Science Translational Medicine
AAAS

Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA

Tim Forshew *et al.*

Sci Transl Med **4**, 136ra68 (2012);
DOI: 10.1126/scitranslmed.3003726

Editor's Summary

Deep Sequencing Tumor DNA in Plasma

Five liters of circulating blood contain millions of copies of the genome, broken into short fragments; in cancer patients, a small fraction is circulating tumor DNA (ctDNA). An even smaller number harbor mutations that affect cancer outcome. Looking for diagnostic answers in circulating DNA is a challenge, but Forshew, Murtaza, and colleagues have risen to the occasion by developing a tagged-amplicon deep sequencing (TAm-Seq) method that can amplify and sequence large genomic regions from even single copies of ctDNA. By sequencing such large regions, the authors were able to identify low-level mutations in the plasma of patients with high-grade serous ovarian carcinomas.

Forshew *et al*. designed primers to amplify 5995 bases that covered select regions of cancer-related genes, including *TP53*, *EGFR*, *BRAF*, and *KRAS*. In plasma obtained from 38 patients with high levels of ctDNA, the authors were able to identify mutations in *TP53* at allelic frequencies of 2% to 65%. In plasma samples from one patient, they also identified a de novo mutation in *EGFR* that had not been detected 15 months prior in the tumor mass itself. Finally, the TAm-Seq approach was used to sequence ctDNA in plasma samples collected from two women with ovarian cancer and one woman with breast cancer at different time points, tracking as many as 10 mutations in parallel. Forshew and coauthors showed that levels of mutant alleles reflected the clinical course of the disease and its treatment—for example, stabilized disease was associated with low allelic frequency, whereas patients at relapse exhibited a rise in frequency.

Through several experiments, the authors were able to show that TAm-Seq is a viable method for sequencing large regions of ctDNA. Although this provides a new way to noninvasively identify gene mutations in our blood, TAm-Seq will need to achieve a more sensitive detection limit (<2% allele frequency) to identify mutations in the plasma of patients with less advanced cancers. Nevertheless, once optimized, this "liquid biopsy" approach will be amenable to personalized genomics, where the level and type of mutations in ctDNA would inform clinical decision-making on an individual basis.

Downloaded from stm.sciencemag.org on May 31, 2012

A complete electronic version of this article and other services, including high-resolution figures, can be found at:
http://stm.sciencemag.org/content/4/136/136ra68.full.html

Supplementary Material can be found in the online version of this article at:
http://stm.sciencemag.org/content/suppl/2012/05/25/4.136.136ra68.DC1.html

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:
http://www.sciencemag.org/about/permissions.dtl

Science Translational Medicine (print ISSN 1946-6234; online ISSN 1946-6242) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2012 by the American Association for the Advancement of Science; all rights reserved. The title *Science Translational Medicine* is a registered trademark of AAAS.

Foresight EX1030
Foresight v Personalis
Personalis EX2194

CANCER GENOMICS

Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA

Tim Forshew,[1]* Muhammed Murtaza,[1,2]* Christine Parkinson,[1,2,3]* Davina Gale,[1]* Dana W. Y. Tsui,[1]* Fiona Kaper,[4†] Sarah-Jane Dawson,[1,2,3] Anna M. Piskorz,[1,2] Mercedes Jimenez-Linan,[3,5] David Bentley,[6] James Hadfield,[1] Andrew P. May,[4] Carlos Caldas,[1,2,3,7] James D. Brenton,[1,2,3,7‡] Nitzan Rosenfeld[1,2‡]

Plasma of cancer patients contains cell-free tumor DNA that carries information on tumor mutations and tumor burden. Individual mutations have been probed using allele-specific assays, but sequencing of entire genes to detect cancer mutations in circulating DNA has not been demonstrated. We developed a method for tagged-amplicon deep sequencing (TAm-Seq) and screened 5995 genomic bases for low-frequency mutations. Using this method, we identified cancer mutations present in circulating DNA at allele frequencies as low as 2%, with sensitivity and specificity of >97%. We identified mutations throughout the tumor suppressor gene *TP53* in circulating DNA from 46 plasma samples of advanced ovarian cancer patients. We demonstrated use of TAm-Seq to noninvasively identify the origin of metastatic relapse in a patient with multiple primary tumors. In another case, we identified in plasma an *EGFR* mutation not found in an initial ovarian biopsy. We further used TAm-Seq to monitor tumor dynamics, and tracked 10 concomitant mutations in plasma of a metastatic breast cancer patient over 16 months. This low-cost, high-throughput method could facilitate analysis of circulating DNA as a noninvasive "liquid biopsy" for personalized cancer genomics.

INTRODUCTION

Circulating cell-free DNA extracted from plasma or other body fluids has potentially transformative applications in cancer management (*1*–*7*). Characterization of tumor mutation profiles is required for informed choice of therapy, given that biological agents target specific pathways and effectiveness may be modulated by specific mutations (*8*–*11*). Yet, mutation profiles in different metastatic clones can differ significantly from each other or from the parent primary tumor (*12*, *13*). Evolutionary changes within the cancer can alter the mutational spectrum of the disease and its responsiveness to therapies, which may necessitate repeat biopsies (*14*–*17*). Biopsies are invasive and costly and only provide a snapshot of mutations present at a given time and location. For some applications, mutation detection in plasma DNA as a "liquid biopsy" could potentially replace invasive biopsies as a means to assess tumor genetic characteristics (*2*–*7*). Sensitive methods for detecting cancer mutations in plasma may find use in early detection screening (*1*), prognosis, monitoring tumor dynamics over time, or detection of minimal residual disease (*3*, *18*, *19*). In high-grade serous ovarian carcinomas (HGSOC), mutations in the tumor suppressor gene *TP53* have been observed in 97% of cases (*20*, *21*), but these are located throughout the gene and are difficult to assay. A cost-effective method that could detect and measure allele frequency (AF) of *TP53* mutations in plasma may be highly applicable as a biomarker for HGSOC (*22*).

Circulating DNA is fragmented to an average length of 140 to 170 base pairs (bp) and is present in only a few thousand amplifiable copies per milliliter of blood, of which only a fraction may be diagnostically relevant (*2*, *3*, *23*–*25*). Recent advances in noninvasive prenatal diagnostics highlight the clinical potential of circulating DNA (*25*–*28*), but also the challenges involved in analysis of circulating tumor DNA (ctDNA), where mutated loci and AFs may be more variable. Various methods have been optimized to detect extremely rare alleles (*1*, *2*, *6*, *7*, *29*–*31*), and can assay for predefined or hotspot mutations. These methods, however, interrogate individual or few loci and have limited ability to identify mutations in genes that lack mutation hotspots, such as the *TP53* and *PTEN* tumor suppressor genes (*32*). In patients with more advanced cancers, ctDNA can comprise as much as 1% to 10% or more of circulating DNA (*2*), presenting an opportunity for more extensive genomic analysis. Targeted resequencing has been recently used to identify mutations in selected genes at AFs as low as 5% (*33*–*35*). However, identifying mutations across sizeable genomic regions spanning entire genes at an AF as low as 2%, or in few nanograms of fragmented template from circulating DNA, has been more challenging.

In response, we describe a tool for noninvasive mutation analysis on the basis of tagged-amplicon deep sequencing (TAm-Seq), which allows amplification and deep sequencing of genomic regions spanning thousands of bases from as little as individual copies of fragmented DNA. We applied this technique for detection of both abundant and

[1]Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK. [2]Department of Oncology, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, UK. [3]Addenbrooke's Hospital, Cambridge University Hospital NHS Foundation Trust and National Institute for Health Research Cambridge Biomedical Research Centre, Cambridge CB2 2QQ, UK. [4]Fluidigm Corporation, 7000 Shoreline Court, Suite 100, South San Francisco, CA 94080, USA. [5]Department of Histopathology, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK. [6]Illumina Cambridge, Chesterford Research Park, Little Chesterford, Cambridge CB10 1XL, UK. [7]Cambridge Experimental Cancer Medicine Centre, Cambridge CB2 0RE, UK.
*These authors contributed equally to this work.
†Present address: Illumina, Inc., 5200 Illumina Way, San Diego, CA 92122, USA.
‡To whom correspondence should be addressed. E-mail: nitzan.rosenfeld@cancer.org.uk (N.R.); james.brenton@cancer.org.uk (J.D.B.)

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

rare mutations in circulating DNA from blood plasma of ovarian and breast cancer patients. This sequencing approach allowed us to monitor changes in tumor burden by sampling only patient plasma over time. Combined with faster, more accurate sequencing technologies or rare allele amplification strategies, this approach could potentially be used for personalized medicine at point of care.

RESULTS

Targeted deep sequencing of fragmented DNA by TAm-Seq

To amplify and sequence fragmented DNA, we designed primers to generate amplicons that tile regions of interest in short segments of about 150 to 200 bases (Fig. 1A and table S1), incorporating universal

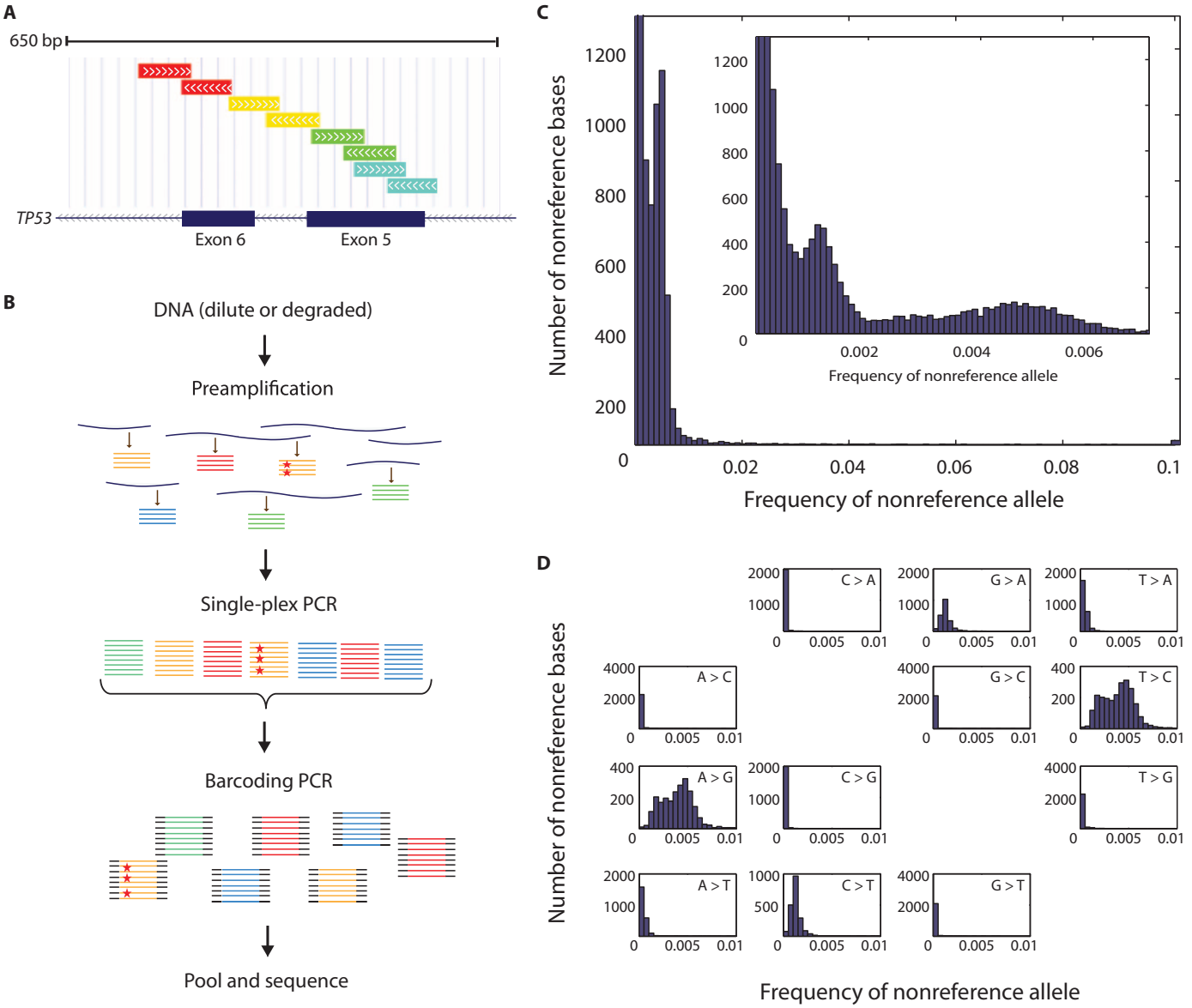

Fig. 1. Overview of tagged amplicon sequencing (TAm-Seq). (**A**) Illustration of amplicon design. Primers were designed to amplify regions of interest in overlapping short amplicons (table S1). Amplicon design is illustrated for a region covering exons 5 to 6 of *TP53*. Colored bars, segmented into forward and reverse reads, show regions covered by different amplicons (excluding primer regions). Sequencing adaptors are attached at either end, such that a single-end read generates separate sets of forward and reverse reads (fig. S1). Because amplicons are mostly shorter than 200 bp, the forward and reverse reads also partially overlap. Figure adapted from University of California, Santa Cruz, Genome Browser (http://genome.ucsc.edu/). (**B**) Workflow overview. Multiple regions were amplified in parallel. An initial preamplification step was performed for 15 cycles using a pool of the target-specific primer pairs to preserve representation of all alleles in the template material. The schematic diagram shows DNA molecules that carry mutations (red stars) being amplified alongside wild-type molecules. Regions of interest in the preamplified material were then selectively amplified in individual (single-plex) PCR, thus excluding nonspecific products. Finally, sequencing adaptors and sample-specific barcodes were attached to the harvested amplicons in a further PCR. (**C**) Distribution of observed nonreference read frequencies, averaged over 47 FFPE samples, across all loci and all nonreference bases. Inset expands the low-frequency range. (**D**) Distribution of the observed background nonreference read frequencies averaged over 47 FFPE samples for the 12 different A/C/G/T base substitutions.

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

adaptors at 5′ ends (fig. S1). Performing single-plex amplification with each of these primer pairs would require dispersing the initial sample into many separate reactions, considerably increasing the probability of sampling errors and allelic loss. Multiplex amplification using a large set of primers could result in nonspecific amplification products and biased coverage. We therefore applied a two-step amplification process: a limited-cycle preamplification step where all primer sets were used together to capture the starting molecules present in the template, followed by individual amplification to purify and select for intended targets (Fig. 1B) (Supplementary Methods). The final concentration of each primer in the preamplification reaction was 50 nM, reducing the potential for interprimer interactions, and 15 cycles of long-extension (4 min) polymerase chain reaction (PCR) were used to remain in the exponential phase of amplification. We used a microfluidic system (Access Array, Fluidigm) to perform parallel single-plex amplification from multiple preamplified samples using multiple primer sets. An additional PCR step attached sequencing adaptors (fig. S1) and tagged each sample by a unique molecular identifier or "barcode" (table S2). Sequencing adaptors were separately attached at either end and the products mixed together, such that single-end sequencing generated separate sets of forward and reverse reads. We performed 100-base single-end sequencing (GAIIx sequencer, Illumina), with an additional 10 cycles using the barcode sequencing primer, generating ~30 million reads per lane. This produced an average read depth of 3250 for each of 96 barcoded samples for 48 amplicons read in two possible orientations.

Validation and sensitivity for mutation identification in ovarian tumor samples

We designed a set of 48 primer pairs to amplify 5995 bases of genomic sequence covering coding regions (exons and exon junctions) of *TP53* and *PTEN*, and selected regions in *EGFR*, *BRAF*, *KRAS*, and *PIK3CA* (table S1) by overlapping short amplicons (Fig. 1A). The sequenced regions cover mutations that account for 38% of all point mutations in the COSMIC database (v55) (*32*). We used TAm-Seq to sequence DNA extracted from 47 formalin-fixed, paraffin-embedded (FFPE) tumor specimens of ovarian cancers (table S3), which were also sequenced for *TP53* by Sanger sequencing (*36*) (Supplementary Methods). DNA extracted from FFPE samples is generally degraded and fragmented as a result of fixation and long-term ambient storage. We amplified DNA from each sample in duplicate, tagging each replicate with a different barcode. Using a single lane of sequencing, we generated 3.5 gigabases of data passing signal purity filters, producing mean read depth of 3200 above Q30 for each of the 9024 expected read groups (48 amplicons × 2 directions × 94 barcoded samples). Background frequencies of nonreference reads were ~0.1% (median, 0.03%; mean, 0.2%; in keeping with Q30 quality threshold applied), yet varied substantially between loci and base substitutions (Fig. 1C) and showed a clear bias toward purine/pyrimidine conservation (Fig. 1D). Sixty-six percent of loci had mean background rate of <0.1%, and 96% of loci had background rate of <0.6%.

The data set interrogated nearly 18,000 possible single-base substitutions for each sample, which introduces a risk of false detection. To control for sporadic PCR errors and reduce false positives, we called point mutations in a sample only if nonreference AFs were above the respective substitution-specific background distribution at a high confidence margin (0.9995 or greater), and ranked high in the list of nonreference AFs, in both replicates (Supplementary Methods). Duplicate sequencing data were obtained for 44 samples, and 43 single-base substitutions were called (table S3). These matched 100% of mutations identified by Sanger sequencing and included three additional mutations at low AFs that were below detection thresholds of Sanger sequencing (fig. S2). The upper bound of AFs that may have been missed was estimated (Supplementary Methods) at <5% for 36 of 44 FFPE samples (82%) and <10% for 42 of 44 samples (95%), with median value of 1.3% and mean value of 2.7%. Mutant AFs were highly reproducible in duplicate samples. For 42 of 43 mutations called, the difference in measured frequency between duplicates was less than 0.08, and the relative difference was 25% or less (Fig. 2A). Mutant AFs correlated significantly with tumor cellularity in the FFPE block (correlation coefficient = 0.422; $P = 0.0049$, t test) (Fig. 2B).

In a separate run, we sequenced libraries prepared from six different diluted mixtures of six FFPE samples, with a different known point mutation in *TP53* in each, to mean read depth of 5600. Of more than 100,000 possible non-SNP (single-nucleotide polymorphism) substitutions, we identified all 33 expected point mutations present at AF >1%, including 6 mutations present at AF <2%, with one false-positive called with AF = 1.9%. Using less stringent parameters (Supplementary Methods), we identified three additional mutations present at AF = 0.6% (Fig. 2C), with no additional false positives. Thus, we obtained 100% sensitivity, identifying mutations at AFs as low as 0.6%. A positive predictive value (PPV) of 100% was calculated for mutations at AF >2%, and a PPV of 90% for mutations identified at AF <2% (Fig. 2D).

Quantitative limitations of mutation detection

When applying TAm-Seq to measure a predefined mutation (as opposed to screening thousands of possible substitutions), the frequency of the mutant allele can be read out directly from the data at the desired locus. False detection is less likely, and criteria for confident mutation detection for a predefined substitution can be less stringent than those described above for de novo mutation identification (Supplementary Methods). The minimal nonreference AFs that could be detected depend on the read depth and background rates of nonreference reads, which vary per locus and substitution type. Minimal detectable frequencies increase when higher confidence margins are used (Supplementary Methods) and had a median value of 0.14% at confidence margin of 0.95 and 0.18% at confidence margin of 0.99 (fig. S3). The minimal detectable frequency would also be limited if a minimal number of reads is applied for confident mutation detection; for example, a minimum of 10 reads implies that sequencing depth of 5000 would be required to detect mutations at AF as low as 0.2%. For alleles present at ~10 or fewer copies in the starting template, reproducibility would also be limited by sampling noise, because these alleles may be over- or underrepresented in any particular reaction.

To characterize the quantitative accuracy of TAm-Seq as applied to circulating DNA, we simulated rare circulating tumor mutations by mixing plasma DNA from two healthy individuals. Using the same set of primers as used for the FFPE experiment, we identified that these two individuals differed at five known SNP loci (table S4). Total amplifiable copies in both plasma DNA samples were determined by digital PCR and mixed to obtain minor AFs ranging from 0.16% to 40% (Supplementary Methods). We sequenced diluted templates containing between 250 and <1 expected copy of the minor allele (table S5). The coefficient of variation (CV) of the observed AFs was equal on average to the inverse square root ($1/\sqrt{n}$) of the expected number of copies of the rare allele (Fig. 3A), which is the theoretical

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

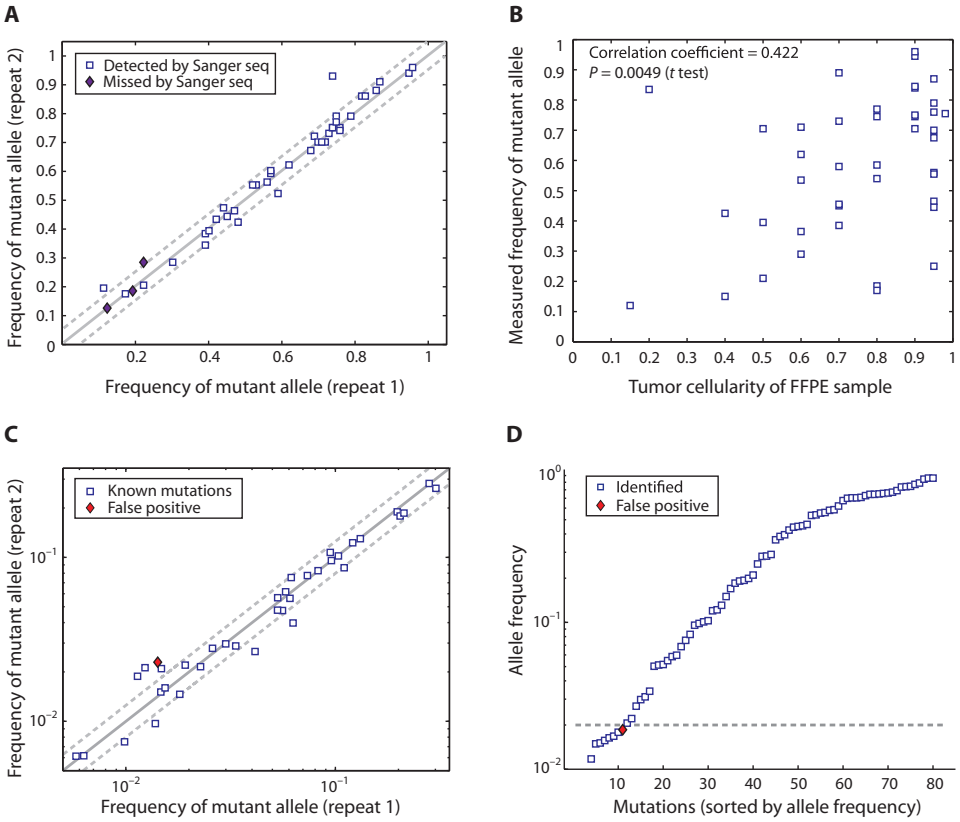

Fig. 2. Identification of mutations in ovarian cancer FFPE samples by TAm-Seq. (**A**) Concordance between duplicate measurements of AFs of mutations identified in fragmented DNA extracted from FFPE samples. The mutation frequency in each library was calculated as the fraction of reads with the mutant (nonreference) base. Solid line indicates equality. Dotted lines indicate a difference in AF of 0.05. (**B**) Correlation of AF with FFPE tumor cellularity. The measured mutant AF (average of both repeats) correlated significantly with the cellularity, estimated from histology (table S3). (**C**) Concordance between duplicate measurements of AFs of mutations identified in a mixture of DNA extracted from different FFPE samples. (**D**) Summary of mutations called in FFPE using TAm-Seq, sorted by increasing AF. Dotted line indicates AF of 2%.

limit of accuracy set by the Poisson distribution for independently segregating molecules. We compared the observed AF to the expected AF for cases where more than six copies of the minor allele were expected. Of 24 such cases, the root mean square (RMS) relative error between the expected and the observed frequency was 14%, with only 2 of 24 cases exhibiting more than 20% discrepancy. For samples with expected minor AF of 0.025, the RMS error was 23% (Fig. 3B).

Noninvasive identification of cancer mutations in plasma circulating DNA

We applied TAm-Seq to directly identify mutations in plasma of cancer patients. We studied a cohort of samples from individuals with HGSOC. These samples were first analyzed for tumor-specific mutations using digital PCR (Supplementary Methods), a method that is highly accurate (*2*, *3*, *7*, *37*) but requires design and validation of a different assay for every mutation screened and relies on previous identification of mutations in tumor samples from the same patients (*2*, *3*). We initially selected for analysis seven cases that had relatively high levels of circulating mutant *TP53* DNA in the plasma (as assessed by digital PCR). Using the equivalent amount of DNA present in 30 to 120 μl of plasma, we performed duplicate preamplification reactions for each sample. For all seven patients, *TP53* tumor mutations were identified in the circulating DNA at frequencies of 4% to 44% (Table 1). In one plasma sample collected from an ovarian cancer patient at relapse, we also identified a de novo mutation in the tyrosine kinase domain of *EGFR* (exon 21), at AF of 6% (patient 27, Table 1). We subsequently validated the presence of this mutation in plasma by performing replicate Sanger sequencing reactions of highly diluted template (Supplementary Methods), and 4 of 91 wells that were successfully Sanger-sequenced contained the *EGFR* mutation (fig. S4). We further validated the presence of this mutation by designing a sequence-specific TaqMan probe targeting this mutation and performing digital PCR (Table 1). The mutation was also identified by TAm-Seq in additional plasma collected from the same individual (sample 16, Table 2). This mutation in *EGFR* was not found in the ovarian mass removed by interval debulking surgery 15 months before the blood sample was collected, although the same sample did contain the concomitant *TP53* mutation found in the same patient's plasma, at AF of 85% (patient 27, table S3). We subsequently used TAm-Seq to sequence seven additional samples collected at the time of initial surgery including deposits in right and left ovaries and omentum. The *EGFR* mutation was detected in the two omental samples above the 0.99 confidence margin (fig. S3) at AF of 0.7%, but was not detected in the six ovarian samples (below the 0.8 confidence margin). Without previous identification in plasma, this mutation would not have been directly identified on screening those samples using high-specificity mutation identification criteria owing to its low AF. In contrast, the *TP53* mutation was identifiable in all biopsy and plasma samples (Fig. 4A). The frequency of mutant alleles in the relapsed tumor could not be directly assessed because a biopsy at relapse was not available.

We validated the TAm-Seq method on a larger panel of plasma samples in which levels of tumor-specific mutations were measured in parallel using patient-specific digital PCR assays. DNA extracted from 62 additional plasma samples collected at different time points from 37 patients with advanced HGSOC was amplified in duplicate (table S6), using DNA present in ~0.15 ml of plasma per reaction (range, 0.06 to 0.2 ml). Amplicon libraries were tagged and pooled together for sequencing with libraries prepared from 24 control samples. This generated an average sequencing depth of 650 for 62 plasma samples, sufficient to detect mutations present at AFs of 1% to 2%. Of >1.5 million possible substitutions, 42 mutations were called using the parameters previously optimized for FFPE analysis (table S6).

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

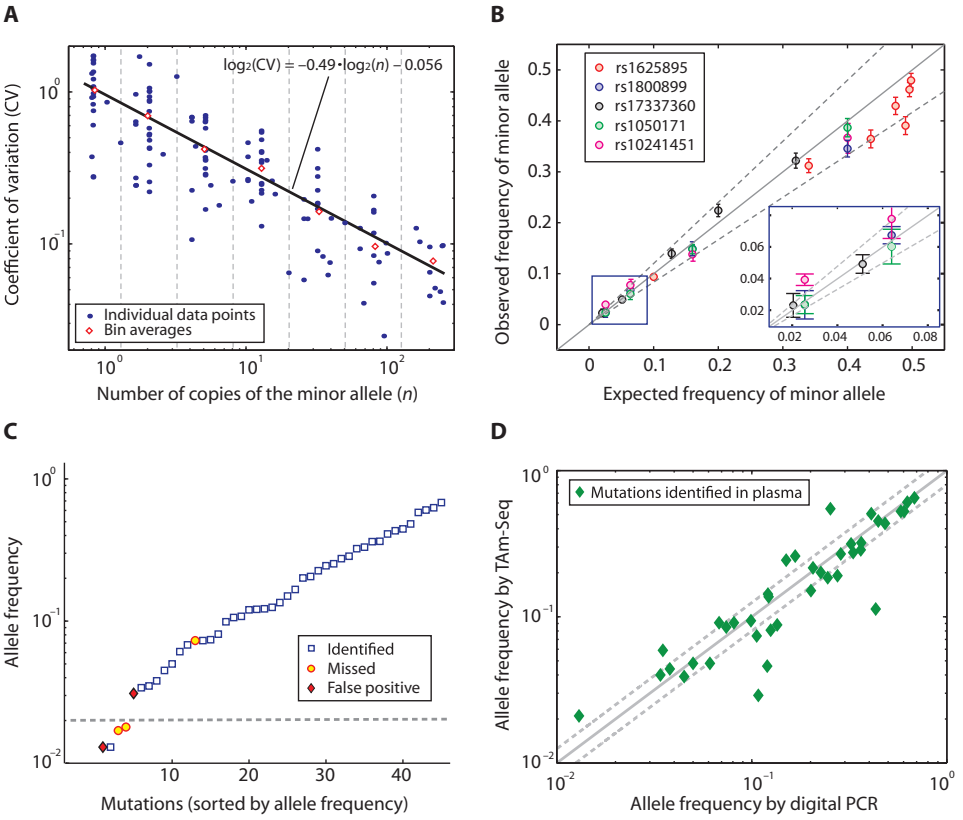

Fig. 3. Noninvasive identification and quantification of cancer mutations in plasma DNA by TAm-Seq. (**A**) Sampling noise in sequencing of sparse DNA using dilutions of plasma DNA from healthy individuals. CV of triplicate AF readings was calculated for each of the five SNPs in each of the mixes, which had varying numbers of copies of the minor allele (n) (blue dots). Bin averages (red diamonds) are the mean CVs calculated for each bin (bin edges denoted by the dotted vertical lines). A linear fit to the $\log_2$ of the mean CV as a function of the $\log_2$ expected copy number was calculated (black line). Two data points, with (n = 100, CV = 0.0064) and (n = 32, CV = 0.0185), were omitted from the figure for enhanced scaling. Three data points with minor allele copies of <0.8 were omitted from the analysis (n = 0.51, CV = 0.62; n = 0.41, CV = 0.86; n = 0.20, CV = 0.99). (**B**) Expected versus observed frequency of rare alleles in a dilution series of circulating DNA. Mean observed frequency was calculated for each of five SNPs for samples, where expected initial number of minor allele copies was greater than 6. Expected frequencies were calculated on the basis of quantification by digital PCR. Dotted lines represent 20% deviation from the expected frequencies. Inset highlights cases with expected minor AF <0.025. (**C**) Mutations identified in 62 plasma samples from patients with advanced HGSOC using TAm-Seq. AFs are based on digital PCR measurement for confirmed mutations (identified or missed by TAm-Seq), and on TAm-Seq for the false positives called using parameters optimized for analysis of FFPE samples. The dashed horizontal line indicates AF of 2%. Mutations detected by digital PCR at AF <1% are not shown. (**D**) AFs measured by TAm-Seq versus digital PCR for mutations identified in plasma DNA.

Table 1. Mutations identified by TAm-Seq in plasma samples from seven ovarian cancer patients. TAm-Seq was used to sequence DNA extracted from plasma of subjects with HGSOC (stage III/IV at diagnosis). Plasma was collected when patients presented with relapse disease, before initiation of chemotherapy. For patient 46, DNA from a formalin-fixed, paraffin-embedded (FFPE) sample was not included in the TAm-Seq set and the mutation was validated in FFPE by Sanger sequencing. CA125 was measured at time of plasma collection. Mean depth of coverage at the mutation locus in the TAm-Seq data was averaged over the repeats (RMS deviation = 850). AF, allele frequency; N, no; Y, yes.

Patient ID	Age at diagnosis	Time elapsed since surgery (months); number of previous lines of chemotherapy	CA125 (U/ml)	Plasma per amplification reaction (μl)	Gene	Mutation and base change (genome build hg19)		Protein change	Detected in FFPE	Mean depth (sequencing reads)	Mean AF using TAm-Seq	Mean AF using digital PCR
8	60	13; 1	2122	50	*TP53*	17:7577120	C>T	p.R273H	Y	5000	0.09	0.10
12	62	27; 3	365	50	*TP53*	17:7577579	G>T	p.Y234*	Y	5000	0.10	0.08
14	58	50; 3	260	120	*TP53*	17:7578212	G>A	p.R213*	Y	5800	0.15	0.12
25	61	9; 1	944	110	*TP53*	17:7578404	A>T	p.C176S	Y	4800	0.04	0.08
27†	68	15; 1	1051	90	*TP53*	17:7578262	C>G	p.R196P	Y	7700	0.06	0.14
					EGFR	7:55259437	G>A	p.R832H	N	5700	0.06	0.05
31	64	12; 1	313	30	*TP53*	17:7578406	C>T	p.R175H	Y	4500	0.44	0.56
46	56	30; 2	1509	30	*TP53*	17:7578406	C>T	p.R175H	Y	4200	0.23	0.30

*Indicates stop codon. †Both a *TP53* and an *EGFR* mutation were identified in this sample (Fig. 4A).

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

Table 2. Mutations identified by TAm-Seq in a set of 62 plasma samples from ovarian cancer patients. Forty mutations were identified by TAm-Seq using stringent parameters for mutation calling. Plasma samples described in this table are distinct from those in Table 1, but patients included overlap. Additional data on patients and mutations are provided in table S6.

Sample number	Plasma volume per amplification reaction (μl)	DNA amount per amplification reaction (ng)	Gene	Protein change	Mean depth (sequencing reads)	Mean AF using TAm-Seq	Mean AF using digital PCR
1	70	0.9	*TP53*	p.R273C	640	0.260	0.167
2	160	4.2	*TP53*	p.R248Q	340	0.244	0.150
3	160	5.7	*TP53*	p.R248Q	640	0.507	0.410
4	120	9.9	*TP53*	p.R213X	810	0.059	0.035
5	120	1.4	*TP53*	p.C141Y	680	0.021	0.013
6	120	2.1	*TP53*	p.C141Y	720	0.044	0.038
7	190	17.9	*TP53*	p.I195N	800	0.091	0.081
8	160	14.8	*TP53*	p.R175H	510	0.608	0.627
9	160	10.7	*TP53*	p.R175H	550	0.526	0.604
10	160	6.1	*TP53*	p.R175H	530	0.651	0.682
11	160	4.9	*TP53*	p.R175H	490	0.526	0.581
13	160	2.8	*TP53*	p.C135R	480	0.039	0.045
14	160	2.5	*TP53*	p.C135R	610	0.046	0.120
15	160	3.0	*TP53*	p.C135R	470	0.091	0.068
16†	130	3.7	*TP53*	p.R196P	1070	0.088	0.135
			EGFR	p.R832H	614	0.048	0.050
17	160	4.2	*TP53*	p.C176S	580	0.113	0.432
18	160	4.4	*TP53*	p.C176S	620	0.029	0.108
20	140	5.2	*TP53*	p.R175H	650	0.201	0.226
21	140	3.6	*TP53*	p.R175H	650	0.085	0.074
22	140	4.1	*TP53*	p.R175H	630	0.081	0.125
23	140	3.7	*TP53*	p.R175H	710	0.074	0.106
24	140	7.1	*TP53*	p.R175H	760	0.269	0.286
25	130	3.9	*TP53*	p.R273H	750	0.094	0.099
26	160	5.7	*TP53*	p.R282W	640	0.048	0.061
27	150	3.6	*TP53*	p.C141Y	480	0.321	0.364
29	150	9.5	*TP53*	p.E258K	190	0.548	0.253
31	160	3.6	*TP53*	p.C135Y	620	0.040	0.034
32	140	2.4	*TP53*	p.E56X	1480	0.137	0.122
33	160	13.2	*TP53*	p.K132N	740	0.216	0.206
34	60	5.3	*TP53*	p.K132N	570	0.151	0.201
36	160	5.8	*TP53*	p.K132N	620	0.191	0.275
37	160	9.4	*TP53*	p.K132N	530	0.287	0.362
38	160	10.1	*TP53*	p.K132N	590	0.275	0.331
39	160	16.4	*TP53*	p.K132N	700	0.315	0.323
40	160	19.7	*TP53*	p.K132N	830	0.435	0.482
41	160	15.0	*TP53*	p.K132N	730	0.452	0.445
42	160	8.5	*TP53*	p.K132N	560	0.185	0.245
43	150	3.6	*TP53*	Splicing	680	0.143	0.121
44‡	170	5.2	*TP53*	p.C238R	1543	0.071	0.073

†Both a *TP53* and an *EGFR* mutation were identified in this sample, collected from patient 27 (Table 1), 25 months after initial surgery (Fig. 4A). ‡The amplicon containing the mutation failed amplification in this sample in the initial experiment and was identified successfully in repeat analysis.

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

Thirty-nine of these matched mutations detected by digital PCR in those samples (Fig. 3C). Three potential false positives were called, at AF of 3.1%, 1.3%, and 0.7% (the latter in a control sample). Using higher-stringency parameters for mutation identification (Supplementary Methods), we retained only the 39 validated mutations called, with no false positives (Table 2).

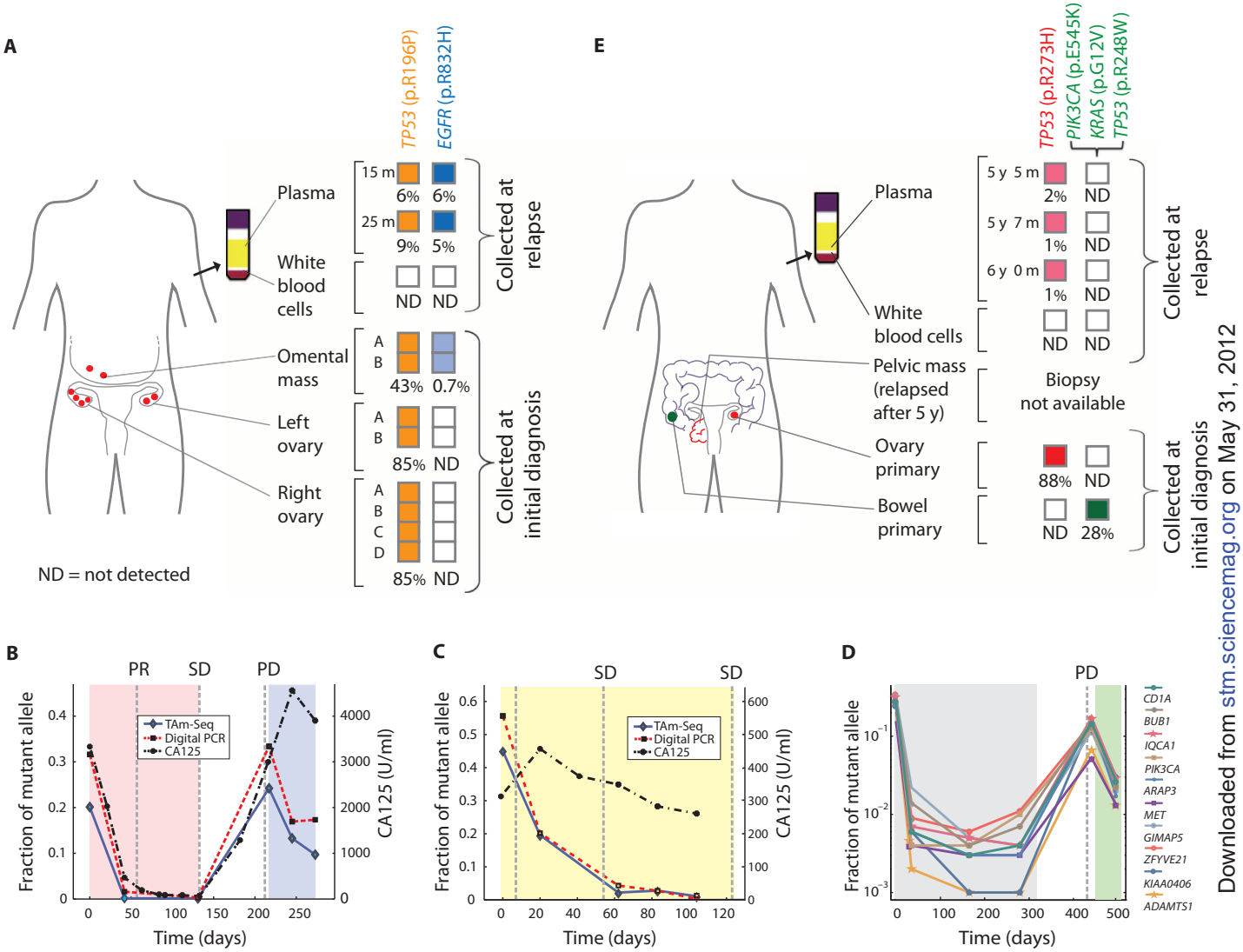

Fig. 4. Clinically relevant applications of plasma DNA sequencing using TAm-Seq. (**A**) Retrospective analysis by TAm-Seq of plasma samples collected during patient follow-up and biopsy specimens collected at initial surgery. We identified a mutation in exon 21 of *EGFR* (dark blue boxes) in two separate plasma samples, collected 15 and 25 months after initial surgery from patient 27 (Tables 1 and 2). This mutation was not directly identified in eight tumor biopsy specimens collected at the time of initial surgery (two from omental mass, two from left ovary, and four from right ovary). Having identified the mutation in the plasma samples, we examined this mutation using the lower-specificity criteria defined for mutation detection (Supplementary Methods) and detected the mutation in the two specimens that had been collected from the omentum at the time of surgery (light blue boxes) but not in the six ovarian specimens. A mutation in *TP53* was identified in all tumor and plasma samples collected from this patient (Tables 1 and 2 and table S3), but not in white blood cells (buffy coat). Percentages indicate mutant AFs. Empty boxes and "ND" indicate samples where a mutation was not identified or detected (below 0.8 confidence margin). (**B**) Monitoring frequency of mutant DNA in plasma of an ovarian cancer patient (patient 46) over time using TAm-Seq and digital PCR. TAm-Seq results are reported as the mean frequency of duplicate analyses. Parallel data are shown for digital PCR and serum CA125. Shaded regions indicate periods of chemotherapy, and vertical dashed lines indicate radiological assessment of patient responses: PR, partial response; SD, stable disease; PD, progressive disease. (**C**) Monitoring frequency of mutant DNA in plasma of an ovarian cancer patient (patient 31) over time. (**D**) Dynamics of 10 tumor-specific mutations in plasma of a breast cancer patient (not included in the other sets of samples analyzed). (**E**) Retrospective analysis of samples from synchronous primary tumors (bowel and ovarian) collected at the time of initial surgery and three plasma samples collected at relapse. In primary tumors from this patient (not included in the other sets of samples analyzed), a *TP53* mutation was identified in the ovarian cancer (red box), and mutations in *PIK3CA*, *KRAS*, and *TP53* were identified in the bowel cancer (green box). At relapse, a biopsy was not performed on the pelvic mass. The *TP53* mutation that was identified in the ovarian primary tumor (p.R273H) was detected in plasma, whereas the bowel-associated mutations were not detected.

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

Of 40 point mutations detected at AF >2% by digital PCR, 38 (95%) were identified by TAm-Seq in a single experiment (Fig. 3C). One additional mutation was located in an amplicon that failed in that sample and was identified in repeated analysis; the other was likely missed by TAm-Seq owing to sampling noise, because it was found in one of the duplicate preamplified libraries but not the other (table S6). One of three mutations detected by digital PCR at 1% < AF < 2% was identified by TAm-Seq (Fig. 3C). Eleven additional point mutations detected by digital PCR at AF <1% were not detected by TAm-Seq at these settings. TAm-Seq and digital PCR measurements of AF had excellent agreement, with correlation coefficient of 0.90, increasing to 0.97 when discarding the two strongest outliers (Fig. 3D). Thus, we screened 62 samples across sizeable genomic stretches, using minute amounts of plasma DNA (median, 4 ng), and obtained 97.5% sensitivity with PPV of 100% for identifying mutations at AF >2% in plasma by TAm-Seq. Using parameters optimized for FFPE samples, one potential false positive was called at AF >2%, reducing the PPV to 97.5% (Table 3).

Monitoring levels of ctDNA

Various methods have been suggested to monitor changes in mutation load in plasma. These can have enhanced sensitivity compared to TAm-Seq for tracking individual mutations, but require design of personalized assays (*3*, *18*, *19*). None of these methods have been widely adopted. We therefore applied TAm-Seq as a generic tool to measure changes in the frequency of ctDNA over time. We studied serial plasma samples collected during follow-up and treatment of two patients with relapsed HGSOC, collected during 104 and 273 days of follow-up and treatment, respectively. Frequencies of mutant *TP53* alleles were measured by TAm-Seq and in parallel by digital PCR using a mutation-specific probe. The two methods of quantification had excellent agreement. Mutant AFs in plasma of ovarian cancer patients reflected well the clinical course of the disease compared to the serum marker CA125, showed marked decrease when systemic treatment was initiated, and increased in parallel to disease progression. In the first case (Fig. 4B), a 56-year-old woman with relapsed ovarian cancer (patient 46) was treated with fourth-line carboplatin + paclitaxel chemotherapy for six cycles (pink-shaded region). Radiology showed partial response on mid-treatment computed tomography (CT) scan. End-of-treatment CT showed stable disease. Twelve weeks from the end of her fourth-line treatment, the patient developed progressive disease. The patient then initiated fifth-line chemotherapy with liposomal doxorubicin (purple-shaded region). In the second case (Fig. 4C), a 64-year-old woman with relapsed ovarian cancer (patient 31) was treated with second-line ECX (epirubicin, cisplatin, and capecitabine) chemotherapy for six cycles. Radiology showed stable disease on mid- and end-of-treatment CT scans. The patient then remained off treatment, until she progressed 3 months later.

TAm-Seq can be flexibly adapted to sequence different genomic regions by designing primers to amplify regions of interest. We used this capability to study dynamics of multiple mutations in parallel.

Downloaded from stm.sciencemag.org on May 31, 2012

Table 3. Summary of mutations identified in 69 plasma samples of ovarian cancer patients. Samples were analyzed by TAm-Seq and in parallel by digital PCR. Using parameters optimized for plasma DNA, false-positive calls were lost, whereas all confirmed calls were retained, resulting in specificity and PPV of 100%.

First set of plasma samples	
Plasma samples analyzed	7
Point mutations originally detected by digital PCR, using patient-specific assays targeting mutations identified in tumor samples	7
Point mutations identified directly in plasma by TAm-Seq	8
De novo mutations identified by TAm-Seq only, subsequently confirmed by digital PCR	1
Second set of plasma samples	
Plasma samples analyzed	62
Point mutations detected by digital PCR at AF >2%	40
Point mutations with AF >2% (by digital PCR) identified by TAm-Seq	39
Point mutations missed by TAm-Seq due to sampling error	1
Sensitivity of TAm-Seq for identifying mutations at AF >2%	97.5%
PPV of mutations called by TAm-Seq with AF >2%	97.5%*
ctDNA in ovarian cancer	
Advanced ovarian cancer patients in both sets†	38
Patients where TAm-Seq identified cancer mutations	20

*One unconfirmed substitution was called at AF >2% using parameters optimized for FFPE material.

†The first set included 7 patients (Table 1), and the second set included 37 patients (table S6), 6 of whom overlap.

Foresight EX1030
Foresight v Personalis
Personalis EX2194

Whole-genome sequencing of tumor material was used to identify tumor mutations in a patient with metastatic breast cancer undergoing two phases of chemotherapy. Ten mutations were selected, and short amplicons (<120 bp) were designed to cover the mutation loci (table S7). Serial plasma samples were collected over the course of 497 days, both before and after treatment. We performed TAm-Seq in duplicate, using DNA from 0.08 ml of plasma per amplification, and tracked dynamics of all mutations in parallel (Fig. 4D). The patient was treated with single-agent epirubicin (gray-shaded region). After 4 months off treatment, a CT scan showed progressive disease and the patient commenced further treatment with paclitaxel chemotherapy. The 10 mutations followed a common pattern of sharp decline in AF upon onset of therapy and an increase in AF upon disease progression after termination of therapy (Fig. 4D).

Finally, we used TAm-Seq to study plasma from a patient who had a history of two synchronous primary cancers, bowel and ovarian, which were resected simultaneously. After a 5-year remission, a pelvic mass of uncertain origin was detected. A biopsy was considered to guide selection of therapy but was not performed owing to risk of complications and comorbidities. The patient commenced empirically on an ovarian cancer chemotherapy regimen, to which she responded. Retrospective analysis by TAm-Seq of FFPE from the primary tumors collected at initial surgery, and three plasma samples collected serially at the time of relapse (5 years and 5 months, 5 years and 7 months, and 6 years after initial surgery), showed that the patient's plasma at relapse contained the *TP53* (p.R273H) mutation identified in the ovarian primary tumor (exceeding the 0.98, 0.93, and 0.97 confidence margins, respectively), but not the *PIK3CA* (p.E545K), *KRAS* (p.G12V), or *TP53* (p.R248W) mutations identified in the primary bowel cancer (below the 0.8 confidence margin) (Fig. 4E). Had these results been available, uncertainty and treatment delays may have been avoided, as well as the risk of prescribing chemotherapy for an inappropriate tumor site. An alternative possible outcome may have involved a finding of the *PIK3CA* or *KRAS* mutations (present in the primary bowel cancer) in the patient's plasma at the time of relapse. Such a finding, if available to clinicians at the time, may not only have led to alternate chemotherapy being offered but may have also opened the possibility of enrolment into a trial for targeted therapy with mammalian target of rapamycin (mTOR), phosphatidylinositol 3-kinase (PI3K), or mitogen-activated protein kinase kinase (MEK) inhibitors (*11*).

DISCUSSION

Detection of rare mutations in circulating DNA has long been pursued owing to its potentially transformative impact on cancer diagnosis and management. Important progress has been made using sequence-specific assays that target predefined mutations and that detect extremely rare alleles. Assays such as PCR (*6*, *7*), ligation (*5*), and primer extension/mass spectrometry (*27*) can identify specific, predefined mutations in plasma samples. Enhanced detection down to 1 mutant allele in 10,000 or more wild-type alleles can be obtained using a variety of methods, such as peptide nucleic acid and primer extension ("PPEM") (*38*), ligation followed by quantitative PCR ("LigAmp") (*39*), bead-based digital PCR in emulsions ("BEAMing") (*2*, *3*), microfluidic-based (*7*) or droplet-based digital PCR (*40*), or microinsertion/deletion/indel-activated pyrophosphorolysis ("MAP") (*29*). Nonetheless, identification of rare mutations in tumor suppressor genes such as *TP53*, which are widely mutated in cancers but lack a well-defined hotspot region, remains an elusive goal.

In patients with advanced cancers, mutant alleles can reach a sizeable fraction of DNA. For example, Dukes' D colorectal cancers have median 8% mutant AF (*2*). Screening of entire genes for mutations would therefore be useful for some applications, even if analytical selectivity is limited to a few percent. Advances in massively parallel sequencing make new approaches possible. These have largely focused on large-scale analyses, including whole-genome or whole-exome sequencing (*41*). This generates a large amount of data on genomic regions that do not, at present, inform clinical decisions. Moreover, the depth of coverage for clinically significant loci is not sufficient to detect changes that occur at low frequency (<5%). Such approaches have recently been complemented by methods for examination of individual amplicons at great depth (*30*).

The intermediate scale of sequencing is most likely to have immediate impact on clinical genomics. Targeted sequencing has been applied for tumor DNA (*34*, *35*) and cyst fluid (*33*) to detect mutations down to 5% AF, but has not been applied for analysis of circulating tumor nucleic acids. Here, we demonstrate noninvasive identification of mutant alleles in plasma, at AFs as low as 2%, by targeted deep sequencing of circulating DNA. Our TAm-Seq method uses a combination of short amplicons, two-step amplification, sample barcodes, and high-throughput PCR. Because the amplicons are short, this method effectively amplifies even small amounts of fragmented DNA such as are present in circulating DNA. The two-step amplification permits extensive primer multiplexing that enables the amplification and sequencing of sizeable genomic regions by tiling short amplicons without loss of fidelity or efficiency. Duplicate sequencing of each sample is used to avoid false positives stemming from PCR errors. Sample barcodes and high-throughput PCR reduce the per-sample costs to a range where this may be widely applicable. Preparing TAm-Seq libraries for sequencing from 48 samples takes less than 24 hours and involves only few hours of hands-on time. New platforms for massively parallel sequencing allow for fast turnaround times, which make this approach practical in a clinical setting.

The sensitivity presently achieved can provide useful diagnostic information in certain advanced cancers. We studied a cohort of subjects with advanced HGSOC in which the tumor suppressor gene *TP53* is a driver mutation (*20*). Of the 69 plasma samples collected from 38 different individuals with advanced HGSOC, we identified mutations in *TP53* in 46 samples (67%) from 20 of the cases (53%). In contrast, a previous study using a ligase detection reaction with bespoke primers found mutated *TP53* sequences in plasma for only 30% of advanced ovarian cancer patients (*5*), and a study using single-strand conformation polymorphism found no ctDNA in preoperative plasma samples from high-grade serous cancer patients (*42*).

Targeted agents, such as inhibitors of poly(adenosine diphosphate–ribose) polymerase (PARP), or tyrosine kinase inhibitors targeting epidermal growth factor receptor (EGFR), may be applicable for systemic treatment of advanced HGSOC (*8*, *10*, *22*). In a recent study of 203 HGSOC tumors, *EGFR* was found to be the most frequently mutated oncogene and was mutated in nearly 10% of cases (*10*). In one case, we identified in plasma a de novo mutation in the tyrosine kinase domain (exon 21) of *EGFR*, located 26 amino acids upstream of the L858R activating mutation widely documented for lung cancer. In a subset of tumor samples collected from the same patient 15 months earlier, this mutation was detected at AF of 0.7%, but could not have

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

been identified by analysis of those samples alone without previous knowledge of the mutation identified in plasma (Fig. 4A). In a clinical setting, identification of such a mutation could potentially guide treatment with alternative molecularly targeted therapy (*10*). Current clinical recommendations in lung adenocarcinoma suggest mutation assessment in exons 18 to 21 of *EGFR* (a region of ~560 bp) in the tumor tissue to identify patients eligible for treatment with gefitinib or erlotinib (*9*). Using a commercial PCR-based in vitro diagnostic kit (Qiagen), 28 different *EGFR* variants can be assayed (not including the mutation we identified), but the sample needs to be subdivided into seven different reactions. When sample is limited or mutant alleles are rare, this could introduce sampling errors.

Using standard amplification primers tailored to the mutation loci, we also used TAm-Seq to monitor the dynamics of 10 mutations in plasma DNA of a single patient with metastatic breast cancer, using minute amounts of input DNA. Previous studies have followed up to two mutations in any individual patient (*3*, *19*). Tracking multiple mutations can provide insight into clonal evolution and, at the same time, increases the robustness for tumor monitoring by compensating for effects of sampling noise or mutational drift. For example, if a patient has only five copies of a mutant allele per milliliter of plasma (on average), there is a 37% probability that this mutation will not be present in a 0.2-ml sample, and even a perfect assay will fail to detect residual tumor, whereas a method that measures multiple mutations in parallel can have a low likelihood of a false-negative result even if the detection rate for each mutation is less than 50%.

A current limitation of TAm-Seq is the detection limit compared to assays that target individual loci (*2*, *3*, *7*, *40*), which have been shown to detect two to three orders of magnitude lower frequencies. Our approach may be sufficient for analyzing plasma from patients with certain advanced cancers, but further improvement may be necessary before this method can be more widely used in the clinic. Higher read depth or fidelity, additional replicates, or improved algorithms could allow for enhanced mutation detection without change to protocols. An alternative strategy is through rare allele enrichment, for example, by combining TAm-Seq with protocols such as COLD-PCR (co-amplification at lower denaturation temperature PCR) (*31*).

Previously proposed methods for personalized monitoring of tumor dynamics relied on expensive custom-designed probes (*3*) or identification of rearrangements using whole-genome sequencing (*18*, *19*). These have better analytical sensitivity than currently achieved by TAm-Seq, but are difficult to implement on a routine basis. TAm-Seq strikes a balance between sensitivity and ease of use and could facilitate study and application of circulating DNA. Using TAm-Seq, we identified cancer mutations in the plasma of most advanced ovarian cancer patients and tracked dynamics of *TP53* mutations without requiring any specially designed probes. In summary, TAm-Seq is a flexible and cost-effective platform for applications in noninvasive cancer genomics and diagnostics. We have shown that this method can be used for high-throughput sequencing of plasma samples to identify and monitor levels of multiple cancer mutations in circulating DNA. This could also be applied to screen for rare mutations in a variety of heterogeneous sample types such as low-cellularity tumor specimens, cytological samples, or circulating tumor cells (*16*). With further developments, this and derivative methods may be applied in molecular screening for earlier detection or for differential diagnosis of cancer from benign masses. For genetic analysis of FFPE or small biopsy samples, TAm-Seq can be applied as is, as a cost-effective clinical aid.

MATERIALS AND METHODS

Sample collection

FFPE blocks were obtained from the pathology archives at Addenbrooke's Hospital (Cambridge, UK). Plasma samples were collected upon disease relapse, before and during chemotherapy treatment. Sample collection for this study was approved by Cambridgeshire Research Ethics Committee (REC 08/H0306/61 and 07/Q0106/63). Peripheral blood samples were collected into EDTA tubes and centrifuged at 820*g* for 10 min within 1 hour of collection to limit degradation of cell-free DNA and leukocyte lysis. Aliquots (1 ml) of plasma were centrifuged in a benchtop microfuge at 14,000 rpm for 10 min. The supernatant was transferred to sterile 1.5-ml tubes and stored at −80°C before extraction.

Extraction of DNA from FFPE and blood plasma

Paraffin blocks were cut as 8-μm sections on plain glass slides. Targeted regions for sampling were marked on adjacent hematoxylin and eosin sections by the study pathologist and recovered by scrape macrodissection. Between 3 and 20 sections were macrodissected depending on the tissue sample's size. DNA from FFPE sections was extracted with QIAamp DNA FFPE Tissue Kit (Qiagen) according to the manufacturer's instructions.

Circulating DNA was extracted from between 0.85 and 2.2 ml of plasma with the QIAamp Circulating Nucleic Acid kit (Qiagen), following the manufacturer's instructions, and with the QIAvac 24 Plus vacuum manifold. Carrier RNA was added to ACL lysis buffer to enhance binding of nucleic acids to the QIAamp membrane with the aim to enhance yields.

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/4/136/136ra68/DC1
Methods
Fig. S1. PCR strategy and primer design.
Fig. S2. Sanger traces for mutations identified by tagged-amplicon sequencing.
Fig. S3. Background frequencies and detection limits for base substitutions.
Fig. S4. Replicate dilute Sanger sequencing of a mutation identified in plasma.
Table S1. Target-specific primers.
Table S2. Unique sequencing barcodes.
Table S3. Mutations identified in FFPE samples.
Table S4. SNPs identified in circulating DNA from two plasma control samples.
Table S5. Frequency of SNP alleles in dilution series of DNA from control plasma.
Table S6. Additional data for Table 2 for mutations identified in plasma samples.
Table S7. Mutations and amplicons studied in one breast cancer patient.

REFERENCES AND NOTES

1. E. Gormally, E. Caboux, P. Vineis, P. Hainaut, Circulating free DNA in plasma or serum as biomarker of carcinogenesis: Practical aspects and biological significance. *Mutat. Res.* **635**, 105–117 (2007).
2. F. Diehl, M. Li, D. Dressman, Y. He, D. Shen, S. Szabo, L. A. Diaz Jr., S. N. Goodman, K. A. David, H. Juhl, K. W. Kinzler, B. Vogelstein, Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16368–16373 (2005).
3. F. Diehl, K. Schmidt, M. A. Choti, K. Romans, S. Goodman, M. Li, K. Thornton, N. Agrawal, L. Sokoll, S. A. Szabo, K. W. Kinzler, B. Vogelstein, L. A. Diaz Jr., Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* **14**, 985–990 (2008).
4. H. Schwarzenbach, D. S. Hoon, K. Pantel, Cell-free nucleic acids as biomarkers in cancer patients. *Nat. Rev. Cancer* **11**, 426–437 (2011).
5. E. M. Swisher, M. Wollan, S. M. Mahtani, J. B. Willner, R. Garcia, B. A. Goff, M. C. King, Tumor-specific p53 sequences in blood and peritoneal fluid of women with epithelial ovarian cancer. *Am. J. Obstet. Gynecol.* **193**, 662–667 (2005).

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

6. R. E. Board, A. M. Wardley, J. M. Dixon, A. C. Armstrong, S. Howell, L. Renshaw, E. Donald, A. Greystoke, M. Ranson, A. Hughes, C. Dive, Detection of *PIK3CA* mutations in circulating free DNA in patients with breast cancer. *Breast Cancer Res. Treat.* **120**, 461–467 (2010).
7. T. K. Yung, K. C. Chan, T. S. Mok, J. Tong, K. F. To, Y. M. Lo, Single-molecule detection of epidermal growth factor receptor mutations in plasma by microfluidics digital PCR in non–small cell lung cancer patients. *Clin. Cancer Res.* **15**, 2076–2084 (2009).
8. S. Banerjee, S. Kaye, The role of targeted therapy in ovarian cancer. *Eur. J. Cancer* **47** (Suppl. 3), S116–S130 (2011).
9. V. L. Keedy, S. Temin, M. R. Somerfield, M. B. Beasley, D. H. Johnson, L. M. McShane, D. T. Milton, J. R. Strawn, H. A. Wakelee, G. Giaccone, American Society of Clinical Oncology provisional clinical opinion: Epidermal growth factor receptor (*EGFR*) mutation testing for patients with advanced non–small-cell lung cancer considering first-line EGFR tyrosine kinase inhibitor therapy. *J. Clin. Oncol.* **29**, 2121–2127 (2011).
10. U. A. Matulonis, M. Hirsch, E. Palescandolo, E. Kim, J. Liu, P. van Hummelen, L. MacConaill, R. Drapkin, W. C. Hahn, High throughput interrogation of somatic mutations in high grade serous cancer of the ovary. *PLoS One* **6**, e24433 (2011).
11. J. A. Engelman, L. Chen, X. Tan, K. Crosby, A. R. Guimaraes, R. Upadhyay, M. Maira, K. McNamara, S. A. Perera, Y. Song, L. R. Chirieac, R. Kaur, A. Lightbown, J. Simendinger, T. Li, R. F. Padera, C. García-Echeverría, R. Weissleder, U. Mahmood, L. C. Cantley, K. K. Wong, Effective use of PI3K and MEK inhibitors to treat mutant *Kras* G12D and *PIK3CA* H1047R murine lung cancers. *Nat. Med.* **14**, 1351–1356 (2008).
12. P. J. Campbell, S. Yachida, L. J. Mudie, P. J. Stephens, E. D. Pleasance, L. A. Stebbings, L. A. Morsberger, C. Latimer, S. McLaren, M. L. Lin, D. J. McBride, I. Varela, S. A. Nik-Zainal, C. Leroy, M. Jia, A. Menzies, A. P. Butler, J. W. Teague, C. A. Griffin, J. Burton, H. Swerdlow, M. A. Quail, M. R. Stratton, C. Iacobuzio-Donahue, P. A. Futreal, The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109–1113 (2010).
13. S. P. Shah, R. D. Morin, J. Khattra, L. Prentice, T. Pugh, A. Burleigh, A. Delaney, K. Gelmon, R. Guliany, J. Senz, C. Steidl, R. A. Holt, S. Jones, M. Sun, G. Leung, R. Moore, T. Severson, G. A. Taylor, A. E. Teschendorff, K. Tse, G. Turashvili, R. Varhol, R. L. Warren, P. Watson, Y. Zhao, C. Caldas, D. Huntsman, M. Hirst, M. A. Marra, S. Aparicio, Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809–813 (2009).
14. M. Inukai, S. Toyooka, S. Ito, H. Asano, S. Ichihara, J. Soh, H. Suehisa, M. Ouchida, K. Aoe, M. Aoe, K. Kiura, N. Shimizu, H. Date, Presence of *epidermal growth factor receptor gene* T790M mutation as a minor clone in non–small cell lung cancer. *Cancer Res.* **66**, 7854–7858 (2006).
15. S. L. Edwards, R. Brough, C. J. Lord, R. Natrajan, R. Vatcheva, D. A. Levine, J. Boyd, J. S. Reis-Filho, A. Ashworth, Resistance to therapy caused by intragenic deletion in *BRCA2*. *Nature* **451**, 1111–1115 (2008).
16. S. Maheswaran, L. V. Sequist, S. Nagrath, L. Ulkus, B. Brannigan, C. V. Collura, E. Inserra, S. Diederichs, A. J. Iafrate, D. W. Bell, S. Digumarthy, A. Muzikansky, D. Irimia, J. Settleman, R. G. Tompkins, T. J. Lynch, M. Toner, D. A. Haber, Detection of mutations in *EGFR* in circulating lung-cancer cells. *N. Engl. J. Med.* **359**, 366–377 (2008).
17. B. Norquist, K. A. Wurz, C. C. Pennil, R. Garcia, J. Gross, W. Sakai, B. Y. Karlan, T. Taniguchi, E. M. Swisher, Secondary somatic mutations restoring *BRCA1/2* predict chemotherapy resistance in hereditary ovarian carcinomas. *J. Clin. Oncol.* **29**, 3008–3015 (2011).
18. R. J. Leary, I. Kinde, F. Diehl, K. Schmidt, C. Clouser, C. Duncan, A. Antipova, C. Lee, K. McKernan, F. M. De La Vega, K. W. Kinzler, B. Vogelstein, L. A. Diaz Jr., V. E. Velculescu, Development of personalized tumor biomarkers using massively parallel sequencing. *Sci. Transl. Med.* **2**, 20ra14 (2010).
19. D. J. McBride, A. K. Orpana, C. Sotiriou, H. Joensuu, P. J. Stephens, L. J. Mudie, E. Hämäläinen, L. A. Stebbings, L. C. Andersson, A. M. Flanagan, V. Durbecq, M. Ignatiadis, O. Kallioniemi, C. A. Heckman, K. Alitalo, H. Edgren, P. A. Futreal, M. R. Stratton, P. J. Campbell, Use of cancer-specific genomic rearrangements to quantify disease burden in plasma from patients with solid tumors. *Genes Chromosomes Cancer* **49**, 1062–1069 (2010).
20. A. A. Ahmed, D. Etemadmoghadam, J. Temple, A. G. Lynch, M. Riad, R. Sharma, C. Stewart, S. Fereday, C. Caldas, A. Defazio, D. Bowtell, J. D. Brenton, Driver mutations in *TP53* are ubiquitous in high grade serous carcinoma of the ovary. *J. Pathol.* **221**, 49–56 (2010).
21. Cancer Genome Atlas Research Network, Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
22. R. C. Bast Jr., Molecular approaches to personalizing management of ovarian cancer. *Ann. Oncol.* **22** (Suppl. 8), viii5–viii15 (2011).
23. K. C. Chan, S. F. Leung, S. W. Yeung, A. T. Chan, Y. M. Lo, Persistent aberrations in circulating DNA integrity after radiotherapy are associated with poor prognosis in nasopharyngeal carcinoma patients. *Clin. Cancer Res.* **14**, 4141–4145 (2008).
24. H. C. Fan, Y. J. Blumenfeld, U. Chitkara, L. Hudgins, S. R. Quake, Analysis of the size distributions of fetal and maternal cell-free DNA by paired-end sequencing. *Clin. Chem.* **56**, 1279–1286 (2010).
25. Y. M. Lo, K. C. Chan, H. Sun, E. Z. Chen, P. Jiang, F. M. Lun, Y. W. Zheng, T. Y. Leung, T. K. Lau, C. R. Cantor, R. W. Chiu, Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Sci. Transl. Med.* **2**, 61ra91 (2010).
26. R. W. Chiu, R. Akolekar, Y. W. Zheng, T. Y. Leung, H. Sun, K. C. Chan, F. M. Lun, A. T. Go, E. T. Lau, W. W. To, W. C. Leung, R. Y. Tang, S. K. Au-Yeung, H. Lam, Y. Y. Kung, X. Zhang, J. M. van Vugt, R. Minekawa, M. H. Tang, J. Wang, C. B. Oudejans, T. K. Lau, K. H. Nicolaides, Y. M. Lo, Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: Large scale validity study. *BMJ* **342**, c7401 (2011).
27. Y. M. Lo, N. B. Tsui, R. W. Chiu, T. K. Lau, T. N. Leung, M. M. Heung, A. Gerovassili, Y. Jin, K. H. Nicolaides, C. R. Cantor, C. Ding, Plasma placental RNA allelic ratio permits non-invasive prenatal chromosomal aneuploidy detection. *Nat. Med.* **13**, 218–223 (2007).
28. Y. M. Lo, F. M. Lun, K. C. Chan, N. B. Tsui, K. C. Chong, T. K. Lau, T. Y. Leung, B. C. Zee, C. R. Cantor, R. W. Chiu, Digital PCR for the molecular detection of fetal chromosomal aneuploidy. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 13116–13121 (2007).
29. Z. Chen, J. Feng, C. H. Buzin, Q. Liu, L. Weiss, K. Kernstine, G. Somlo, S. S. Sommer, Analysis of cancer mutation signatures in blood by a novel ultra-sensitive assay: Monitoring of therapy or recurrence in non-metastatic breast cancer. *PLoS One* **4**, e7220 (2009).
30. I. Kinde, J. Wu, N. Papadopoulos, K. W. Kinzler, B. Vogelstein, Detection and quantification of rare mutations with massively parallel sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9530–9535 (2011).
31. J. Li, L. Wang, H. Mamon, M. H. Kulke, R. Berbeco, G. M. Makrigiorgos, Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. *Nat. Med.* **14**, 579–584 (2008).
32. S. A. Forbes, N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, J. W. Teague, P. J. Campbell, M. R. Stratton, P. A. Futreal, COSMIC: Mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945–D950 (2011).
33. J. Wu, H. Matthaei, A. Maitra, M. Dal Molin, L. D. Wood, J. R. Eshleman, M. Goggins, M. I. Canto, R. D. Schulick, B. H. Edil, C. L. Wolfgang, A. P. Klein, L. A. Diaz Jr., P. J. Allen, C. M. Schmidt, K. W. Kinzler, N. Papadopoulos, R. H. Hruban, B. Vogelstein, Recurrent *GNAS* mutations define an unexpected pathway for pancreatic cyst development. *Sci. Transl. Med.* **3**, 92ra66 (2011).
34. O. Harismendy, R. B. Schwab, L. Bao, J. Olson, S. Rozenzhak, S. K. Kotsopoulos, S. Pond, B. Crain, M. S. Chee, K. Messer, D. R. Link, K. A. Frazer, Detection of low prevalence somatic mutations in solid tumors with ultra-deep targeted sequencing. *Genome Biol.* **12**, R124 (2011).
35. N. Wagle, M. F. Berger, M. J. Davis, B. Blumenstiel, M. DeFelice, P. Pochanard, M. Ducar, P. Van Hummelen, L. E. MacConaill, W. C. Hahn, M. Meyerson, S. B. Gabriel, L. A. Garraway, High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer Discov.* **2**, 82 (2012).
36. T. Sjöblom, S. Jones, L. D. Wood, D. W. Parsons, J. Lin, T. D. Barber, D. Mandelker, R. J. Leary, J. Ptak, N. Silliman, S. Szabo, P. Buckhaults, C. Farrell, P. Meeh, S. D. Markowitz, J. Willis, D. Dawson, J. K. Willson, A. F. Gazdar, J. Hartigan, L. Wu, C. Liu, G. Parmigiani, B. H. Park, K. E. Bachman, N. Papadopoulos, B. Vogelstein, K. W. Kinzler, V. E. Velculescu, The consensus coding sequences of human breast and colorectal cancers. *Science* **314**, 268–274 (2006).
37. B. Vogelstein, K. W. Kinzler, Digital PCR. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9236–9241 (1999).
38. X. Sun, K. Hung, L. Wu, D. Sidransky, B. Guo, Detection of tumor mutations in the presence of excess amounts of normal DNA. *Nat. Biotechnol.* **20**, 186–189 (2002).
39. C. Shi, S. H. Eshleman, D. Jones, N. Fukushima, L. Hua, A. R. Parker, C. J. Yeo, R. H. Hruban, M. G. Goggins, J. R. Eshleman, LigAmp for sensitive detection of single-nucleotide differences. *Nat. Methods* **1**, 141–147 (2004).
40. L. B. Pinheiro, V. A. Coleman, C. M. Hindson, J. Herrmann, B. J. Hindson, S. Bhat, K. R. Emslie, Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal. Chem.* **84**, 1003–1011 (2012).
41. M. Meyerson, S. Gabriel, G. Getz, Advances in understanding cancer genomes through second-generation sequencing. *Nat. Rev. Genet.* **11**, 685–696 (2010).
42. J. Otsuka, T. Okuda, A. Sekizawa, S. Amemiya, H. Saito, T. Okai, M. Kushima, Detection of p53 mutations in the plasma DNA of patients with ovarian cancer. *Int. J. Gynecol. Cancer* **14**, 459–464 (2004).

Acknowledgments: We thank H. Biggs, C. Hodgkin, S. Richardson, and L. Jones for assistance in sample collection, S. Aldridge for assistance in genomic analysis, and S. Tavaré, B. Davis, and M. Dunning for assistance in data analysis. **Funding:** We acknowledge the support of Cancer Research UK, the University of Cambridge, National Institute for Health Research Cambridge Biomedical Research Centre, Cambridge Experimental Cancer Medicine Centre, and Hutchison Whampoa Limited. C.P. was supported in part by the Academy of Medical Sciences, Wellcome Trust, British Heart Foundation, and Arthritis Research UK. **Author contributions:** T.F., M.M., C.P., D.G., D.W.Y.T., C.C., J.D.B., and N.R. designed the study. T.F., M.M., D.W.Y.T., F.K., J.H., A.P.M.,

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194

and N.R. developed methods. T.F., D.G., D.W.Y.T., A.M.P., and S.-J.D. collected the data. T.F., M.M., and N.R. analyzed TAm-Seq data. C.P., S.-J.D., C.C., and J.D.B. designed clinical studies and collected samples and clinical data. M.J.-L. performed pathological analysis. D.B. contributed sequencing data. T.F., M.M., C.P., D.G., D.W.Y.T., J.H., A.P.M., J.D.B., and N.R. interpreted the data. T.F., M.M., and N.R. wrote the manuscript with assistance from C.P., D.G., D.W.Y.T., A.P.M., J.D.B., and other authors. All authors approved the final manuscript. **Competing interests:** A.P.M. and F.K. hold equity in Fluidigm and may stand to gain by publication of these findings. D.B. and F.K. hold equity in Illumina and may stand to gain by publication of these findings.

Submitted 24 October 2011
Accepted 18 April 2012
Published 30 May 2012
10.1126/scitranslmed.3003726

Citation: T. Forshew, M. Murtaza, C. Parkinson, D. Gale, D. W. Y. Tsui, F. Kaper, S.-J. Dawson, A. M. Piskorz, M. Jimenez-Linan, D. Bentley, J. Hadfield, A. P. May, C. Caldas, J. D. Brenton, N. Rosenfeld, Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci. Transl. Med.* **4**, 136ra68 (2012).

Downloaded from stm.sciencemag.org on May 31, 2012

Foresight EX1030
Foresight v Personalis
Personalis EX2194