

Supplementary Materials for

Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies

Chetan Bettegowda, Mark Sausen, Rebecca J. Leary, Isaac Kinde, Yuxuan Wang, Nishant Agrawal, Bjarne R. Bartlett, Hao Wang, Brandon Lubber, Rhoda M. Alani, Emmanuel S. Antonarakis, Nilofer S. Azad, Alberto Bardelli, Henry Brem, John L. Cameron, Clarence C. Lee, Leslie A. Fecher, Gary L. Gallia, Peter Gibbs, Dung Le, Robert L. Giuntoli, Michael Goggins, Michael D. Hogarty, Matthias Holdhoff, Seung-Mo Hong, Yuchen Jiao, Hartmut H. Juhl, Jenny J. Kim, Giulia Siravegna, Daniel A. Laheru, Calogero Lauricella, Michael Lim, Evan J. Lipson, Sueli Kazue Nagahashi Marie, George J. Netto, Kelly S. Oliner, Alessandro Olivi, Louise Olsson, Gregory J. Riggins, Andrea Sartore-Bianchi, Kerstin Schmidt, le-Ming Shih, Sueli Mieko Oba-Shinjo, Salvatore Siena, Dan Theodorescu, Jeanne Tie, Timothy T. Harkins, Silvio Veronese, Tian-Li Wang, Jon D. Weingart, Christopher L. Wolfgang, Laura D. Wood, Dongmei Xing, Ralph H. Hruban, Jian Wu, Peter J. Allen, C. Max Schmidt, Michael A. Choti, Victor E. Velculescu,* Kenneth W. Kinzler,* Bert Vogelstein,* Nickolas Papadopoulos,* Luis A. Diaz Jr.

*Corresponding author. E-mail: ldiaz1@jhmi.edu (L.A.D.); npapad1@jhmi.edu (N.P.); vogelbe@jhmi.edu (B.V.); kinzlk@jhmi.edu (K.W.K.); velculescu@jhmi.edu (V.E.V.)

Published 19 February 2014, *Sci. Transl. Med.* **6**, 224ra24 (2014)
DOI: 10.1126/scitranslmed.3007094

This PDF file includes:

Materials and Methods

Fig. S1. Comparison of methods for analysis of point mutations in plasma DNA.

Fig. S2. ctDNA in advanced malignancies, ranking of the fraction of patients with detectable ctDNA.

Fig. S3. Diagram of the assay used to confirm rearrangements in plasma DNA.

References (52–58)

Other Supplementary Material for this manuscript includes the following:

(available at www.sciencetranslationalmedicine.org/cgi/content/full/6/224/224ra24/DC1)

Table S1 (Microsoft Excel format). Summary of tissue mutations, plasma mutations, and clinical information for 410 early- and advanced-stage subjects.

Table S2 (Microsoft Excel format). Comparison of ctDNA levels assessing point mutations and rearrangements in the same specimens.

Table S3 (Microsoft Excel format). Comparison between plasma and tumor tissue *KRAS* status in 206 patients with metastatic CRC.

Table S4 (Microsoft Excel format). Clinical characteristics of patients with discordant tissue and plasma *KRAS* mutation data.

Table S5 (Microsoft Excel format). Clinical characteristics of patients with false negatives in plasma *KRAS* mutation compared with tissue samples.

Table S6 (Microsoft Excel format). Association between clinical characteristics and absolute levels of ctDNA.

Table S7 (Microsoft Excel format). Association between clinical characteristics and ctDNA concentration (log scale) in metastatic CRC patients.

Table S8 (Microsoft Excel format). Patient characteristics and plasma mutations detected after EGFR blockade.

Filename: generate.table.r

Filename: Table_S1_analysis.r

Filename: Table_S3_analysis.r

Materials and Methods

Samples

All samples were collected after Institutional Review Board (IRB) approval at participating institutions, under full compliance with HIPAA guidelines. Tumors and adjacent normal tissues were either frozen at a minimum of -80 C or formalin-fixed and paraffin- embedded (FFPE) according to standard histopathologic procedures. Tumors were macro-dissected under a dissecting microscope to ensure a neoplastic cellularity of >60%. DNA was purified from the macrodissected frozen tumors using AllPrep (Qiagen, cat #80204) and from macrodissected paraffin-embedded tumors with a Qiagen FFPE Kit (Qiagen cat #56494). Translocations, but not point mutations, were previously reported for three of the CRCs (42, 52). Translocation data, but not all clinical correlatives, were previously reported for eight of the nine neuroblastomas recorded in table S1 (53); these cases were included in the current study for comparative purposes only. For white blood cell DNA extraction, cells were pelleted at 1000 g prior to the preparation of plasma. DNA from these cells was purified using AllPrep (Qiagen, cat #80204). Plasma was used for ctDNA measurements in all experiments except in 17 of the 24 cases described in table S6, in which serum was used. DNA from plasma or serum was purified using QIAamp Circulating Nucleic Acid kit (Qiagen cat# 55114). Total plasma DNA concentration was measured using quantitative PCR as described (54). The amounts of plasma available from all patients except those used for the liquid biopsy studies are listed in table S1; for the liquid biopsy study in table S3, 2 ml plasma was available and for the liquid biopsy study in table S6, 1 ml serum or 2 ml plasma was used. One plasma draw was used for each patient except those described in table S6, in which two plasma draws were obtained: one prior to initiating EGFR blockade and one sample when the tumors had recurred after a clinical response.

Tumor Mutational Profiling

A tiered approach was used to identify somatic mutations in tumors. For pancreatic cancers, genomic regions encompassing *KRAS* codons 12,13, 59, 60 and 61 were amplified and the sequence of the PCR products determined via ligation of mutant-specific probes (53) or via SafeSeqS (40), as it is well known that nearly all pancreatic ductal adenocarcinomas harbor mutations in the *KRAS* gene (44), (55). For colorectal cancers, PCR was used to amplify the *KRAS*, *BRAF*, *TP53*, *SMAD4*, *PIK3CA*, and *APC* genes and the sequence of the PCR products was determined, generally using SafeSeqS, as described below. For all other cancers, paired-end libraries were generated and regions encompassing 100 genes commonly mutated in cancers were captured as described previously (53). For tumors that did not contain

detectable mutations of these genes, exomic sequencing was performed after capture of the same libraries via SureSelect (Agilent), as previously described (56, 57). In cases in which rearrangements were analyzed using PARE (personalized analysis of rearranged ends), (42, 52), genomic libraries were constructed for whole genome sequencing with a physical coverage of ~10x. Whenever possible, we selected rearrangements within amplified segments of the genome. Such rearrangements would be represented more often in tumor DNA than in DNA from normal cells, theoretically increasing the sensitivity of detection of the altered fragment in plasma. Once putative rearrangements were identified on the basis of sequencing data, PCR primers were designed to amplify PCR products of 100 bp that spanned the rearrangement. The rearrangements were confirmed to be somatic by demonstrating that PCR products were generated from the DNA of the tumor but not from DNA of non-neoplastic cells of the same patient.

Mutation Detection in ctDNA or CTCs

In the early phases of this study, single base substitutions and small insertions or deletions (indels) were assessed either by BEAMing or (21, 53) or by a PCR/ligation method (21, 53). For the latter method, 25% of the plasma DNA was aliquotted into wells of a 384-well plate so that an average of 1 ng was contained in each well. After PCR and ligation as described (21, 53), all wells were individually evaluated via gel electrophoresis and fluorescence imaging. If all wells contained a mutation, the plasma DNA was re-diluted for more precise quantification. If no wells contained a mutation, then a further 65% of the plasma was aliquotted and the assay repeated; ~10% of the plasma was used to determine DNA concentration and to confirm that the plasma and tumor were derived from the same patient via single nucleotide polymorphism analysis. In the latter stages of this study, mutations were assessed by SafeSeqS, an approach in which template molecules are individually assessed via massively parallel sequencing (40). For SafeSeqS, 25% of the plasma DNA was aliquotted into wells of a 96-well plate so that an average of 3 ng DNA was contained in each well. The DNA from each well was then amplified using well-specific index primers, and the DNA from all wells was pooled and subjected to massively parallel sequencing and analysis as described (40). If no mutations were detected, a further 65% of the plasma was aliquotted and the assay repeated; ~10% of the plasma was used to determine DNA concentration and to confirm that the plasma and tumor were derived from the same patient via single nucleotide polymorphism analysis. In each experiment, equivalent amounts of DNA from non-neoplastic cells were included in adjacent wells performed to ensure that the identified mutations were not the result of errors generated during PCR or other steps of the procedures (40). SafeSeqS was used to assess all of the 206 metastatic colorectal cancer patients assayed for mutations in the liquid biopsy

study (table S3) as well as to assess the 24 patients assayed for resistance mutations after EGFR blockade (table S6).

To assess differences in assay performance among the methods used to assess ctDNA, we quantified the amount of mutations in 20 plasma samples that had been evaluated by all three methods used for detecting point mutations (BEAMing (58), PCR-Ligation (53), or Safe-SeqS (40)). We found that the results were comparable, as evident from the data in fig. S1. All three methods could detect one mutant template in the DNA from 5 ml plasma, as determined by spiking known amounts of mutant KRAS DNA in plasma DNA from normal individuals.

Rearrangements in ctDNA or CTCs were detected and quantified by digital PCR, using PARE (Paired Analysis of Rearranged Ends) as described previously (40, 42) with the following modifications. First, 25% of the plasma DNA was aliquotted into wells of a 384-well plate so that an average of 3 ng (plasma) or 300 ng (WBCs containing CTCs) were contained in each well. After amplification, a portion of each well was evaluated by polyacrylamide gel electrophoresis to determine whether a PCR product of the predicted size was present. If all wells contained a mutation, the DNA was re-diluted for more precise quantification. If no wells contained a mutation, then a further 65% of the plasma was aliquotted and the assay repeated. To further verify that the PCR fragments of the expected size contained the intended rearrangement, ligation reactions were performed on each PCR fragment as described (21, 53). The two oligonucleotides used in the ligation reaction spanned the breakpoint so that ligation only occurred if the PCR products assessed contained the rearrangement (fig. S3). Control experiments with DNA from non-neoplastic cells of the same patients showed that each rearrangement reported in this study was not found in the germ line of that patient.

Statistical Analyses

Proportions of patients with detectable ctDNA, with 95% Wilson confidence intervals, the rank of the proportion and ctDNA concentration by cancer type, with 95% bootstrapped confidence intervals, are listed. Proportions of patients with detectable ctDNA were compared across cancer types with a likelihood ratio chi-square test from a logistic model of detectable ctDNA, across stage of disease with Somers' Dxy rank correlation, and across both cancer type and stage of disease for breast, colon, pancreas, and gastroesophageal cancers using a Cochran-Mantel-Haenszel chi-squared test.

For the liquid biopsy cohort of metastatic colorectal cancer patients, the sensitivity and specificity, along with 95% confidence intervals, for detecting a plasma KRAS mutation compared to the detection of a

tissue KRAS mutation were calculated. We also report the percent concordance and kappa statistic for the agreement between liquid biopsy and tissue samples.

Clinical characteristics of the false negative and true negative groups were compared with Pearson's chi-square test for categorical variables and Wilcoxon Mann Whitney tests for continuous variables. Those variables which had <20% missingness of the dependent variable were included in a multivariable logistic regression model of true negative status using lasso penalties.

In cases with detectable levels of mutant KRAS fragments in the plasma, the association of clinical characteristics with the log ctDNA levels was evaluated using univariable linear regression models. Logarithm transformations were made for the dependent variable and some continuous predictor variables to correct for skewness. Those variables which had <20% missingness of the dependent variable were included in a multivariable linear regression using lasso penalties.

Overall survival was calculated from the time of ctDNA measurement to the date of death or last follow-up. The known prognostic factors (age, ECOG PS and CEA), linearity assumed, were included in a multivariable Cox proportional hazards model with ctDNA concentration level transformed with a natural spline function. The 2-year survival probability estimates were plotted against ctDNA concentration levels, fixing the other covariates at the mean (continuous variables) or mode (categorical variables). The other prognostic factors, MSI and BRAF status, had more than 20% missing values and were not adjusted in the multivariable model. The variables were selected based on their clinical relevance, and none were removed by statistical significance testing.

Statistical analyses were performed using the R statistical package (version 2.15.1). A complete self-contained software script that can be used to reproduce all data manipulation, transformation, and analysis steps reported in this paper is available as part of the supplementary materials.

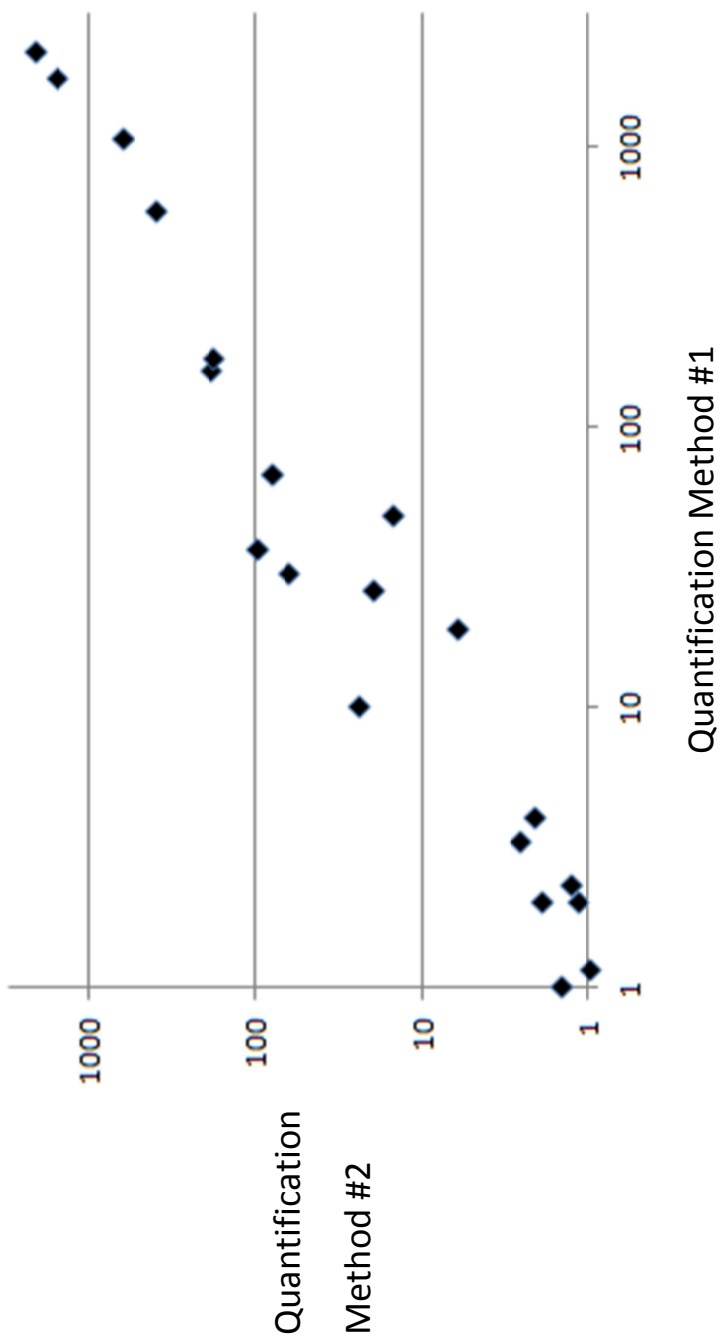


Fig. S1. Comparison of methods for analysis of point mutations in plasma DNA. Duplicate aliquots of plasma or serum from 20 different patients were collected. A mutation present in the corresponding tumors was quantified in the first aliquot by Quantification Method #1 (PCR-ligation) and in the second aliquot by Quantification Method #2 (BEAMing in 11 patients and SafeSeqS in 9 patients). Mutant templates per 5 ml plasma or serum are plotted on both the x- and y-axes. In addition to the 20 samples displayed in this graph, each of which yielded at least one mutation with both quantification methods, we tested 10 other duplicate samples containing low amounts of mutant DNA. In five of these cases, neither of the aliquots tested by the two methods revealed any mutations. In five other cases, one method revealed a single mutant template molecule while the other method revealed zero. These results are consistent with expectations based on a Poisson distribution of mutant templates in the circulation.

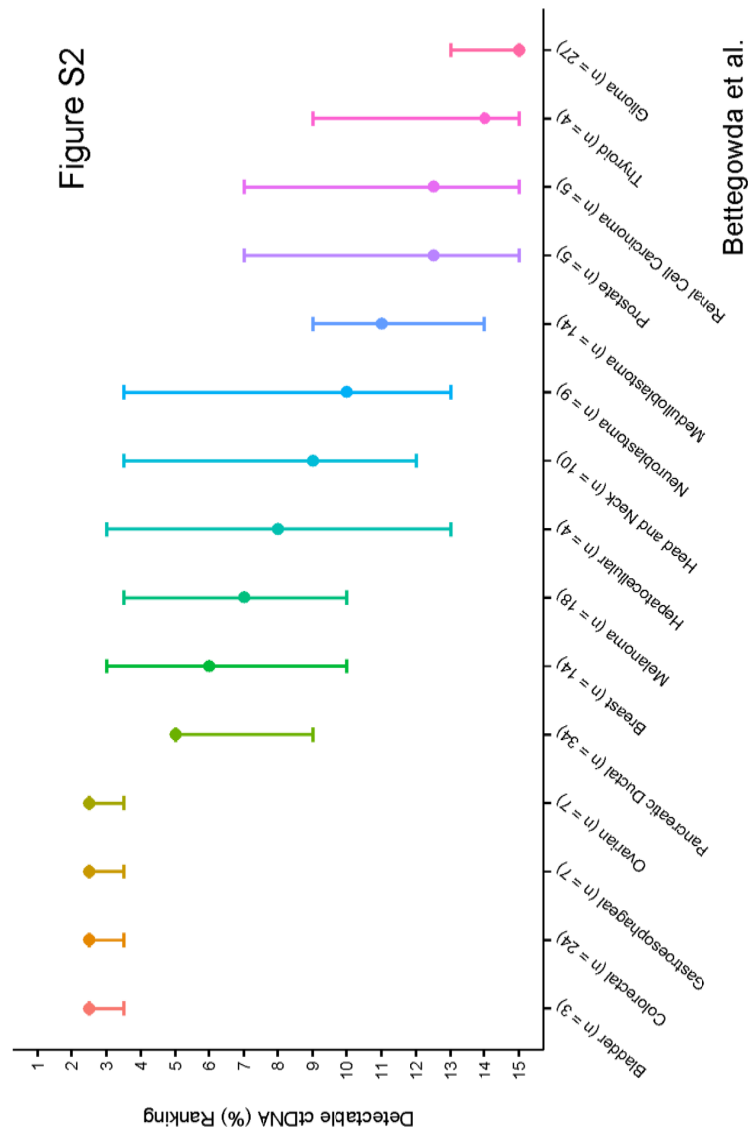


Fig. S2. *ctDNA in advanced malignancies, ranking of the fraction of patients with detectable ctDNA.* For each tumor type, the rank of the proportion of detectable ctDNA is reported, where a rank of 1 means having the highest proportion and a rank of 15 means having the lowest proportion among the 15 tumor types. Ties were handled by taking the average ranking. The error bars represent the 95% bootstrap confidence interval of the rank of each tumor type.

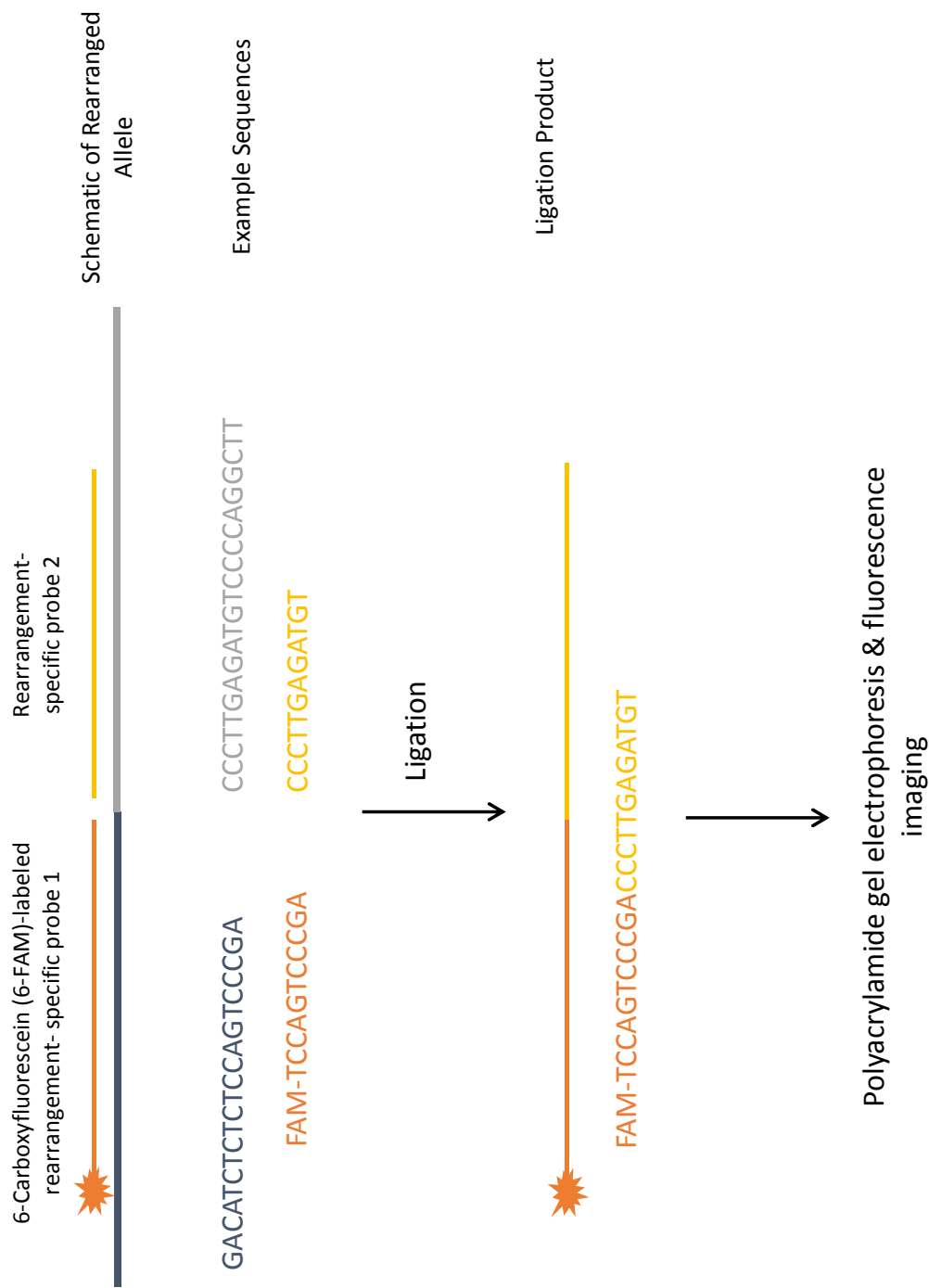


Fig. S3. Diagram of the assay used to confirm rearrangements in plasma DNA.