* ectodomain mutations. *Cancer Discov; 8(2); 164–73. ©2017 AACR.*

INTRODUCTION

Next-generation sequencing (NGS) of tumor tissue specimens from large patient cohorts has led to major advances in elucidating the genomic landscape of cancer. Despite the many benefits and therapeutic insights offered by this approach, large-scale genomic profiling through tissue-based sequencing methods is not without limitations. Importantly, to be suitable for sequencing, tissue samples must be of

[1]Duke University Medical Center, Durham, North Carolina. [2]The University of Texas MD Anderson Cancer Center, Houston, Texas. [3]Massachusetts General Hospital Cancer Center and Department of Medicine, Harvard Medical School, Boston, Massachusetts. [4]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, California. [5]Caris Life Sciences, Phoenix, Arizona. [6]Guardant Health, Inc., Redwood City, California. [7]University of California, San Diego, San Diego, California. [8]CureMatch Inc., San Diego, California.

Note: Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

Corresponding Authors: Ryan B. Corcoran, Harvard Medical School, 149 13th Street, Boston, MA 02129. Phone: 617-726-8599; Fax: 617-643-0798; E-mail: rbcorcoran@partners.org; and Scott Kopetz, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 426, Houston, TX 77030. Phone: 713-792-2828; Fax: 713-563-2957; E-mail: skopetz@mdanderson.org

doi: 10.1158/2159-8290.CD-17-1009

adequate quantity and must possess a sufficient percentage of tumor cells among a background of normal cells and stroma. Although many surgical resection specimens are of suitable quality, needle-based biopsy samples from metastatic lesions often yield insufficient tumor cells to produce robust sequencing data (1, 2). As a result, many tumor tissue sequencing databases are derived from earlier-stage, largely untreated cancers, as opposed to tumors from patients with metastatic disease, where small biopsies are typically performed (3, 4). Thus, tissue-based sequencing compendia may not accurately represent the genomic landscape and biology of advanced cancers.

With the recent increase in clinical cell-free DNA (cfDNA) testing, large cfDNA-based sequencing databases could represent an attractive resource for genomic studies. However, it is not yet known whether the genomic landscape defined by population-based cfDNA sequencing accurately reflects the genomic landscape of mutations identified in large tissue-based sequencing studies. cfDNA represents DNA present in the noncellular portion of blood that originates from either normal tissue or tumor sources. The tumor-derived portion of cfDNA is termed circulating tumor DNA (ctDNA) and represents DNA that is shed into the bloodstream by tumor cells throughout the body. cfDNA testing offers access to tumor DNA, regardless of whether a surgery or biopsy is clinically indicated, thus protecting patients with metastatic disease from potential complications of an invasive procedure. Furthermore, as cfDNA can be shed into the blood from multiple lesions in an individual patient, it can offer an anatomically unbiased sampling of primary and metastatic tumor lesions, thereby generating insights into intra- and intertumoral heterogeneity (5–12). Therefore, understanding the key similarities and differences between cfDNA and tissue-based tumor databases would be informative in determining whether large-scale cfDNA profiling might represent an effective and efficient approach to define the mutational landscape of specific cancer types.

Here, we report results from clinical cfDNA testing of 1,397 deidentified individual patients with advanced colorectal cancer and compare these results to three independent large-scale tissue-based sequencing databases. We find that cfDNA profiling detects genomic alterations at frequencies comparable to those previously reported by direct tumor sequencing. Our findings also reveal how analysis of large cfDNA sequencing databases can provide novel and clinically relevant insights into tumor heterogeneity and therapeutic resistance.

RESULTS

Comparison of the Mutational Landscape of cfDNA and Tumor Tissue

To assess the potential utility of a large cfDNA sequencing database, we analyzed 1,772 consecutive blood specimens from patients with colorectal cancer who underwent testing with a targeted NGS assay (Guardant360, Guardant Health) between June 1, 2014, and May 18, 2016. There were three versions of the assay during the study time period, covering 54, 68, and 70 genes, respectively (Supplementary Table S1). In all, 1,500 cases (85%) had at least one genomic alteration detectable in cfDNA. Of these, 103 samples represented serial assays from the same patient, leaving 1,397 unique patients with genomic data for analysis (Supplementary Fig. S1).

The prevalence of nonsynonymous single-nucleotide variants (SNV) detected in cfDNA was compared with those observed in three publicly available colorectal cancer tissue-based databases, including The Cancer Genome Atlas (TCGA; $n = 228$; ref. 3), the Nurses Health Study/Health Professionals Follow-up Study (NHS/HPFS; $n = 619$; ref. 4), and the AACR Project Genomics Evidence Neoplasia Information Exchange (GENIE; $n = 1{,}149$; ref. 13). For consistency, the analysis of all cohorts was adjusted to include only those SNVs covered by the cfDNA assay. Mutational prevalence was largely consistent among the three tissue cohorts (Fig. 1A), although some small differences were observed, which may be due to expected statistical variability or to differences in demographics among the three tissue cohorts. For example, the increased prevalence of *BRAF* mutations in the NHS/HPFS cohort is likely due to the increased representation of women in this cohort (*BRAF* mutations are more common in women than men with colorectal cancer; Supplementary Table S2). Overall, there were striking similarities in mutational prevalence between cfDNA and the three tissue-based databases (Fig. 1A; Supplementary Table S3). The mutational prevalence of the 20 most commonly mutated genes in cfDNA was strongly associated with the mutational prevalence in tumor tissue ($R^2 = 0.95$; $P < 0.0001$; Fig. 1B). Still, there were some notable differences between cfDNA and tissue. For example, *EGFR* mutations were significantly more common in cfDNA (11.2%) than tissue [4.8% for TCGA ($P < 0.003$); 4.4% for NHS/HPFS ($P < 0.0001$); and 2.9% for GENIE ($P < 0.0001$), by χ^2 test; Fig. 1A]. These differences may reflect real biological differences between the cfDNA and tissue-based cohorts, as patients who received cfDNA profiling were more likely to have metastatic disease and to have received prior therapies.

In addition, *JAK2*^{V617F} mutations were detected in the cfDNA of 16 patients, but none of the 1,996 combined colorectal cancer specimens profiled from the tissue-based sequencing databases (Fig. 1C). The most likely explanation for this discrepancy is that *JAK2*^{V617F} mutations detected in cfDNA were not actually present in the patients' tumors, but were present in a hematopoietic clone of indeterminate clinical potential, which is observed in 0.2% of the general population (14). Indeed, *JAK2*^{V617F} mutations are one of the most common mutations observed in hematopoietic clones of indeterminate clinical potential, and these mutations increase in prevalence with age (15). Consistent with this hypothesis, the median age for patients with a *JAK2*^{V617F} mutation detected in cfDNA was 73 years, whereas the median age among patients without this mutation was 60 years ($P = 0.0006$; Fig. 1D). This example highlights an important limitation of blood-based genomic profiling, in that one cannot be certain that a mutation detected in cfDNA is actually derived from the patient's tumor. Still, these data overall show a remarkably high similarity between cfDNA-based and tissue-based profiling, and as a whole support the potential utility and validity of large-scale cfDNA genomic profiling approaches.

Comparison of Clonal and Subclonal Mutations

We then evaluated the clonal versus subclonal landscape of mutation variants detected in the colorectal cancer cfDNA

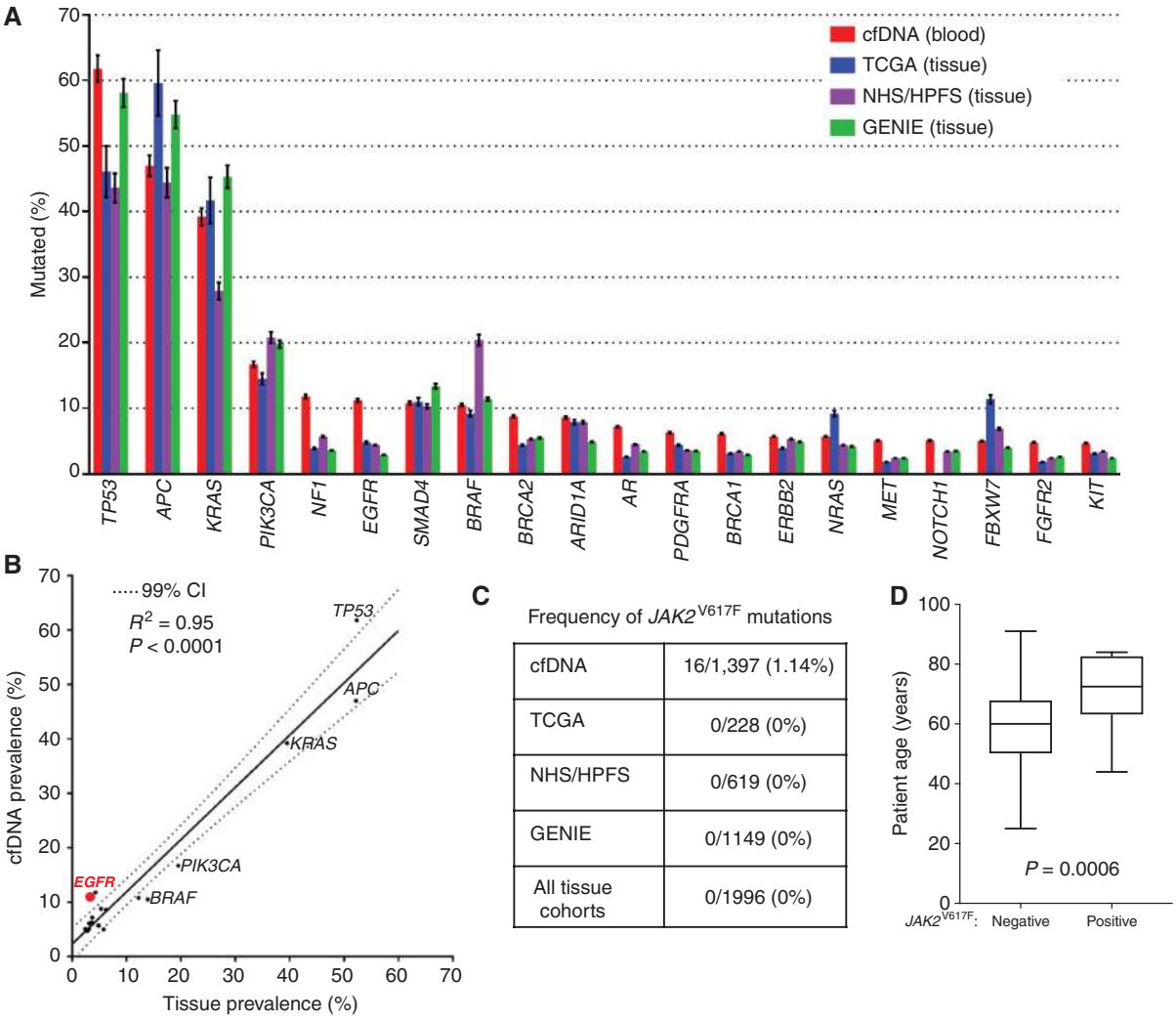

Figure 1. Genomic profiling by cfDNA or tumor tissue sequencing in colorectal cancer cohorts. **A,** Comparison of mutation frequencies in cfDNA and tissue cohorts (SNVs only). Top 20 gene mutations in cfDNA listed. Includes missense and nonsense mutations only (splice site mutations, insertions, and deletions excluded). **B,** Correlation between mutation frequencies in cfDNA versus tissue (top 20 genes in cfDNA listed). **C,** Comparison of *JAK2*^{V617F} mutation frequency in cfDNA and tumor tissue databases. **D,** Comparison of age between all patients with cfDNA profiling versus patients with detectable *JAK2*^{V617F} mutation in blood.

cohort. A mutation was defined as "subclonal" if the mutant allele frequency (MAF) was less than 25% of the highest MAF in the sample and was defined as "clonal" if it was above this threshold. At least one subclonal mutation was found in 51% of patients (range of 1 to 54 subclonal mutations). Among the 20 genes with the highest mutational prevalence in cfDNA, the six genes most likely to be clonal (in order of most to least clonal) include *KRAS, FBXW7, APC, SMAD4, BRAF*, and *TP53* (Fig. 2A), all of which are known to play early and critical roles in the oncogenesis of colorectal cancer. Furthermore, clonal SNVs were significantly more likely to represent mutations predicted to be activating or inactivating truncal driver mutations, whereas subclonal SNVs were more likely to be nonfunctioning "passenger" mutations or variants of unknown significance [OR, 3.65; 95% confidence interval (CI), 3.24–4.10; $P < 0.0001$; Fig. 2B].

We hypothesized that another class of subclonal mutations could represent nontruncal acquired resistance mutations emerging during prior therapy. Interestingly, *EGFR* was the gene with the highest percentage of subclonal mutations (Fig. 2A). When the MAFs of two predominantly clonal mutations, *KRAS* and *APC*, were compared in individual tumors, a linear relationship was observed (Fig. 2C), suggesting that these mutations often coexist as clonal events in colorectal cancer, although some subclonal *KRAS* and *APC* mutations were noted. However, when the MAFs of *EGFR* and *APC* SNVs were compared, a linear relationship was not observed, with most *EGFR* mutations occurring at subclonal frequencies (Fig. 2D). This suggests that the *EGFR* mutations detected in cfDNA are not likely to be founding clonal events in the development of these colorectal cancers, but rather are likely to be mutations emerging in specific tumor subclones, either as part of the process of tumor progression or metastasis, or perhaps in response to the selective pressure of anticancer therapies.

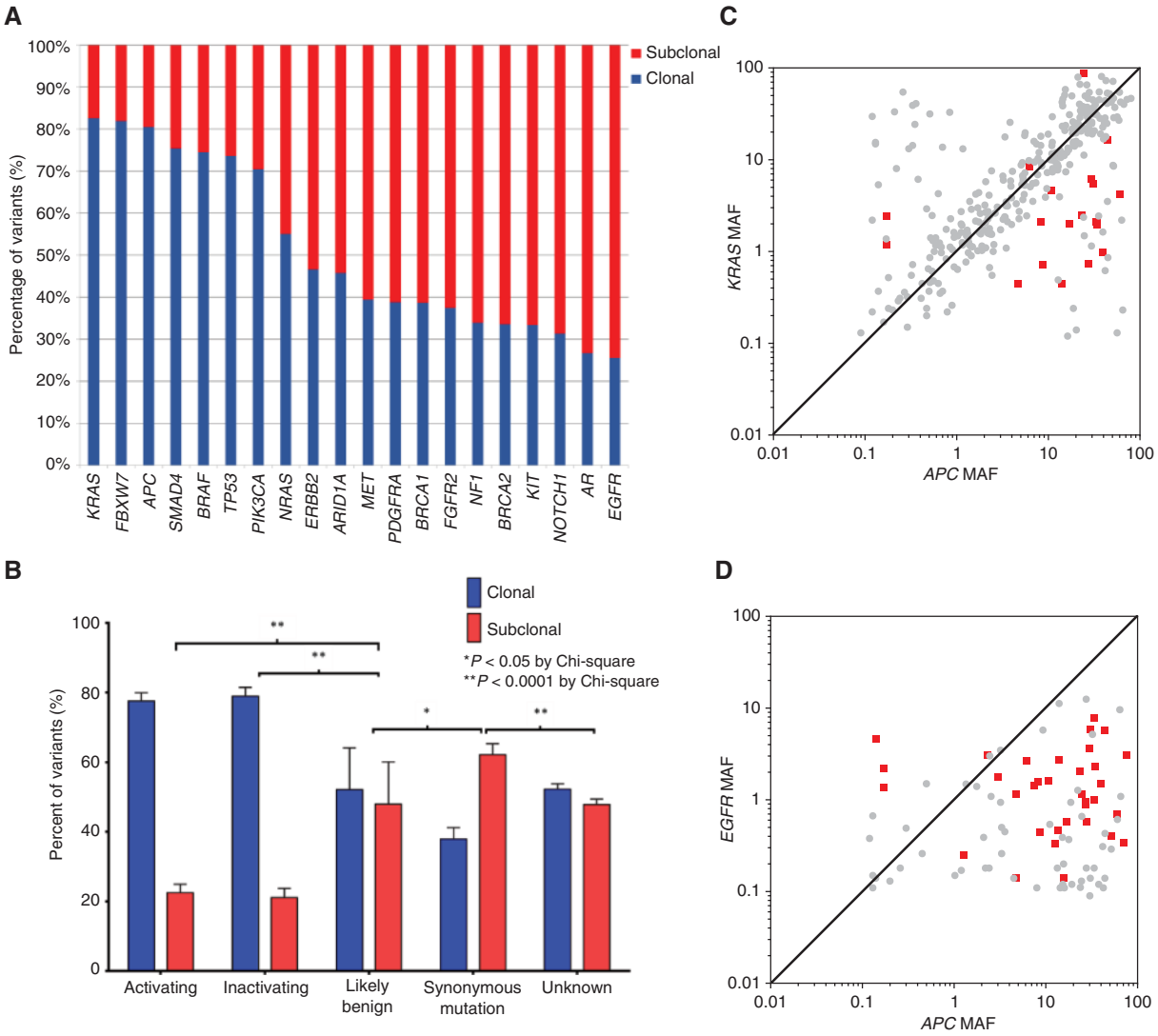

Figure 2. Clonality of common gene mutations in cfDNA from patients with colorectal cancer. **A,** Proportion of clonal versus subclonal mutations in commonly mutated genes in cfDNA (top 20 genes in cfDNA listed). **B,** Impact of variant functional significance on clonality of the alteration. **C,** Scatter plot of *KRAS* MAF versus *APC* MAF. Patients with *EGFR* ECD mutations are labeled in red. **D,** Scatter plot of *EGFR* MAF versus *APC* MAF. Patients with *EGFR* ECD mutations are labeled in red.

Consistent with this latter hypothesis, many of the patients with subclonal *EGFR* mutations harbor specific mutations in the extracellular domain (ECD) of EGFR. *EGFR* ECD mutations have been implicated in driving clinical acquired resistance to anti-EGFR antibodies, such as cetuximab and panitumumab, which are approved for the treatment of *KRAS* and *NRAS* (*RAS*) wild-type metastatic colorectal cancer (7, 12, 16–18). Similarly, many of the subclonal *KRAS* mutations observed also occurred in these same patients with *EGFR* ECD mutations (Fig. 2C, red labels). *KRAS* and *NRAS* mutations are another common mechanism of acquired resistance that can emerge during therapy with anti-EGFR antibodies, and prior studies have shown that subclonal *RAS* and *EGFR* ECD mutations can emerge in the same patient during anti-EGFR therapy (9, 19). The substantially higher mutation prevalence of *EGFR* mutations in cfDNA compared with tissue (Figs. 1A and B) may thus reflect the emergence of subclonal resistance mutations, particularly in the setting of acquired EGFR antibody resistance.

EGFR ECD Mutations in cfDNA of Patients with Colorectal Cancer

To evaluate the potential for large-scale cfDNA sequencing to provide clinically relevant insights into therapeutic resistance, we performed a focused analysis on patients with *EGFR* mutations. In total, 85 of 157 patients with *EGFR* mutations had mutations in the ECD (amino acids 1 through 649; Fig. 3A). To identify mutations with the greatest likelihood of functional relevance, we focused on *EGFR* ECD mutations that were recurrent (observed in more than one patient) in this cohort. In total, 58 patients harbored a recurrent *EGFR* ECD mutation, and in 42 patients, these *EGFR* ECD mutations

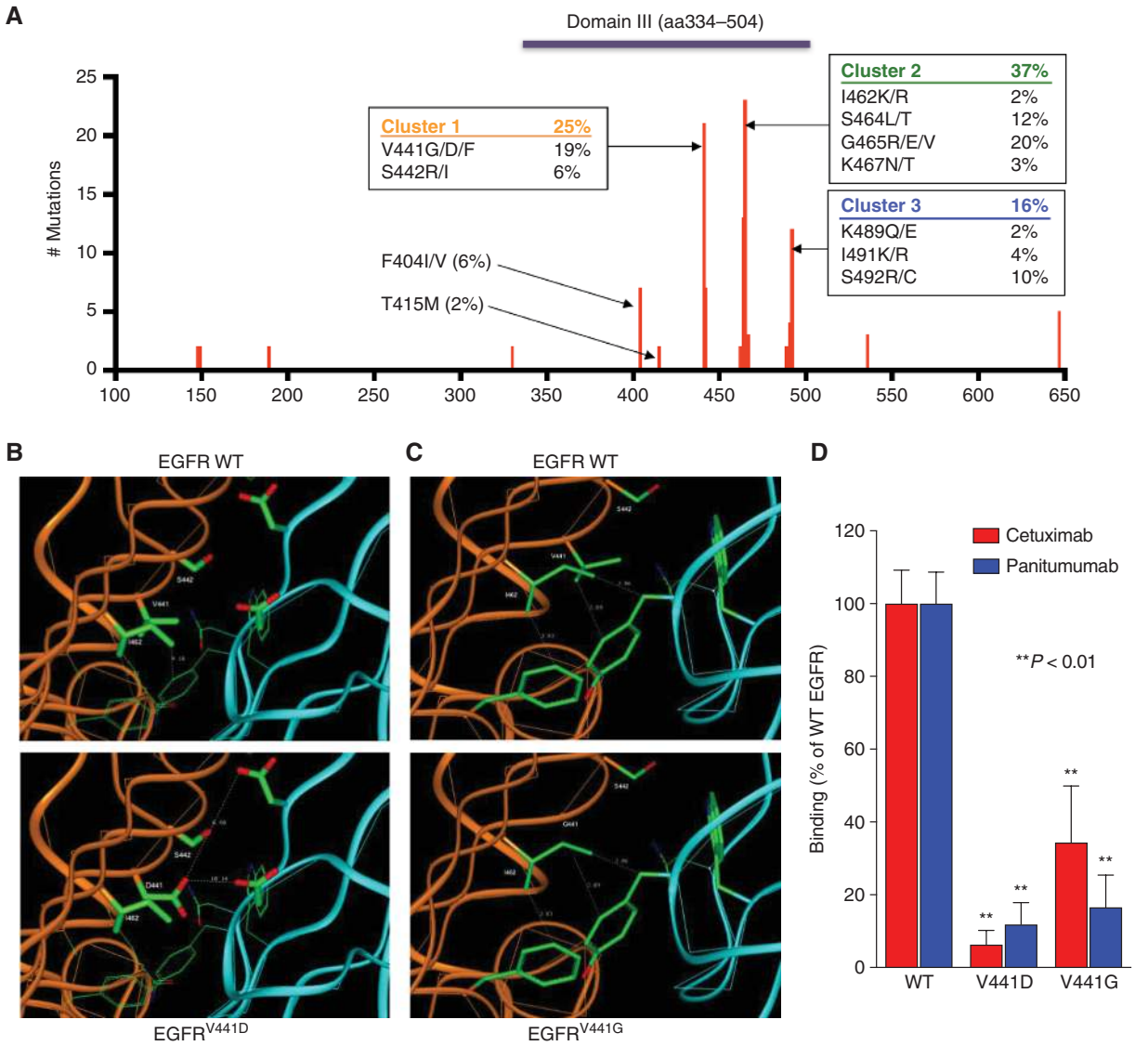

Figure 3. EGFR ECD mutations in cfDNA. **A,** *EGFR* ECD mutations occurring more than once in cfDNA cohort. Three dominant clusters of amino acid substitutions in the binding domain of anti-EGFR antibodies are shown. **B,** Molecular model of cetuximab bound to wild-type EGFR V441 and EGFR V441D. **C,** Molecular model of cetuximab bound to wild-type EGFR V441 and EGFR V441G. **D,** Binding assay of cetuximab and panitumumab to wild-type, V441D, and V441G EGFR. **, $P < 0.01$ by one-way ANOVA with Tukey *post hoc* test for cetuximab and panitumumab.

occurred in domain III of EGFR (amino acids 334–504), which represents the binding epitope of both cetuximab and panitumumab (Fig. 3A). Of these 42 patients, all 24 patients (57%) with available treatment histories were confirmed to have received anti-EGFR antibody treatment prior to blood collection for cfDNA analysis, supporting a likely role for these mutations in driving therapeutic resistance. Treatment histories for the remaining 18 patients were not available.

Analysis of the *EGFR* ECD mutations from these 42 patients revealed 23 distinct mutations in 11 amino acids in domain III. Seven of these amino acids clustered into two regions, Cluster 2 (I462, S464, G465, K467) and Cluster 3 (K489, I491, S492), previously reported to be associated with cetuximab and/or panitumumab resistance (Fig. 3A; refs. 7, 10, 12, 18). Mutations in two amino acids, F404 and T415, were noted in a small percentage of patients and have not been previously reported in anti-EGFR antibody resistance. However, mutations in two additional amino acids not previously implicated in anti-EGFR antibody resistance, V441 and S442, formed a prominent new cluster (Cluster 1) accounting for 25% of all *EGFR* ECD mutations. The majority of these novel but highly recurrent mutations affected V441, with V441D and V441G being the most common amino acid changes observed.

To explore the hypothesis that V441D and V441G represent novel mechanisms of acquired resistance to anti-EGFR antibodies, we performed molecular modeling to predict the effects of these mutations on cetuximab binding to

EGFR. Compared with wild-type EGFR, the V441D mutation introduces a negatively charged residue that is predicted to decrease the interaction between cetuximab and EGFR (Fig. 3B). Similarly, the V441G mutation is predicted to destroy a critical hydrophobic node, decreasing the interaction between cetuximab and EGFR (Fig. 3C). Consistent with these models, both the V441D and V441G EGFR mutants showed significantly reduced binding of both cetuximab and panitumumab, relative to wild-type EGFR (Fig. 3D), supporting their role as novel mechanisms of acquired resistance to anti-EGFR antibodies.

Heterogeneity of anti-EGFR Antibody Resistance

To further investigate the impact of tumor heterogeneity on therapeutic strategies to overcome anti-EGFR antibody resistance, we performed an in-depth analysis of the cfDNA profiles from the 42 patients with *EGFR* ECD mutations. In addition to *EGFR* ECD mutations, multiple mechanisms of acquired resistance to anti-EGFR antibodies that bypass the need for EGFR signaling have been previously identified, including alterations in *KRAS, NRAS, BRAF, MAP2K1, ERBB2, MET*, and *KIT* (5–7, 10–12, 16, 18, 20–25). We observed that in 91% of patients with *EGFR* ECD mutations in cfDNA, at least one additional cooccurring resistance alteration was also detected in cfDNA. On average, these patients harbored 5 distinct resistance alterations to anti-EGFR antibodies (median 4), with as many as 13 co-occurring resistance alterations being detected in a single patient, indicating a striking degree of heterogeneity (Fig. 4A). These co-occurring resistance alterations frequently involved multiple functionally distinct gene targets with 54 unique alterations noted affecting eight different genes. *KRAS* alterations (including mutations and amplifications) were the most common resistance alterations observed, present in 69% of patients. *KRAS* mutations occurred in 62% of patients, with $KRAS^{Q61H}$ being the most common, observed in 52% of patients. *MET* amplification, *NRAS* mutations, and *KRAS* amplification were the next most common resistance alterations observed, present in 38%, 33%, and 24% of cases, respectively. *BRAF* and *MAP2K1* mutations were each present in 9.5% of cases.

The presence of multiple resistance alterations in the cfDNA of an individual patient is thought to represent the existence of multiple resistant tumor subclones that have emerged during therapy, which perhaps reside in different metastases throughout the body. Consistent with this model, resistance mutations were significantly more likely to be subclonal in patients with multiple concurrent mechanisms of resistance compared with patients with only a single mechanism of resistance to anti-EGFR therapy (54.5% vs. 10.0%; $P < 0.0001$; OR, 10.81; 95% CI, 7.83–14.93; Supplementary Fig. S2). Indeed, we observed that patients with *EGFR* ECD mutations showed an enrichment of subclonal *EGFR* and *KRAS* mutations relative to the overall cohort (Figs. 2C and D, red labels). Our analysis also revealed profound heterogeneity occurring in the context of acquired resistance. For example, one patient (Fig. 4B, pt #1) harbored 13 distinct resistance alterations, including four *EGFR* ECD mutations, four *KRAS* mutations, *KRAS* amplification, two *NRAS* mutations, *ERBB2* amplification, and a downstream mutation affecting MEK1 (encoded by *MAP2K1*). Another patient (Fig. 4B, pt #8) harbored eight different resistance alterations, including two *EGFR* ECD mutations, as well as *MET* amplification, *ERBB2* amplification, *KIT* amplification, $KRAS^{Q61H}$, and $BRAF^{V600E}$. A third patient (Fig. 4B, pt #12) harbored seven different resistance alterations, including two *EGFR* ECD mutations, *MET* amplification, and *KRAS, NRAS, BRAF*, and *MAP2K1* mutations. In another patient with serial cfDNA analyses and detailed treatment history, multiple resistance alterations were observed to emerge during anti-EGFR therapy (Supplementary Fig. S3). This degree of heterogeneity and the co-occurrence of multiple distinct resistance alterations present a daunting challenge for therapeutic strategies designed to overcome resistance to anti-EGFR antibodies.

In an effort to overcome acquired resistance, anti-EGFR antibody mixtures, such as Sym004 and MM-151, have been developed that are capable of binding multiple epitopes on EGFR and can thereby overcome the effects of individual *EGFR* ECD resistance mutations. However, these agents may not overcome non-ECD resistance alterations that bypass the requirement for EGFR signaling via activation at other points in the RAS–RAF–MEK–ERK pathway. Importantly, of the 42 patients harboring domain III *EGFR* ECD mutations identified in our analysis, 88% harbored at least one additional non-ECD resistance alteration detectable in cfDNA that would be predicted to drive resistance to an EGFR antibody mixture alone (Fig. 4C), with an average of 2.9 non-ECD resistance alterations per patient (range, 1–9; median, 3).

DISCUSSION

With the growing utilization of clinical cfDNA testing, large cfDNA sequencing databases could represent a valuable resource for genomic discovery. Here, we present one of the first studies assessing whether large-scale genomic profiling of cfDNA can accurately reproduce the genomic landscape of driver mutations defined by direct tumor tissue sequencing studies. In our analysis of the cfDNA profiles of 1,397 patients with colorectal cancer, we find that the spectrum and frequency of genomic alterations identified in cfDNA demonstrate a striking similarity to results from three large colorectal cancer tumor tissue sequencing cohorts. These data provide a key proof-of-concept supporting the feasibility and validity of large-scale genomic analysis of cfDNA.

Genomic profiling from cfDNA offers potential advantages and limitations when compared with tumor tissue–based sequencing. One key limitation to tumor tissue sequencing is that up to 25% of patients with advanced cancer may have insufficient quantity of tissue available for molecular analysis (2). In contrast, as peripheral blood can be collected efficiently and noninvasively, cfDNA testing can be performed on almost any patient, including those with tumor lesions that are difficult to biopsy. Furthermore, as tumor tissue sequencing often relies on archival tissue obtained prior to the development of metastatic disease, cfDNA profiling may more readily facilitate analysis of patients with metastatic disease and may better capture the presence of tumor heterogeneity.

However, cfDNA profiling also has limitations. In this dataset, genomic alterations in cfDNA were not detected in 15% of cases. This result is similar to rates of cfDNA detection in other colorectal cancer series (5, 19), and is comparable to

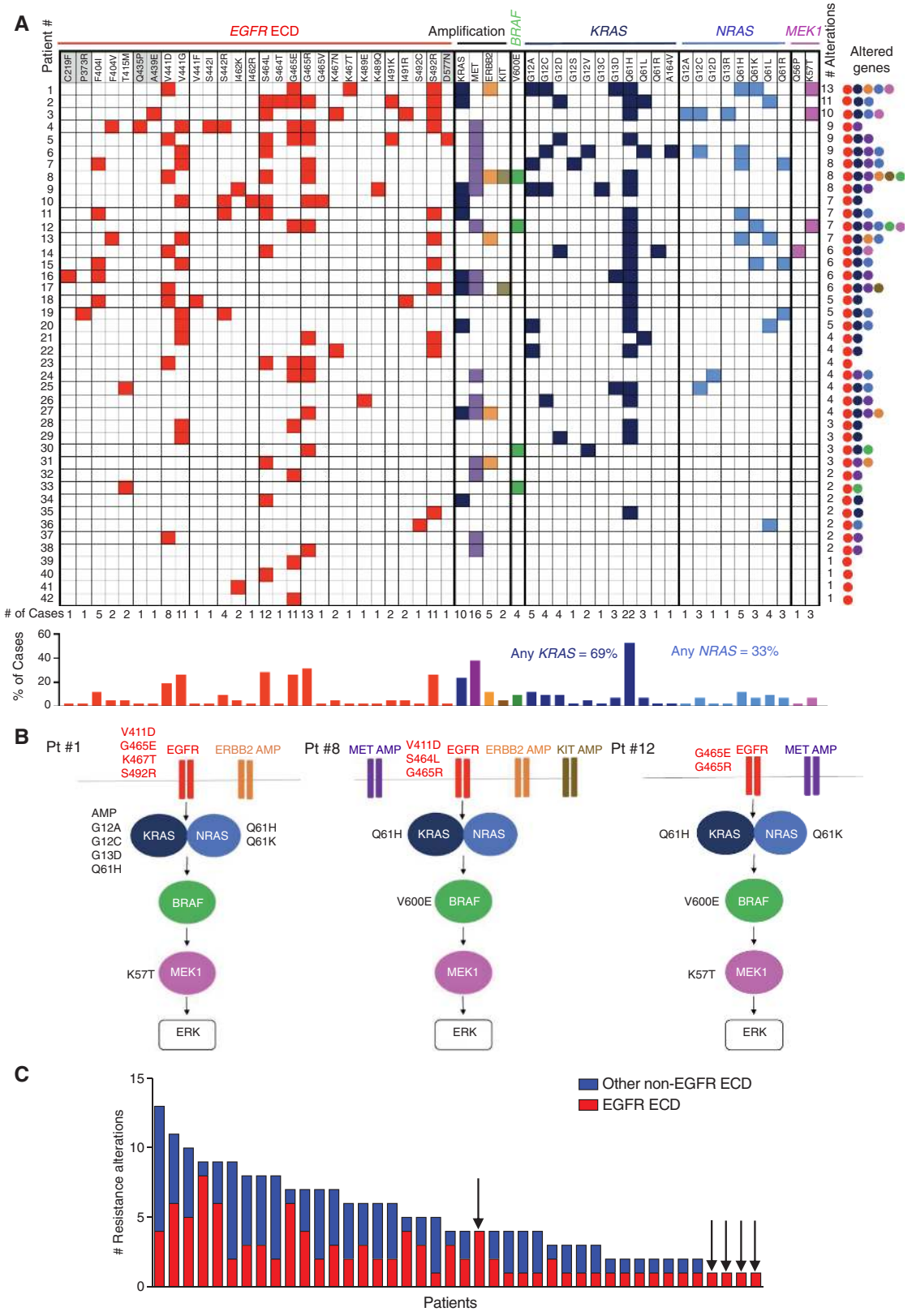

Figure 4. Heterogeneity of anti-EGFR resistance alterations in patients with *EGFR* ECD mutations. **A,** Top, known anti-EGFR antibody resistance alterations identified in cfDNA for patients with *EGFR* ECD mutations, with each row representing an individual patient. Bottom, percentage of cases with alteration, with each bar representing mutational frequency. Shaded EGFR ECD mutations represent those observed only once in the overall cohort. **B,** Case examples of patients with multiple EGFR pathway alterations. **C,** Number of ECD (blue) and non-ECD (red) resistance alterations identified in cfDNA for each patient. Arrows indicate patients with *EGFR* ECD mutations only.

the rate of tissue insufficiency in tissue-based NGS profiling (2). Although it is possible that some patients did not have alterations in genes covered by the NGS assay, in most cases, lack of detection of genomic alterations in cfDNA was likely due to other factors, including low tumor burden, lack of cfDNA shedding by some tumors, and timing of blood collection (ctDNA is reduced after surgical resection and while on active treatment; ref. 26). Optimizing the timing of cfDNA testing, for example, prior to initiation of therapy or at the time of disease progression, may be an important means of increasing the yield of cfDNA testing.

Another limitation of cfDNA profiling highlighted in our study is that it is possible for both tumor-derived and non–tumor-derived genomic alterations to be detected in cfDNA, which has the potential to confound analyses (27). For example, in our study, *JAK2*^{V617F} mutations were detected in cfDNA from 1.1% of patients, but in none of the colorectal cancer cases from tissue-based cohorts (Fig. 1C). These *JAK2*^{V617F} mutations are most likely derived from a hematopoietic clone of indeterminate clinical potential, which is observed subclinically in the peripheral blood in a small percentage of the general population, but with increasing prevalence with age (14). Indeed, these mutations were found predominantly in older patients in our cohort (Fig. 1D). Similarly, a recent case report identified an *IDH2* mutation in cfDNA from a patient with metastatic colorectal cancer that was not readily detectable in a matched tumor biopsy. The same *IDH2* mutation was identified in a bone marrow biopsy, supporting that the alteration originated from a hematopoietic clone (28). Going forward, methods such as parallel sequencing of mononuclear cells isolated from peripheral blood may help to delineate whether specific alterations detected in cfDNA are derived from clonal hematopoiesis or tumor (27).

In addition to illustrating the close relationship between the genomic landscape of cfDNA and tissue, our study also offers unique insights into therapeutic resistance. Because tissue-based sequencing compendia rely primarily on early-stage and treatment-naïve tumors, these databases have generated limited insights into acquired resistance. Conversely, large cfDNA cohorts, which can more readily provide non-invasive access to patients with advanced disease, may offer unique insight into resistance mechanisms emerging under the selective pressure of systemic therapies. For example, the potential for *EGFR* ECD mutations to drive resistance to anti-EGFR antibodies has been documented through cfDNA and tumor biopsies from small patient cohorts (7, 10, 12, 16–18). These efforts have identified key amino acid mutations that drive acquired resistance, including I462, S464, G465, K467 (Cluster 2) and K489, I491, S492 (Cluster 3; refs. 7, 18). Our cfDNA database analysis confirmed the recurrent alteration of these previously identified residues, but also identified a previously unreported cluster of *EGFR* ECD mutations involving V441 and S442 (Cluster 1) that accounted for 25% of all ECD mutations, representing an important and novel mechanism of resistance to EGFR blockade.

A key limitation of our study is the lack of clinical annotation for the cfDNA cohort. Indeed, as treatment history was not available for some patients with *EGFR* ECD mutations, it is not possible to confirm that all patients had received prior anti-EGFR antibody therapy. However, *EGFR* ECD mutations have not been observed in colorectal cancer prior to EGFR blockade, and, accordingly, in the 57% of *EGFR* ECD patients with available treatment history, every patient was confirmed to have received prior anti-EGFR therapy. Although these factors support the likelihood that these alterations emerged in the setting of acquired resistance to anti-EGFR therapy, the absence of paired baseline samples makes this impossible to confirm. Still, it is notable that even without detailed clinical data, this analysis was able to generate new insights into resistance in colorectal cancer. Collectively, these findings demonstrate the potential of large-scale cfDNA profiling as a tool for discovery and underscore the potential benefits of ongoing academic efforts to create publicly available cfDNA databases with clinical annotation for future studies (29).

Our study also provides key insights into the role of tumor heterogeneity in the setting of acquired resistance to anti-EGFR antibodies in colorectal cancer. Previous studies have illustrated that multiple, heterogeneous resistance alterations can be detected in the cfDNA of individual patients, which are thought to represent the existence of multiple resistant tumor subclones, often residing in different metastases throughout the body (5, 7, 8, 16, 18, 19, 24). Our study supports the frequent cooccurrence of multiple resistance alterations in individual patients following EGFR blockade, but suggests that the degree of molecular heterogeneity present may be even more profound and complex than anticipated. Indeed, we observed that patients harbored an average of 5 unique resistance alterations (median 4), with as many as 13 distinct resistance alterations observed in a single patient. In only 9% of patients was a single resistance alteration detected.

The degree of molecular heterogeneity observed following anti-EGFR therapy highlights the difficulty of devising a single therapeutic strategy capable of overcoming a broad array of resistance mechanisms, particularly as these alterations frequently affected multiple functionally distinct targets in an individual patient. These findings have profound clinical implications for efforts designed to overcome *EGFR* ECD mutations by binding multiple epitopes on EGFR (16, 30, 31). Indeed, our results suggest that the percentage of patients who harbor *EGFR* ECD mutations alone may be exceedingly small, only 12% in this limited series (Fig. 4C). Therefore, tumor heterogeneity at the time of acquired resistance to EGFR blockade represents a significant obstacle to the development of precision medicine strategies and suggests that therapeutic strategies that target a key convergent signaling node capable of overcoming the multiplicity of resistant clones present in an individual patient may be required (9, 18). Collectively, these studies support the potential utility of large-scale cfDNA profiling databases to define the genomic landscape of patients with cancer and to provide novel and clinically relevant insights into tumor heterogeneity and therapeutic resistance.

METHODS

cfDNA Cohort

A total of 1,772 consecutive blood specimens were analyzed from patients with colorectal cancer using the Guardant360 cfDNA assay (Guardant Health). The Guardant360 assay is a Clinical Laboratory Improvement Amendments (CLIA)–certified targeted digital sequencing panel designed to detect SNVs, as well as selected insertions/deletions, amplifications, and fusions (Supplementary Table S1).

Subjects provided informed, written consent when appropriate. This research was conducted in accordance with the Declaration of Helsinki and was performed with Institutional Review Board approval (MDACC LAB09-0373).

Analysis of cfDNA Mutational Prevalence

To determine cfDNA mutational prevalence, we analyzed samples in which at least one mutation was detected ($N = 1{,}500$). When more than one sample was available for the same patient, the most recent sample was selected for analysis. The final cfDNA analysis cohort included samples from 1,397 unique patients. Mutations considered in this analysis include SNPs that resulted in protein-coding changes (i.e., missense/nonsense). Synonymous, splice site, intron, intergenic, and untranslated region variants were not included in the prevalence calculation. Insertions/deletions (indels), fusions, and amplifications/deletions were also excluded. The primary focus for this analysis is on the 20 genes with the highest mutation prevalence.

Analysis of Tissue Mutational Prevalence

To compare cfDNA mutational prevalence to that of tissue-based datasets, we obtained sequencing results from three previously reported and publicly available tumor tissue cohorts. The TCGA cohort consisted of 228 unique cases with tumor and matched normal tissue pairs with whole-exome sequencing (WES) from the supplement of the original TCGA characterization of colon and rectal cancer (3). The NHS/HPFS cohort included WES of tumor and matched normal tissue pairs from 619 untreated colorectal cancers (4). The third cohort consisted of colorectal cancer cases from the AACR Project GENIE (13). Because profiling techniques differ across centers within the AACR Project GENIE, we limited our analysis to those centers that analyzed all exons of sequenced genes. This allowed prevalence calculations based on gene coverage comparable with the Guardant360 assay. The proportions and 99% CIs of specific gene mutations were computed for all tissue and cfDNA cohorts. Ninety-nine percent confidence intervals for binomial proportions were calculated for each gene mutation based on the frequencies and numbers of patients studied using the modified Wald method (32). CIs were compared descriptively between the cfDNA cohort and each tissue-based cohort. Median age was compared between patients with and without a $JAK2^{V617F}$ mutation detected in cfDNA using the *t* test. Bar charts and scatter plots were used to illustrate the data.

Analysis of Clonal and Subclonal Mutations

The 20 genes with the highest mutation prevalence were analyzed for percentage of variants that were either clonal or subclonal. This analysis included all nonsynonymous SNV gene variants. In this analysis, when serial samples were available for the same patient, the sample with the highest MAF was selected as the most representative for analysis. A mutation was defined as clonal if the MAF was greater than or equal to 25% of the highest MAF in the sample. A mutation was defined as "subclonal" if the MAF was less than 25% of the highest MAF in the sample.

Analysis of Functional Significance of Alterations

Functional annotation of variants for Fig. 2B was provided by the Precision Oncology Decision Support Core at Sheikh Khalifa Bin Zayed Al Nahyan Institute for Personalized Cancer Therapy of The University of Texas MD Anderson Cancer Center (Houston, TX; https://pct.mdanderson.org/). This annotation utilizes a large curated database of variants with literature based and *in vitro* assessment of functional significance (33, 34).

Determination of EGFR Antibody Binding

NIH3T3 cells were obtained from the ATCC (CRL-1658) in 2016 and were passaged for less than 6 months after receipt. Full-length human *EGFR* cDNA (Addgene 23935) was mutagenized using the Stratagene Site-Directed Mutagenesis Kit to produce V441D and V441G mutations. *EGFR*-containing pLenti-puro (Addgene 17452) was used to produce virus to infect NIH3T3 cells. Cells were infected with empty pLenti plasmid, wild-type *EGFR*, $EGFR^{V441D}$, and $EGFR^{V441G}$, and analyzed by flow cytometry for antibody binding as performed by others (17). Cells were analyzed by Western blot analysis to demonstrate equal levels of EGFR expression. 1×10^6 cells of each type were resuspended in 100 μL of 1% BSA in PBS and incubated with 1 μg of either cetuximab or panitumumab for 1 hour at 4°C. Cells were washed with 1% BSA in PBS and incubated with 1 μg of anti-human PE-conjugated secondary antibody (Thermo Fisher Scientific H10104). Cells were washed and subsequently read on an LSRII flow cytometer. The antibody binding experiment was performed in triplicate. Molecular modeling was conducted using the program Insight II (Accelrys).

Disclosure of Potential Conflicts of Interest

M. McKinney is a consultant/advisory board member for Kite Pharmaceuticals and Spectrum Pharmaceuticals. W.M. Korn is a consultant/advisory board member for Guardant Health, Lilly, and Merck. C.E. Atreya reports receiving commercial research support from Bristol-Myers Squibb, Guardant Health, Merck, and Novartis, is a consultant/advisory board member for Kura Oncology and Roche/Genentech, and has provided expert testimony for Cancer Panels and Champions Oncology. K.C. Banks has ownership interest (including patents) in Guardant Health. F. Meric-Bernstam reports receiving commercial research grants from Aileron, AstraZeneca, Bayer, Calithera, Curis, CytoMx, Debiopharma, eFFECTOR, Genentech, Novartis, Pfizer, PUMA, Taiho, and Zymeworks, and is a consultant/advisory board member for Clearlight Diagnostics, Darwin Health, Dialecta, GRAIL, Inflection Biosciences, Pieris, and Sumitomo Dainippon Pharma. R.B. Lanman has ownership interest (including patents) in Guardant Health. R.B. Corcoran reports receiving commercial research grants from AstraZeneca and Sanofi, and is a consultant/advisory board member for Amgen, Astex, Avidity Biosciences, BMS, Genentech, Loxo, Merrimack, N-of-one, Roche, Shire, and Taiho. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: J.H. Strickler, K.C. Banks, R.J. Nagy, R.B. Lanman, R.B. Corcoran, S. Kopetz

Development of methodology: J.H. Strickler, J.M. Loree, L.G. Ahronian, M. McKinney, R.B. Lanman, R.B. Corcoran, S. Kopetz

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J.H. Strickler, L.G. Ahronian, A.R. Parikh, A.A.L. Pereira, W.M. Korn, C.E. Atreya, K.C. Banks, R.J. Nagy, F. Meric-Bernstam, R.B. Lanman, R.B. Corcoran, S. Kopetz

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J.H. Strickler, J.M. Loree, L.G. Ahronian, A.R. Parikh, D. Niedzwiecki, A.A.L. Pereira, M. McKinney, W.M. Korn, R.J. Nagy, R.B. Lanman, I.F. Tsigelny, R.B. Corcoran, S. Kopetz

Writing, review, and/or revision of the manuscript: J.H. Strickler, J.M. Loree, A.R. Parikh, D. Niedzwiecki, A.A.L. Pereira, M. McKinney, C.E. Atreya, K.C. Banks, R.J. Nagy, F. Meric-Bernstam, R.B. Lanman, R.B. Corcoran, S. Kopetz

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J.H. Strickler, J.M. Loree, A.A.L. Pereira, K.C. Banks, R.J. Nagy, R.B. Corcoran, S. Kopetz

Study supervision: J.H. Strickler, A. Talasaz, R.B. Corcoran, S. Kopetz

Acknowledgments

The work is supported by NIH R01 CA184843 (to S. Kopetz), CPRIT grant RP150535 (to F. Meric-Bernstam), an ASCO Young Investigator Award, Canadian Association of Medical Oncologists

Research Fellowship, and Detweiler Fellowship (to J.M. Loree), NIH/NCI K08CA175143 (to C.E. Atreya), and a Damon Runyon Clinical Investigator Award, NIH/NCI Gastrointestinal Cancer SPORE P50 CA127003, R01CA208437, and K08CA166510. Research was supported by a Stand Up To Cancer Colorectal Cancer Dream Team Translational Research Grant (grant number SU2C-AACR-DT22-17; to R.B. Corcoran). Research grants are administered by the American Association for Cancer Research, the scientific partner of SU2C.

Received September 8, 2017; revised November 14, 2017; accepted November 29, 2017; published OnlineFirst December 1, 2017.

REFERENCES

1. Kris MG, Johnson BE, Berry LD, Kwiatkowski DJ, Iafrate AJ, Wistuba II, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. JAMA 2014;311:1998–2006.
2. Meric-Bernstam F, Brusco L, Shaw K, Horombe C, Kopetz S, Davies MA, et al. Feasibility of large-scale genomic testing to facilitate enrollment onto genomically matched clinical trials. J Clin Oncol 2015;33:2753–62.
3. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487:330–7.
4. Giannakis M, Mu XJ, Shukla SA, Qian ZR, Cohen O, Nishihara R, et al. Genomic correlates of immune-cell infiltrates in colorectal carcinoma. Cell Rep 2016;17:1206.
5. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014;6:224ra24.
6. Diaz LA Jr, Williams RT, Wu J, Kinde I, Hecht JR, Berlin J, et al. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature 2012;486:537–40.
7. Arena S, Bellosillo B, Siravegna G, Martinez A, Canadas I, Lazzari L, et al. Emergence of multiple EGFR extracellular mutations during cetuximab treatment in colorectal cancer. Clin Cancer Res 2015;21:2157–66.
8. Kwak EL, Ahronian LG, Siravegna G, Mussolin B, Godfrey JT, Clark JW, et al. Molecular heterogeneity and receptor coamplification drive resistance to targeted therapy in MET-amplified esophagogastric cancer. Cancer Discov 2015;5:1271–81.
9. Misale S, Di Nicolantonio F, Sartore-Bianchi A, Siena S, Bardelli A. Resistance to anti-EGFR therapy in colorectal cancer: from heterogeneity to convergent evolution. Cancer Discov 2014;4:1269–80.
10. Pietrantonio F, Vernieri C, Siravegna G, Mennitto A, Berenato R, Perrone F, et al. Heterogeneity of acquired resistance to anti-EGFR monoclonal antibodies in patients with metastatic colorectal cancer. Clin Cancer Res 2017;23:2414–22.
11. Russo M, Siravegna G, Blaszkowsky LS, Corti G, Crisafulli G, Ahronian LG, et al. Tumor heterogeneity and lesion-specific response to targeted therapy in colorectal cancer. Cancer Discov 2016;6:147–53.
12. Siravegna G, Mussolin B, Buscarino M, Corti G, Cassingena A, Crisafulli G, et al. Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. Nat Med 2015;21:795–801.
13. AACR Project GENIE Consortium. AACR project GENIE: powering precision medicine through an international consortium. Cancer Discov 2017;7:818–31.
14. Genovese G, Kahler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med 2014;371:2477–87.
15. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med 2014;371:2488–98.
16. Bertotti A, Papp E, Jones S, Adleff V, Anagnostou V, Lupo B, et al. The genomic landscape of response to EGFR blockade in colorectal cancer. Nature 2015;526:263–7.
17. Montagut C, Dalmases A, Bellosillo B, Crespo M, Pairet S, Iglesias M, et al. Identification of a mutation in the extracellular domain of the epidermal growth factor receptor conferring cetuximab resistance in colorectal cancer. Nat Med 2012;18:221–3.
18. Van Emburgh BO, Arena S, Siravegna G, Lazzari L, Crisafulli G, Corti G, et al. Acquired RAS or EGFR mutations and duration of response to EGFR blockade in colorectal cancer. Nat Commun 2016;7:13665.
19. Morelli MP, Overman MJ, Dasari A, Kazmi SM, Mazard T, Vilar E, et al. Characterizing the patterns of clonal selection in circulating tumor DNA from patients with colorectal cancer refractory to anti-EGFR treatment. Ann Oncol 2015;26:731–6.
20. Bertotti A, Migliardi G, Galimi F, Sassi F, Torti D, Isella C, et al. A molecularly annotated platform of patient-derived xenografts ("xenopatients") identifies HER2 as an effective therapeutic target in cetuximab-resistant colorectal cancer. Cancer Discov 2011;1:508–23.
21. Martin V, Landi L, Molinari F, Fountzilas G, Geva R, Riva A, et al. HER2 gene copy number status may influence clinical efficacy to anti-EGFR monoclonal antibodies in metastatic colorectal cancer patients. Br J Cancer 2013;108:668–75.
22. Raghav K, Morris V, Tang C, Morelli P, Amin HM, Chen K, et al. MET amplification in metastatic colorectal cancer: an acquired response to EGFR inhibition, not a de novo phenomenon. Oncotarget 2016;7:54627–31.
23. Raghav KP, Overman MJ, Yu R, Meric-Bernstam F, Menter D, Kee BK, et al. HER2 amplification as a negative predictive biomarker for anti-epidermal growth factor receptor antibody therapy in metastatic colorectal cancer. J Clin Oncol 34:15s, 2016(suppl; abstr 3517).
24. Bardelli A, Corso S, Bertotti A, Hobor S, Valtorta E, Siravegna G, et al. Amplification of the MET receptor drives resistance to anti-EGFR therapies in colorectal cancer. Cancer Discov 2013;3:658–73.
25. Misale S, Yaeger R, Hobor S, Scala E, Janakiraman M, Liska D, et al. Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. Nature 2012;486:532–6.
26. Sacher AG, Paweletz C, Dahlberg SE, Alden RS, O'Connell A, Feeney N, et al. Prospective validation of rapid plasma genotyping for the detection of EGFR and KRAS mutations in advanced lung cancer. JAMA Oncol 2016;2:1014–22.
27. Razavi P, Li BT, Hou C, Shen R, Venn O, Lim RS, et al. Cell-free DNA (cfDNA) mutations from clonal hematopoiesis: implications for interpretation of liquid biopsy tests. J Clin Oncol 35:15s, 2017(suppl; abstr 11526).
28. Zhang BM, Aleshin A, Lin CY, Ford J, Zehnder JL, Suarez CJ. IDH2 mutation in a patient with metastatic colon cancer. N Engl J Med 2017;376:1991–2.
29. Grossman RL, Abel B, Angiuoli S, Barrett JC, Bassett D, Bramlett K, et al. Collaborating to compete: blood profiling atlas in cancer (BloodPAC) consortium. Clin Pharmacol Ther 2017;101:589–92.
30. Arena S, Siravegna G, Mussolin B, Kearns JD, Wolf BB, Misale S, et al. MM-151 overcomes acquired resistance to cetuximab and panitumumab in colorectal cancers harboring EGFR extracellular domain mutations. Sci Transl Med 2016;8:324ra14.
31. Sanchez-Martin FJ, Bellosillo B, Gelabert M, Dalmases A, Canadas I, Vidal J, et al. The first-in-class anti-EGFR antibody mixture Sym004 overcomes cetuximab-resistance mediated by EGFR extracellular domain mutations in colorectal cancer. Clin Cancer Res 2016;22:3260–7.
32. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. Am Stat 1998;52:119–26.
33. Johnson A, Zeng J, Bailey AM, Holla V, Litzenburger B, Lara-Guerra H, et al. The right drugs at the right time for the right patient: the MD Anderson precision oncology decision support platform. Drug Discov Today 2015;20:1433–8.
34. Kurnit KC, Bailey AM, Zeng J, Johnson AM, Shufean MA, Brusco L, et al. A publically available precision oncology resource. Cancer Res 2017;77:e123–6.