SUPPORTING MATERIALS AND METHODS

Data analysis

Panels and bin set parameters for automatic data analysis were created using GeneMapper Software version 4.0, according to the manufacturer's instructions and are provided as text files. Briefly, for each genetic locus tested by a SNaPshot mutation assay, there are four possible alleles (for deletion and insertion assays only two alleles were considered: the wild-type allele and the expected nucleotide change resulting from the specified deletion or insertion). The position of each of these alleles can be automatically captured by the analysis software upon the creation of specific bins (allele definitions). Bin parameters for each assay were initially established using Primer Focus Kit data (Life Technologies/Applied Biosystems) according to the manufacturer's recommendations and were subsequently adjusted using reference data from wild-type tumor samples and from the mutant controls used for assay validation. The panel and bin set parameters used in this study are provided as text files. Automatic mutation calling was set at a 5% sensitivity threshold. Interpretation of SNaPshot genotyping results was accomplished by automatic analysis of the raw data using the established panels and bin settings, followed by visual inspection of the spectra for all loci by at least two users (one of whom was highly experienced). In addition, if a mutation was detected, a third user reviewed the panel containing the mutation. Since spectral analysis follows a very strict set of scoring guidelines (described below), the concordance in calling between different users was extremely high.

Data analysis was performed using the following scoring criteria: ***Pass.*** For each sample, an individual SNaPshot assay passed if: (1) the peak fluorescent height for the wild type allele was ≥ 1,000 f.u. (this value was selected for being approximately 50-100 times higher than the overall background noise, however, since signal intensities may vary for different genetic analyzer instruments, this value should be adjusted by different users); and (2) the peak fluorescent height for the wild type allele in the negative control (water sample) was < 10% of the height of the wild type allele in the clinical sample. ***Mutant.*** A mutation was called for a specific assay when: (1) the

% of mutant allele for one of the 3 possible nucleotide variants, falling within its corresponding bin, was ≥ 10% (fluorescent peak height ratio of [mutant/(mutant+wild type)] alleles > 0.10), and (2) the peak fluorescence of the mutant allele was > 3 times above the background in the wild type control sample (please refer to Supporting Information Figure S3 for detailed assessment of background). Lower level mutations were also called if the % of mutant allele was ≥ 5% and the peak fluorescence of the mutant allele was > 5 times above background. For all suspected mutant samples, the SNaPshot panel containing the assay in question was repeated to confirm the initial result. ***Repeat.*** A specific panel was repeated if it contained an assay with a suspected mutation, or if it contained an assay that failed (either because: (1) the peak fluorescent height for the wild type allele was < 1,000 f.u. or (2) the negative control produced a peak fluorescent height for the wild type allele that was ≥ 10% of the height of that same peak in the test sample).

For this study, in addition to the first 8-panel SNaPshot run, we had to repeat 1 to 2 panels per sample on average due to QC issues. Since running one panel consumes ~20 ng of DNA, for a complete genotyping data set we used on average ~200 ng of DNA (or 600 ng of total nucleic acid). If a mutation is subsequently confirmed by Sanger sequencing, an additional 40 ng of DNA is required. Hence, we typically used <250 ng of DNA (or <750 ng of total nucleic acid) for genetic profiling of FFPE-derived tumor specimens.

Assay Validation and Sensitivity Assessment

The tumor genotyping assay described here consists of 8 SNaPshot multiplex panels that test for 58 commonly mutated loci in 13 cancer genes. Since multiple nucleotide variants have been described at most of these loci, the assay can detect 120 previously described mutations (Supporting Information Table S1). The frequency of occurrence of each allele variant was calculated using data compiled by the Wellcome Trust Sanger Institute and reported for each cancer gene in the COSMIC database (Bamford et al., 2004) (v42 release). To calculate the frequencies of gene mutation depicted in Table 1 and Supporting Information Table S1, we

included all mutations described in the COSMIC database with available positional information at the amino acid level.

We validated 81 out of the 120 allele variants covered by our panel, using three types of control samples (Supporting Information Table S1): (1) whenever possible, we used primary tumor samples that had been previously tested at the MGH Molecular Diagnostics Pathology Laboratory and shown to carry the mutations of interest; (2) for the majority of the assays, validation was performed using cancer cell lines harboring known mutations, which were identified using the Wellcome Trust Sanger Institute Cancer Cell Line Project database; (3) synthetic oligonucleotides harboring the mutation of interest were designed to validate those allele variants for which we were unable to identify an appropriate tumor sample or cancer cell line control (Supporting Information Table S7).

Genomic DNA was extracted from blood using the QIAamp Blood kit (QIAGEN Inc., Valencia, CA), or from FFPE primary tumor tissue and frozen cancer cell line pellets, using the RecoverALL™ Total Nucleic Acid Isolation Kit (Applied Biosystems, Foster City, CA), according to the manufacturer's recommendations. To prepare the synthetic control samples, 1 to 40 pmol of custom-made oligonucleotides designed to include the mutation of interest, were added to 3 µl of PCR product obtained from amplification of 20ng of male genomic DNA (Promega, Madison, WI) as indicated in Supporting Information Table S7, followed by Exo/SAP treatment and by the extension reaction. Each mutant sample was tested using the SNaPshot genotyping panel containing the assay to be validated, and male genomic DNA (Promega, Madison, WI) was used as a wild-type control for each run.

For those allele-specific assays that could be validated using genomic DNA derived from primary tumor tissue or from cancer cell lines, a sensitivity assessment was also performed (Supporting Information Fig. S2). For sensitivity testing, mutant DNA samples were serially diluted in 1:3

increments with male genomic DNA (Promega), to obtain solutions of 100%, 30%, 10%, 3%, and 1% of mutant DNA input material.

It is well established that cancer cells are prone to genetic instability, which can result in the gain or loss of genetic material. In addition, primary tumor specimens may contain normal (non-cancerous) cells. Due to this heterogeneity, the calculated amount of input mutant DNA material does not accurately reflect the relative amount of mutant vs. wt allele in each tested sample. Thus, the percentage of mutant allele in each sample was calculated by comparing the fluorescent peak heights of the mutant and wild-type alleles, according to the following: % mutation = [mutant allele peak height / (wild-type allele peak height + mutant allele peak height)]*100.

The sensitivity of each assay was established as the lowest % mutation for which the fluorescent peak height of the mutant allele is > 3x background (the background for a specific mutant allele is defined as the height of the fluorescent peak corresponding to that allele, within its assigned bin in the wild type control genomic DNA sample). For a detailed explanation of the process involved in sensitivity assessment, please refer to Supporting Information Figure S3.

Independent confirmation of test results

All of the mutations detected in a primary tumor sample were initially verified by an independent SNaPshot reaction using the genotyping panel containing the assay in question. The cases of chronic myeloproliferative disease and a small number of colorectal adenocarcinomas had been previously sequenced for *JAK2* exon 12 and for *KRAS* exon 2, respectively, as part of standard clinical testing. Once our genotyping analysis was completed, the original data was disclosed and we confirmed that the SNaPshot results matched the previous clinical findings. We evaluated the additional mutations using standard Sanger sequencing. In total, we independently confirmed 90% of the mutations identified by SNaPshot genotyping (inability to independently verify the

presence of mutation in 10% of the cases was due to unsuccessful Sanger sequencing data, which we attribute to the limiting amounts of remaining nucleic acid).

Mutational profiling of 250 primary tumor samples identified a total of 100 mutations that could be classified into 33 distinct mutation groups. We attempted to identify cases with normal matching tissue for each of these 33 independent mutation types, and do a side-by-side comparison between tumor and normal tissue from the same individual, to test the specificity of the SNaPshot assay. We were able to perform this analysis for 25 out of the 33 mutation types (76%), and verified that in all cases the somatic mutant allele was only detected in the tumor specimen and not in the matching normal tissue, which confirmed the specificity of the corresponding SNaPshot assays.

Sequencing analysis

Traditional Sanger sequencing was performed in a volume of 20 μl, containing 1 unit of Taq polymerase (Invitrogen, Carlsbad, CA), 4 nmol of dNTPs (Invitrogen, Carlsbad, CA), 10 pmol of forward (a1) and reverse (a2) primers, 40 nmol of $MgCl_2$ (or the amount indicated in Supporting Information Table S8), and either 40 ng of genomic DNA or 120 ng of total nucleic acid. We initially attempted sequencing with the same amplification primers and cycling parameters used for SNaPshot multiplexed PCR. For those cases where this strategy was not successful, new primers were designed (Supporting Information Table S8) and the cycling conditions were: 94 ºC for 5 min, followed by 38 cycles of 94 ºC for 30 s, a specific annealing temperature for 30 s and 72 ºC for 45 sec, and one last cycle of 72 ºC for 10 min. The annealing temperature and amount of $MgCl_2$ used for each PCR are detailed in Supporting Information Table S8. The resulting PCR products were treated using 1 unit of shrimp alkaline phosphatase (USB, Cleveland, OH) and 5 units of exonuclease I (USB, Cleveland, OH) at 37°C for 20 minutes followed by 80°C for 15 minutes, and tested for the presence of mutations by bi-directional Sanger sequencing using the BigDye Terminator V1.1 Cycle Sequencing Kit (Applied Biosystems), according to the manufacturer's recommendations. The sequencing reaction step was performed with the original

PCR primers or with the incorporated M13 tags. Tumor and control human genomic DNA (Promega, Madison, WI) sequences were compared using the AB Sequencing Analysis Software v5.2 (Applied Biosystems).

EGFR exon 19 sizing assay

A PCR-based strategy was developed to identify insertions or deletion mutations in exon 19 of the *EGFR* gene, which is a hotspot region for deletions. Amplification primer sequences were as follows, with the forward primer being 5'-labeled with the NED fluorophore: NED-EGFR_Ex19_F [0.1 µM]: 5'-NED-GCACCATCTCACAATTGCCAGTTA-3'; EGFR-Ex19-REV1 [0.1 µM]: 5'-AAAAGGTGGGCCTGAGGTTCA-3'. 20 ng of DNA template was amplified using Platinum Taq polymerase in the presence of 2 mM $MgCl_2$ (Invitrogen, Carlsbad, CA). The 20 µl reaction was subjected to 5 minutes of denaturation at 94°C and 40 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds and elongation at 72°C for 60 seconds. Following PCR amplification, products were diluted 1:30 in water and a 1 µl aliquot was added to 9.9 µl of Hi-Di Formamide and 0.1 µl of GeneScan 500 LIZ Size Standard (Applied Biosystems Inc, Foster City, CA). Heat-denatured samples were analyzed through capillary electrophoresis using the automated ABI 3730 DNA Analyzer with GeneMapper software (Applied Biosystems Inc). Insertions or deletions were visualized by shifts in molecular weight of the fluorescently-identifiable PCR amplicon relative to wild-type.

Comparison with other platforms: costs and other considerations

Infrastructure. Data analysis by SNaPshot genotyping uses a capillary-based DNA sequencer. These instruments are ubiquitously present in clinical laboratories for dideoxy (Sanger) sequencing. In contrast, other multiplexed allele-targeted technologies and next-generation sequencing platforms use costly instrumentation, which can reach several hundreds of thousands of US dollars. In addition to equipment costs, the analysis of next-generation sequencing data would require an investment in bioinformatics personnel.

Workflow. As mentioned in the main text, in our early experience implementing this assay, it takes on average two to three weeks from the time of test requisition until genotyping report finalization. This includes specimen identification, collection of the FFPE tissue blocks and clinical slides, pathology review, specimen processing, DNA extraction, SNaPshot genotyping, data analysis, and report preparation.

One round of SNaPshot genotyping and data analysis (testing the full 8 panels once) takes 1.5 to 2 days total including incubation steps. This time is very comparable to Sanger sequencing and other SNP-based allele-targeted sequencing methods (such as Sequenom MasARRAY and the SNPstream technology from Beckman Coulter). Next generation sequencing options (such as SOLiD from ABI and the Illumina platforms) have a longer time frame, including library construction and processing, as well as the actual sequencing and data analysis, which collectively can take ~10 days.

Tissue requirements and assay cost. The SNaPshot test described here assays for recurrent mutations in a total of 23 exons across 13 cancer genes and uses ~20 ng of FFPE-derived tumor DNA for each of the 8 panels. Assuming that optimal Sanger sequencing data from FFPE material would require something in the range of 20-40 ng of DNA per reaction (note that 40 ng is the amount used in the sequencing procedure presented in this study to confirm all mutations identified by SNaPshot), Sanger sequencing of the same 23 cancer exons would use about 3-6 times more tumor DNA than SNapshot. We estimate that the direct costs for running SNaPshot are about 50-80% of Sanger sequencing the 23 cancer exons (considering the price range in Sanger sequencing by different laboratories). In the United States, the amount that is charged to patients and insurance companies for a specific laboratory test is based on a series of Chemistry Pathology and Laboratory (CPT) codes published by the American Medical Association, each of them describing one procedure used in the analysis (for instance: “extraction of highly purified nucleic acid”, “amplification of patient nucleic acid”, etc). Comparison of the fees ascribed by the Massachusetts General Hospital to each of these codes revealed that the charge for SNaPshot genotyping is equivalent to Sanger sequencing of 23 exons. Thus, the main cost difference between the two methods is not financial, but rather the amount of tissue. While Sanger will

provide full sequencing information, SNaPshot will focus on targeted alleles enabling the use of a much lower amount of tumor DNA, which in many cases is the limiting factor of this analysis.
Of note, from our brief experience with the Sequenom MassARRAY platform (described in detail in Supporting Information Fig. S9), we concluded that it is hard to extrapolate assay costs without trying a specific platform on a desired application. List prices for allele-targeted sequencing platforms are generally based on highly multiplexed panels used for SNP detection in high quality DNA. Our own experience and published data (Thomas et al., 2007) show that, when employed for rare mutation profiling, the multiplexing ability of the Sequenom MassARRAY is restricted to less than 10 assays per panel, which would increase the assay cost. Also, there are many options for next generation sequencing, especially at the step of library construction. Until one of these platforms is shown to detect low level mutations on FFPE tumor tissue, it will be hard to estimate assay price and the actual amount of nucleic acid required for the analysis.

SUPPORTING REFERENCES

Bamford, S., Dawson, E., Forbes, S., Clements, J., Pettett, R., Dogan, A., Flanagan, A., Teague, J., Futreal, P.A., Stratton, M.R. *et al*. (2004) The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer*, **91**, 355-358.

Thomas, R.K., Baker, A.C., Debiasi, R.M., Winckler, W., Laframboise, T., Lin, W.M., Wang, M., Feng, W., Zander, T., MacConaill, L. *et al*. (2007) High-throughput oncogene mutation profiling in human cancer. *Nat Genet*, **39**, 347-351.

SUPPLEMENTARY TABLE LEGENDS

Supplementary Table 1. Assays and somatic mutations covered by SNaPshot genotyping. Mutation data was compiled using the COSMIC database v42 release.

Supplementary Table 2. Primary cancer samples and tumor genotyping data. Of note, unless otherwise noted as mutant, all 250 tumor specimens yielded a wild-type allele for each of the 58 cancer loci covered by SNaPshot genotyping. All clinical cases were also tested for common deletions in *EGFR* exon 19 using a PCR-based sizing assay, and the results are indicated.

Supplementary Table 3. Mutation distribution across tumor types. Frequency of mutation for individual genes: observed in this study (top) and previously documented (bottom). Previously reported data was obtained from the COSMIC database v42, using the following tissue (histology - sub-histology) selections: breast (carcinoma, carcinoma in situ, and other carcinoma); haematopoietic and lymphoid (haematopoietic neoplasm - myeloproliferative disease); large intestine (adenoma, carcinoid-endocrine tumor, carcinoid tumor, and carcinoma); lung (carcinoid-endocrine tumor, and carcinoma); all (malignant melanoma); pancreas (carcinoid-endocrine tumor, carcinoma, and endocrine tumor) and prostate (carcinoma - adenocarcinoma).

Supplementary Table 4. Assessment of sample heterogeneity in primary tumors with mutations, and evaluation of patient response to EGFR tyrosine kinase inhibitor therapy. In an attempt to evaluate the degree of sample heterogeneity in our clinical cohort, we conducted histological re-evaluation of the primary tumors harboring mutations to estimate the degree of stromal contamination within the tumor tissue that had been extracted for analysis (column 2), and we calculated the % of mutant allele for each SNaPshot assay (column 3). To begin to address how the abundance of sensitizing mutations in *EGFR* may impact clinical outcome, we

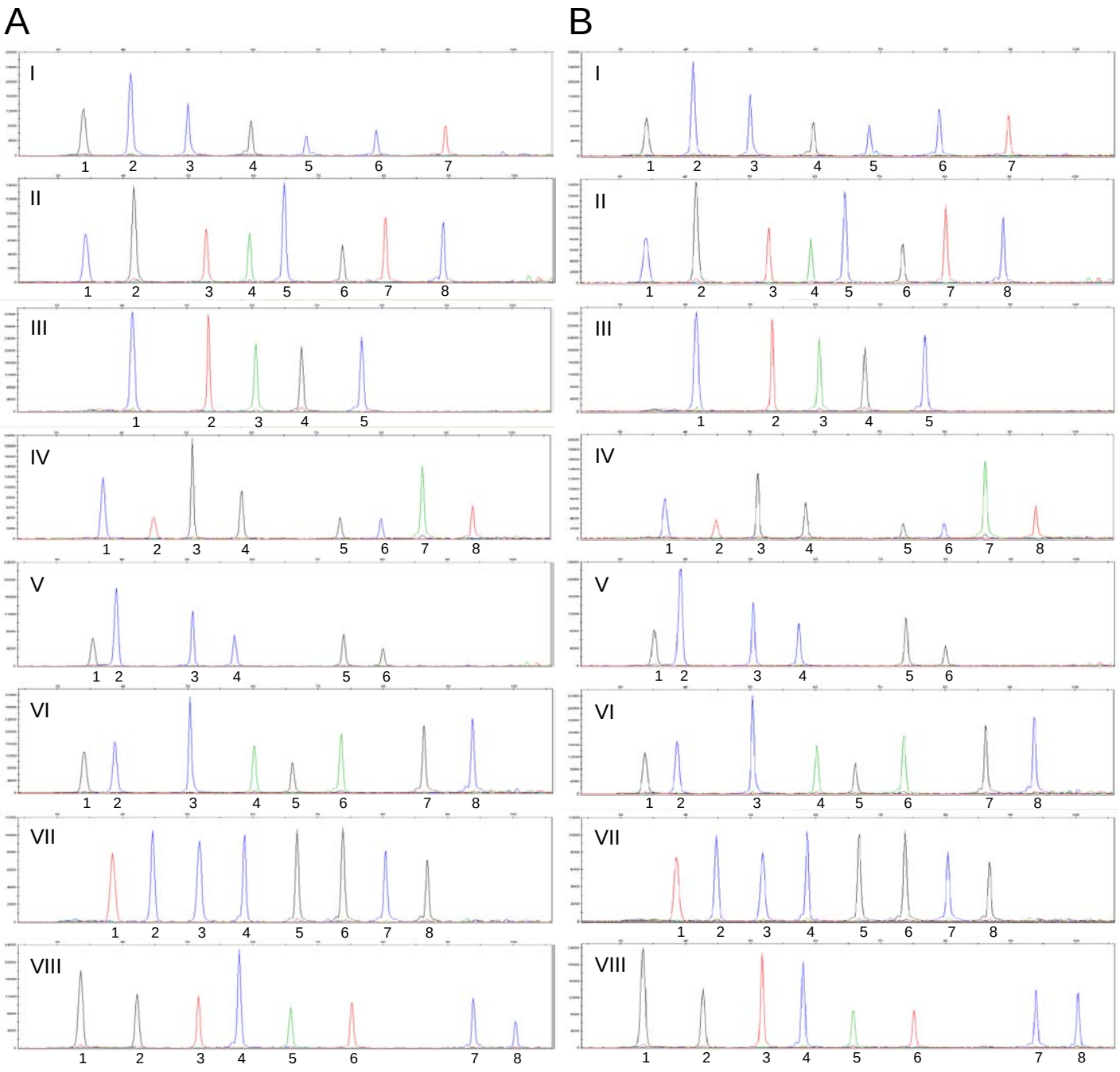
investigated whether patients with EGFR activating mutations responded to EGFR TKIs. Available clinical data is shown in column 5.

Supplementary Table 5. Primer sequences. (A) Amplification primers. **(B)** Extension primers. Note: extF refers to extension assays designed in the Forward orientation (coding strand); and extR refers to extension assays designed in the Antisense orientation (non-coding strand).

Supplementary Table 6. Genotyping conditions. (A) Multiplex PCR. **(B)** Single-base extension reaction.

Supplementary Table 7. Synthetic oligonucleotides used for assay validation.

Supplementary Table 8. Sequencing primers.



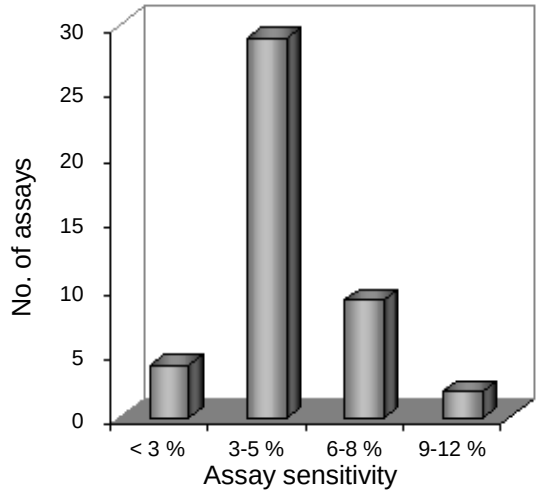
Supplementary Figure 1. Representative spectra of the 58 SNaPshot assays. The test performs equally well with **(A)** 20 ng of commercially available high-quality genomic DNA (Promega) and **(B)** 60 ng of total nucleic acid extracted from FFPE primary tumor tissue. Assays: **I.** (1) KRAS 34; (2) EGFR 2235_49del F; (3) EGFR 2369; (4) NRAS 181; (5) PIK3CA 1633; (6) CTNNB1 94 and (7) CTNNB1 121. **II.** (1) EGFR 2235_49del R; (2) NRAS 38; (3) BRAF 1799; (4) NRAS 182; (5) PIK3CA 263; (6) TP53 742; (7) CTNNB1 95 and (8) CTNNB1 122. **III.** (1) EGFR 2236_50del F; (2) EGFR 2573; (3) CTNNB1 133; (4) PIK3CA 1624 and (5) NRAS 35. **IV.** (1) KRAS 35; (2) EGFR 2236_50del R; (3) PTEN 517; (4) TP53 733; (5) FLT3 2503; (6) PIK3CA 3139; (7) NOTCH1 4724 and (8) NOTCH1 4802. **V.** (1) CTNNB1 110; (2) KRAS 38; (3) CTNNB1 134; (4) TP53 743; (5) TP53 817 and (6) APC 4666_67insA. **VI.** (1) CTNNB1 98; (2) KRAS 37; (3) EGFR 2155; (4) KIT 2447; (5) PIK3CA 3145; (6) PIK3CA 1637; (7) APC 4012 and (8) TP53 818. **VII.** (1) PIK3CA 3140; (2) CTNNB1 101; (3) JAK2 1849; (4) BRAF 1798; (5) NRAS 37; (6) PIK3CA 1636; (7) APC 4348 and (8) APC 3340. **VIII.** (1) NRAS 34; (2) PTEN 388; (3) CTNNB1 109; (4) PTEN 697; (5) PTEN 800delA; (6) NRAS 183; (7) TP53 524 and (8) TP53 916.

A

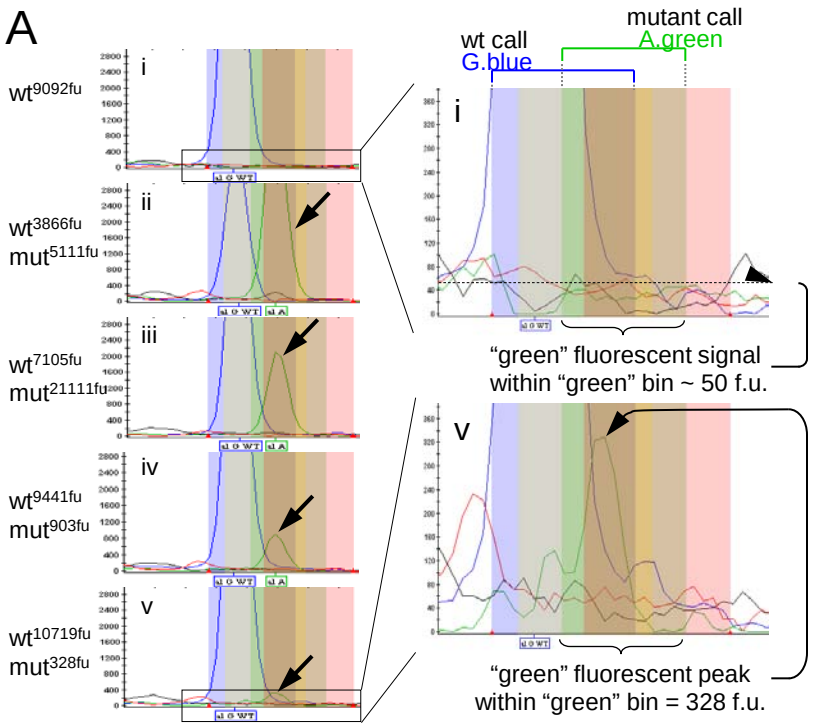
Mutation assay	Validation control sample	Sensitivity
APC 3340C>T	cell line (LoVo)	6.82%
APC 4012C>T	cell line (SW620)	4.60%
BRAF 1799T>A	primary tumor (FFPE_NA08-249)	4.90%
CTNNB1 121A>G	cell line (A-427)	2.80%
CTNNB1 134C>T	cell line (LS174T)	1.70%
CTNNB1 98C>A	cell line (SW48)	5.63%
CTNNB1 98C>T	cell line (SW1573)	4.32%
EGFR 2155G>A	cell line (SW48)	2.13%
EGFR 2235_49del15 F	cell line (PC9)	4.20%
EGFR 2235_49del15 R	cell line (PC9)	1.76%
EGFR 2236_50del15 F	primary tumor (FFPE_NA08-0247)	1.40%
EGFR 2236_50del15 R	primary tumor (FFPE_NA08-0247)	8.51%
EGFR 2369C>T	cell line (NCI-H1975)	3.20%
EGFR 2573T>G	cell line (NCI-H1975)	5.30%
FLT3 2503G>T	cell line (MO-4)*	4.40%
JAK2 1849G>T	primary tumor (blood_NA08-0257)	3.00%
KRAS 34G>A	cell line (A549)	3.93%
KRAS 34G>C	cell line (Cal-62)	5.49%
KRAS 34G>T	cell line (KYSE-410)	7.60%
KRAS 35G>A	cell line (A427)	2.97%
KRAS 35G>T	cell line (LCLC97TMI)	5.80%
KRAS 38G>A	cell line (LoVo)	2.70%
NRAS 34G>T	cell line (MOLT-4)	3.40%
NRAS 35G>A	cell line (PA-1)	5.90%
NRAS 35G>T	cell line (GA-10)	3.20%
NRAS 37G>C	cell line (K052)	6.89%
NRAS 181C>A	cell line (HMV-11)	7.20%
NRAS 182A>T	cell line (BFTC-905)	4.44%
PIK3CA 1624G>A	cell line (Cal51)	4.60%
PIK3CA 1633G>A	cell line (BFTC-909)	3.92%
PIK3CA 1637A>G	cell line (22RVI)	3.15%
PIK3CA 263G>A	cell line (SNG-M)	2.84%
PIK3CA 3139C>T	cell line (MFE-280)	4.66%
PIK3CA 3140A>G	cell line (LS174T)	5.10%
PIK3CA 3145G>C	cell line (HEC-1)	4.32%
PTEN 517C>T	cell line (639V)	5.15%
PTEN 697C>T	cell line (SF295)	3.00%
PTEN 800delA	cell line (MOLT-4)	7.60%
TP53 524G>A	cell line (VM-CUB1)	5.59%
TP53 742C>T	cell line (Colo680N)	3.90%
TP53 743G>A	cell line (639V)	11.43%
TP53 818G>A	cell line (NCI-H1975)	4.45%
TP53 818G>T	cell line (HCC38)	4.90%
TP53 916C>T	cell line (MOLT-4)	5.30%

* College of American Pathologists (proficiency testing control)

B



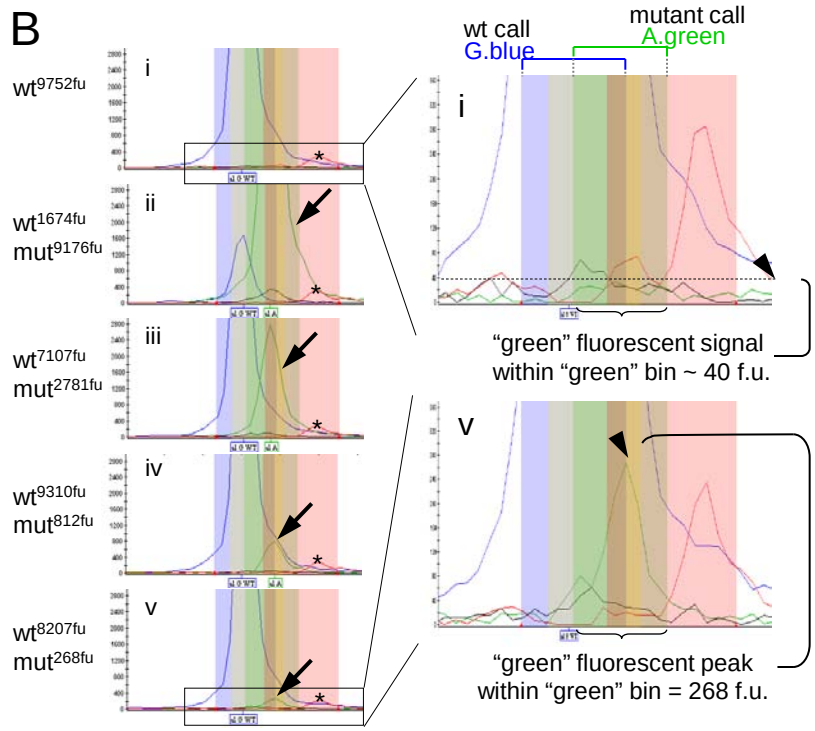
Supplementary Figure 2. Sensitivity assessment. (A) Assay sensitivity and validation results. **(B)** Graphical representation of the sensitivity data, which is on average 4.64%. A few examples of assay sensitivity are presented in Figure 2 and Supplementary Figure 4. A detailed illustration of data collection and the calculations involved in sensitivity assessment can be found in Supplementary Figure 3.



A427 cell line
KRAS G12D (c.35G>A)

Sample	Genomic DNA (Promega)	Cell line DNA (A427)	MUTANT % $\left[\frac{\text{MUT}*100}{\text{WT+MUT}}\right]$
i	100 %	-	0.5%
ii	-	100 %	57%
iii	70 %	30 %	23%
iv	90 %	10 %	8.7%
v	97 %	3 %	3.0%
vi	99 %	1 %	0.6%

Sample	$\frac{\text{MUT peak}_{\text{A427 DNA at each dilution}}}{\text{MUT signal within MUT bin}_{\text{Genomic DNA}}}$
i	-
ii	5111/50 = 102 > 3
iii	2111/50 = 42 >3
iv	903/50 = 18 > 3
v	328/50 = 6.6 >3
vi	70/50 = 1.4 < 3

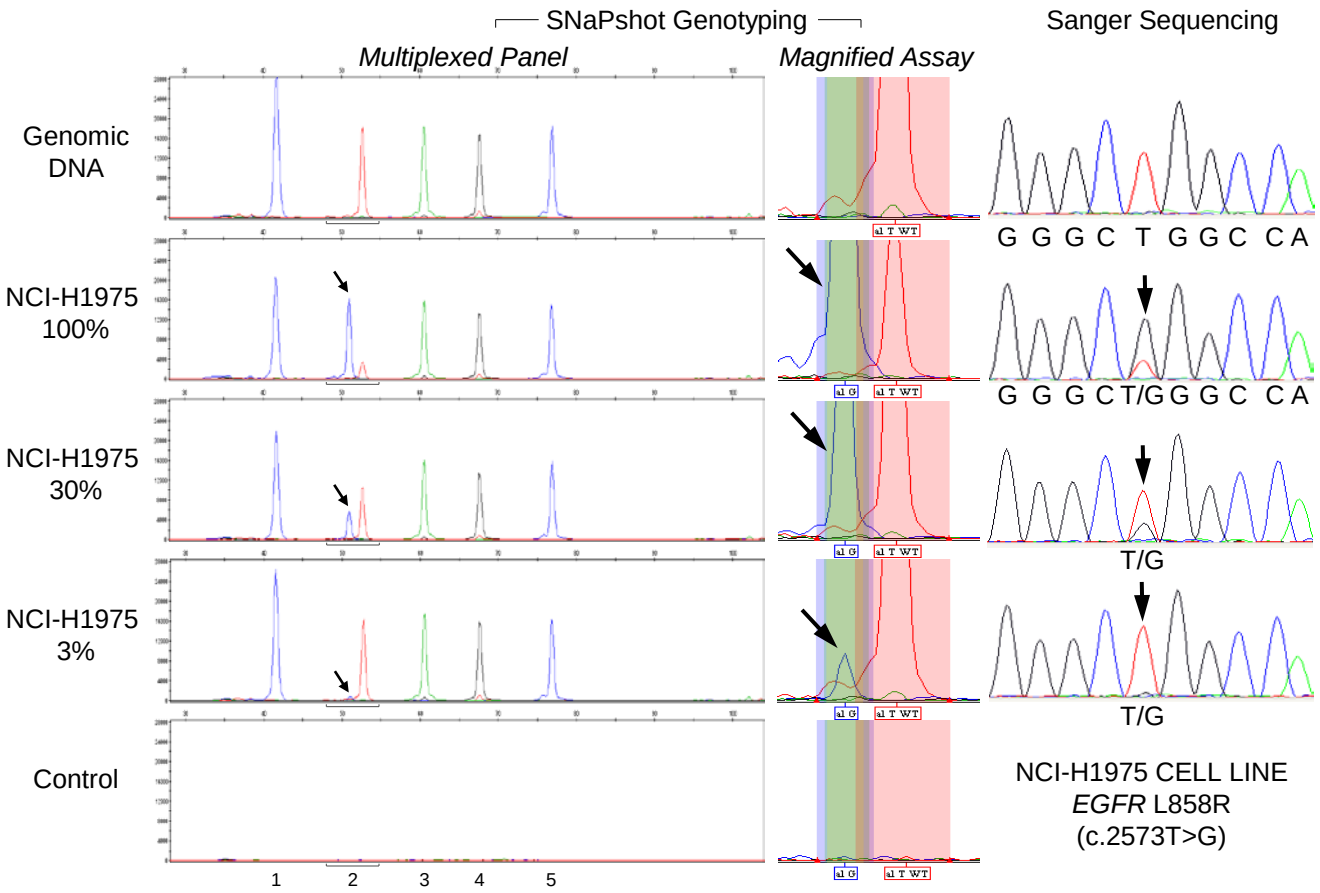


NCI-H1975 cell line
EGFR T790M (c.2369C>T)

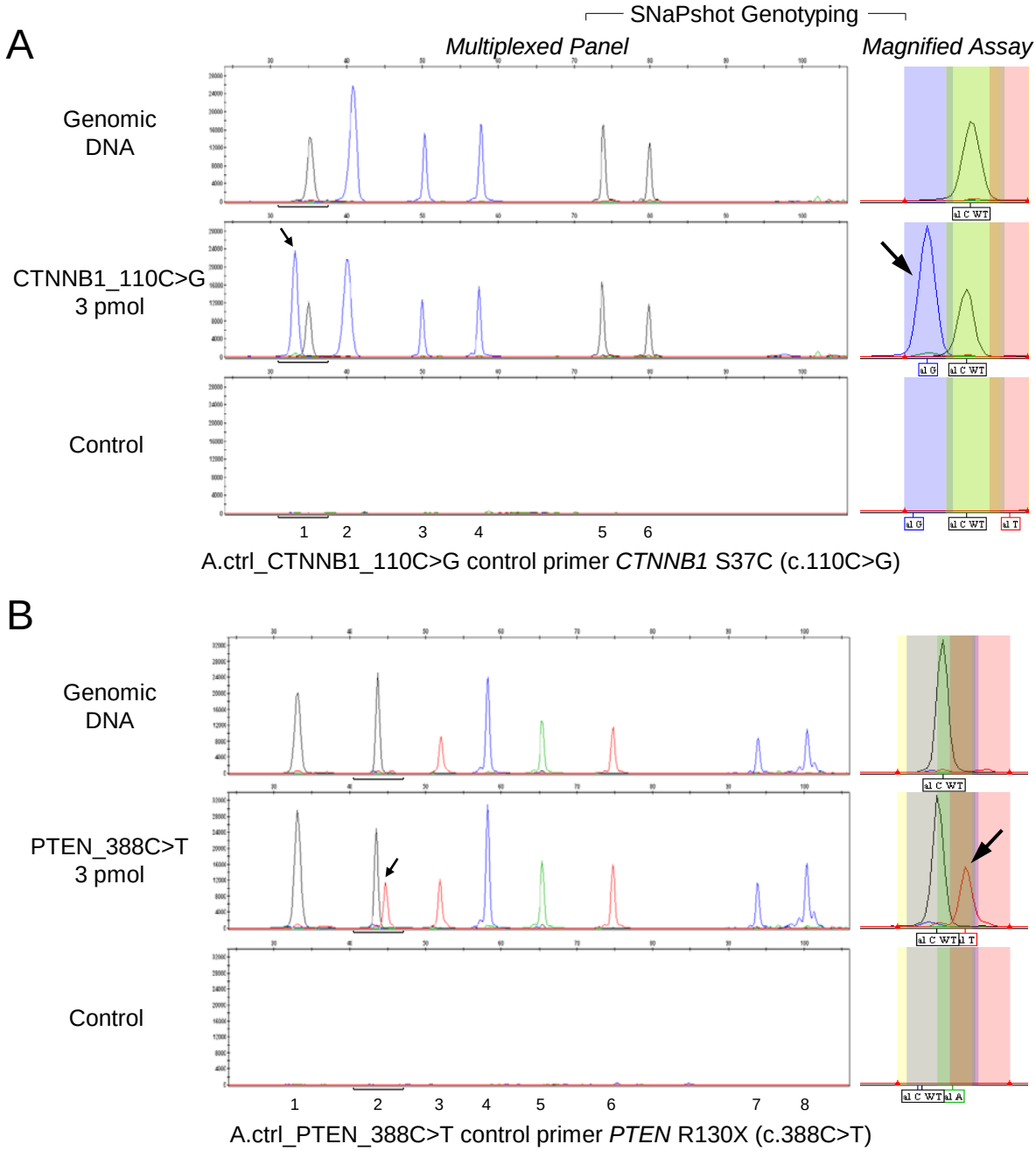
Sample	Genomic DNA (Promega)	Cell line DNA (H1975)	MUTANT % $\left[\frac{\text{MUT}*100}{\text{WT+MUT}}\right]$
i	100 %	-	0.4%
ii	-	100 %	85%
iii	70 %	30 %	28%
iv	90 %	10 %	8.0%
v	97 %	3 %	3.2%
vi	99 %	1 %	1.4%

Sample	$\frac{\text{MUT peak}_{\text{H1975 DNA at each dilution}}}{\text{MUT signal within MUT bin}_{\text{Genomic DNA}}}$
i	-
ii	9176/40 = 229 > 3
iii	2781/40 = 69 >3
iv	812/40 = 20 > 3
v	268/40 = 6.7 >3
vi	150/40 = 3.7 ~ 3

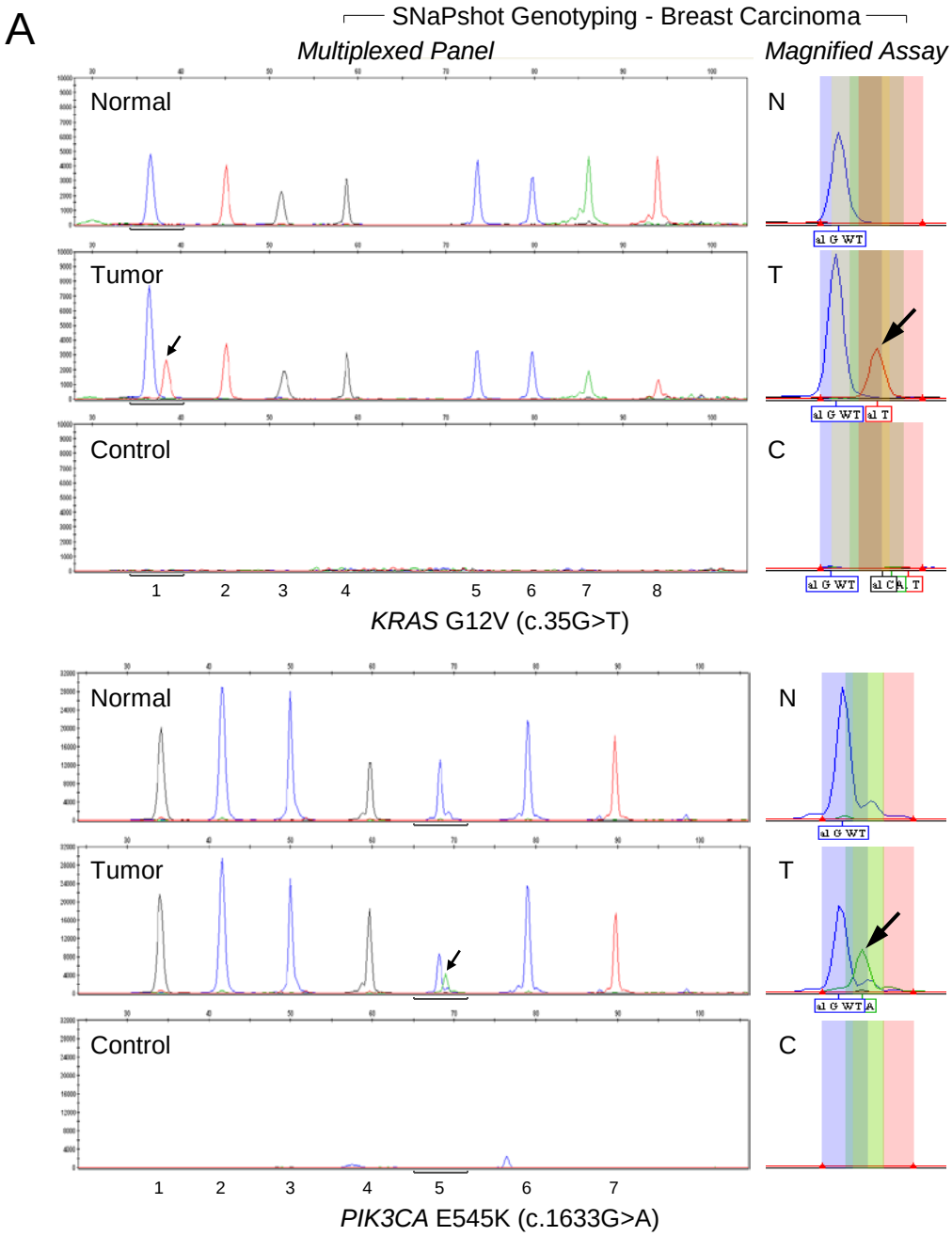
Supplementary Figure 3. In-depth view of the sensitivity assessment illustrated in Figure 2. The section on the left represents the SNaPshot assay being tested, with the sizes of wild-type and mutant alleles indicated on the left (f.u. = fluorescence units). Arrows in the high-power images in the middle section point to the background noise within the mutant bin in the genomic DNA sample (top) and to the mutant allele in the 3% dilution of the mutant sample (bottom). The top table depicts the levels of genomic (wild-type) and cell line (mutant) DNA within each sample, and the percentage of mutant allele obtained for each SNaPshot assay, calculated as a ratio of fluorescent peak heights [mutant*100/(wild type+mutant)]. The bottom table illustrates the calculations that selected the sample used to determine the sensitivity. Sensitivity of an assay was established as the lowest percentage of mutation in the test sample (blue arrow) yielding a mutant allele peak that was > 3 times the background noise in the wild type sample (red arrow). **(A)** The sensitivity of the KRAS G12D (c.35G>A) SNaPshot assay is 3.0% (blue arrow), which was determined using the sample with 3% of A427 cell line DNA (red arrow). **(B)** The sensitivity of the EGFR T790M (c.2369C>T) SNaPshot assay is 3.2% (blue arrow), which was established using the sample containing 3% of NCI-H1975 cell line DNA (red arrow).



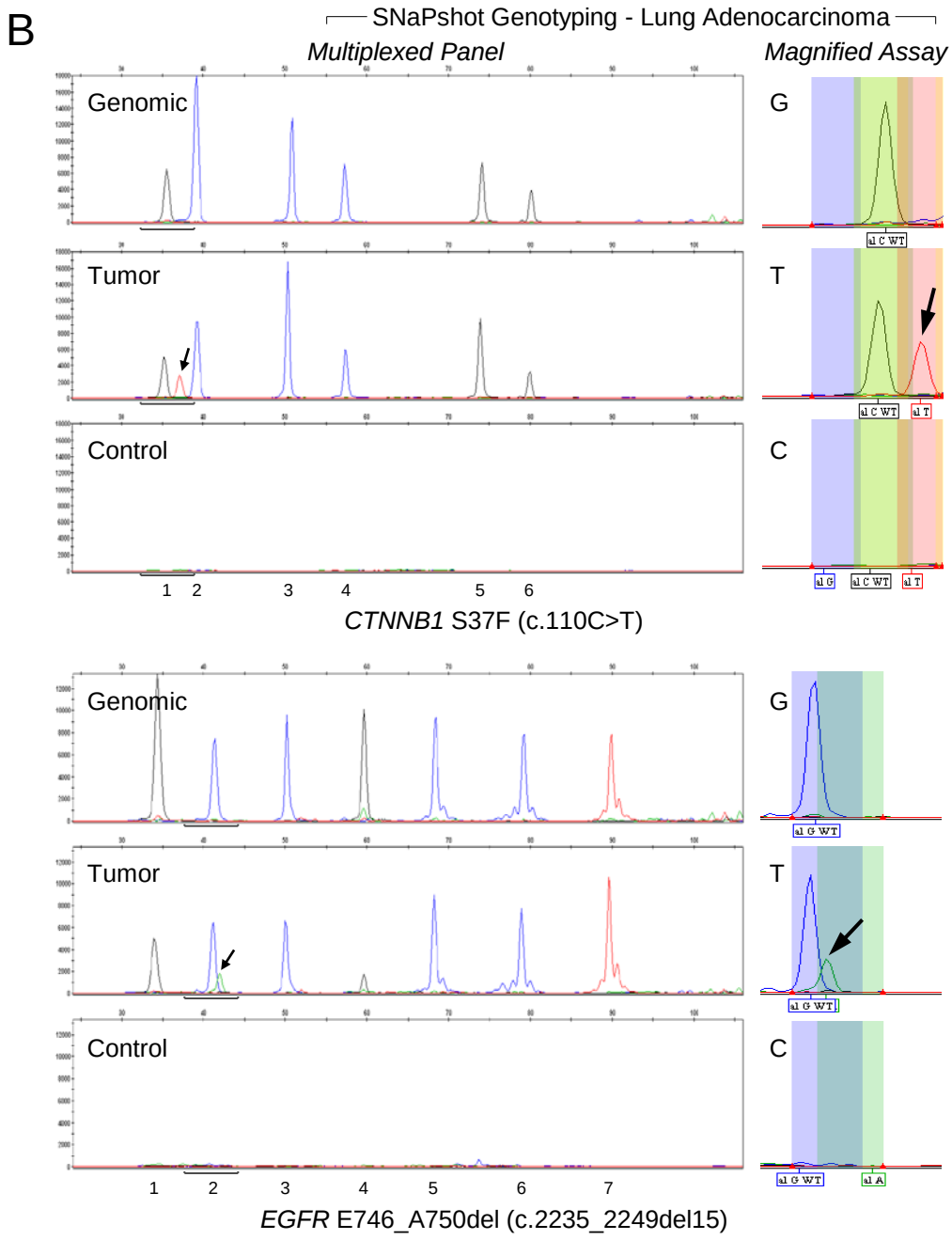
Supplementary Figure 4. Sensitivity testing using cancer cell line DNA. The NCI-H1975 lung adenocarcinoma cell line was used to identify the *EGFR* L858R (c.2573T>G) mutation. Sensitivity was ~5%. Assays: (1) EGFR 2236_50del F; (2) EGFR 2573; (3) CTNNB1 133; (4) PIK3CA 1624 and (5) NRAS 35.



Supplementary Figure 5. Assay validation using synthetic oligonucleotides. Synthetic DNA primers designed to harbor specific mutations (Supplementary Table 8) were used to validate SNaPshot assays for which we lacked primary tumor or cell line controls. Both cases illustrate the genotyping results obtained using wild-type genomic DNA (Promega) (top), 3 pmol of synthetic oligonucleotide added to wild-type genomic DNA (middle), and a no-DNA control (bottom). **(A)** The A.ctrl_CTNNB1_110C>G control primer was used to identify the *CTNNB1* S37C (c.110C>G) mutation. Assays: (1) CTNNB1 110; (2) KRAS 38; (3) CTNNB1 134; (4) TP53 743; (5) TP53 817 and (6) APC 4666_67insA. **(B)** The A.ctrl_PTEN_388C>T control primer was used to detect the *PTEN* R130X (c.388C>T) mutation. Assays: (1) NRAS 34; (2) PTEN 388; (3) CTNNB1 109; (4) PTEN 697; (5) PTEN 800delA; (6) NRAS 183; (7) TP53 524 and (8) TP53 916.



Supplementary Figure 6. Examples of rare mutations detected by SNaPshot genotyping. (A) Co-occurrence of the *KRAS* G12V (c.35G>T) (upper) and *PIK3CA* E545K (1633G>A) (lower) mutations in a case of breast lobular carcinoma. Both images show genotyping data obtained using total nucleic acid extracted from normal (top) and tumor (middle) FFPE tissue from the same individual, and a no-DNA negative control (bottom). Upper image assays: (1) KRAS 35; (2) EGFR 2236_50del R; (3) PTEN 517; (4) TP53 733; (5) FLT3 2503; (6) PIK3CA 3139; (7) NOTCH1 4724 and (8) NOTCH1 4802. Lower image assays: (1) KRAS 34; (2) EGFR 2235_49del F; (3) EGFR 2369; (4) NRAS 181; (5) PIK3CA 1633; (6) CTNNB1 94 and (7) CTNNB1 121.

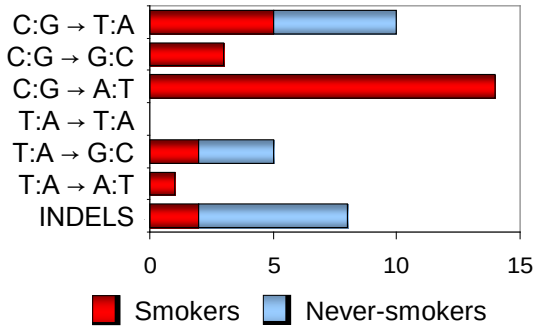


Supplementary Figure 6 (cont'd). Examples of rare mutations detected by SNaPshot genotyping. (B) Co-ocurrence of the *CTNNB1* S37F (c.110C>T) and *EGFR* E746_A750del (c.2235_2249del15) mutations in a case of fetal lung adenocarcinoma. Both images show the results obtained using wild type genomic DNA (Promega) (top), total nucleic acid extracted from FFPE primary tumor tissue (middle), and a no-DNA negative control (bottom). Upper image assays: (1) CTNNB1 110; (2) KRAS 38; (3) CTNNB1 134; (4) TP53 743; (5) TP53 817 and (6) APC 4666_67insA. Lower image assays: (1) KRAS 34; (2) EGFR 2235_49del F; (3) EGFR 2369; (4) NRAS 181; (5) PIK3CA 1633; (6) CTNNB1 94 and (7) CTNNB1 121.

A

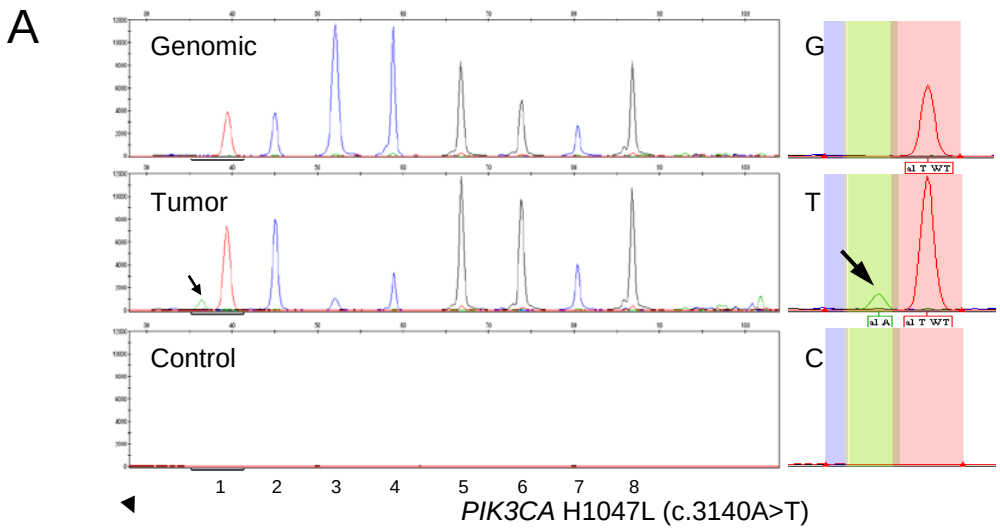
Tumor Type	Breast	CMD	Colorectal	Lung	Melanoma	Pancreatic	Prostate	Other
Total no. of cases	33	10	30	87	11	23	20	36
Total no. of mutations	7	4	21	41	7	13	4	3
C:G → T:A	3	4	15	10	1	5	1	2
C:G → G:C	0	0	0	3	0	2	2	0
C:G → A:T	1	0	4	14	0	6	1	0
T:A → C:G	2	0	0	0	1	0	0	0
T:A → G:C	0	0	0	5	0	0	0	0
T:A → A:T	1	0	2	1	5	0	0	1
INDELS	0	0	0	8	0	0	0	0

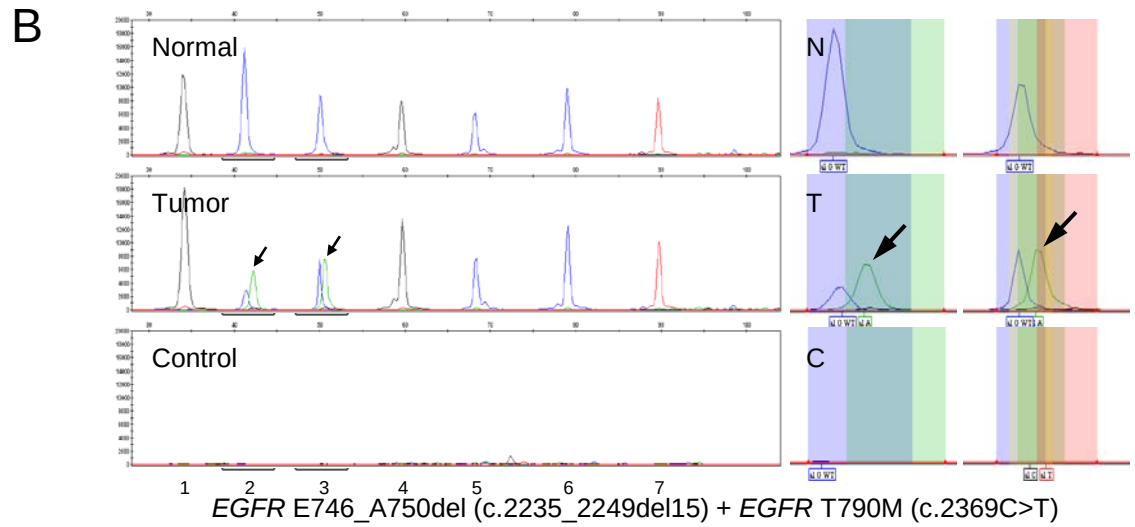
B

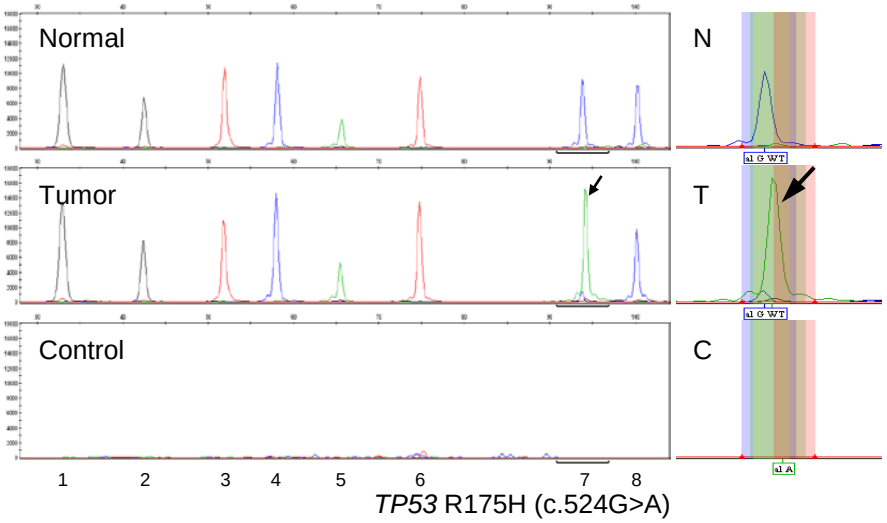


	Mutant cases	
	Lung	Pancreatic
Smokers	49%	67%
Never-smokers	28%	13%

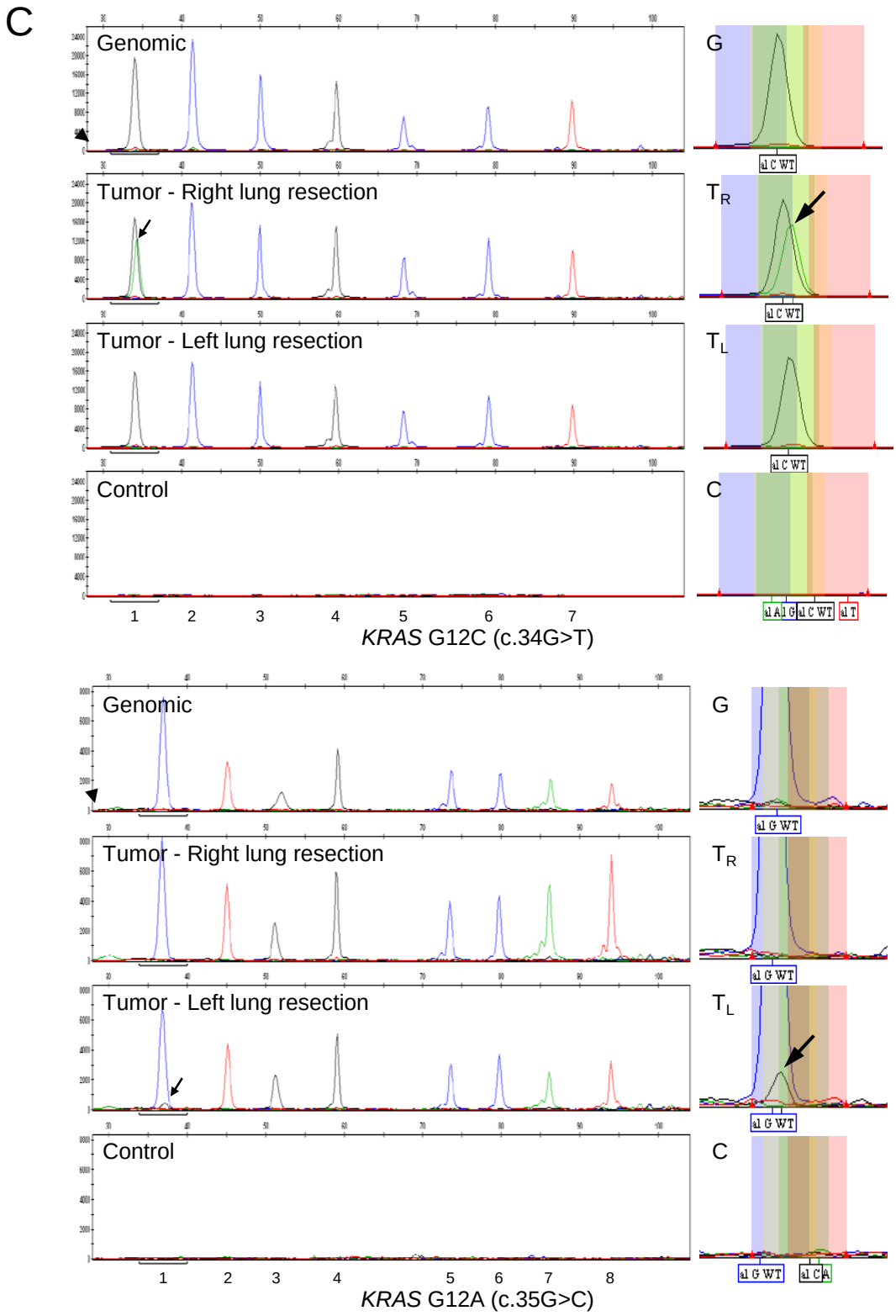
Supplementary Figure 7. Classes of mutations found in primary tumors (A) Across tumor types and **(B)** Correlation with smoking history.



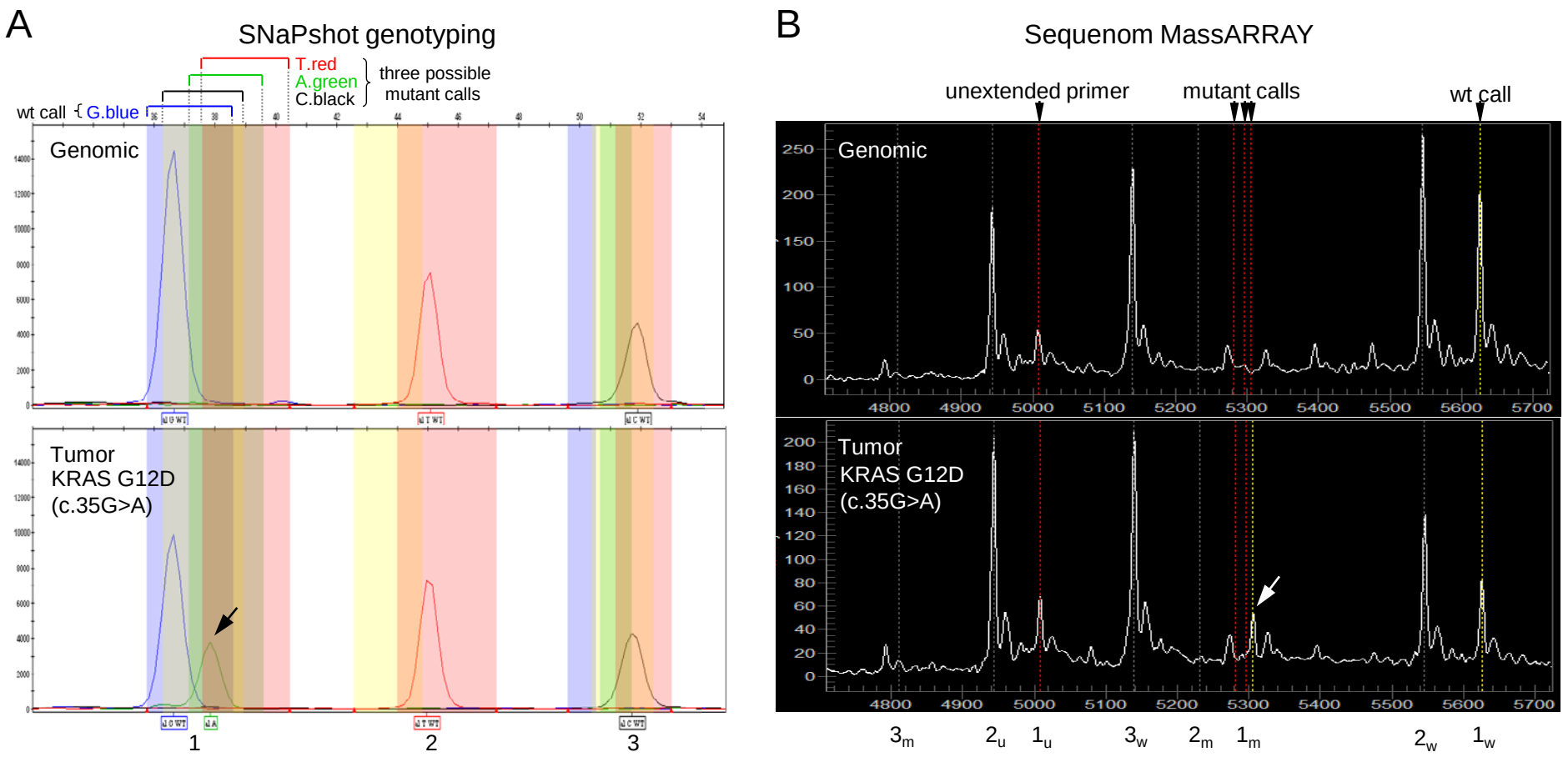




Supplementary Figure 8. Targeted mutational profiling impacts clinical management. Genomic DNA or total nucleic acid extracted from normal (top) and tumor (middle) FFPE tissue from the same patient was run in parallel with a no-DNA negative control (bottom). **(A)** Identification of the *PIK3CA* H1047L (c.3140A>T) mutation in breast cancer. Of note, the PIK3CA c.3140A assay was designed in the reverse orientation, thus the observed alleles are T (red) for the wild-type and A (green) for the mutant. Assays: (1) PIK3CA 3140; (2) CTNNB1 101; (3) JAK2 1849; (4) BRAF 1798; (5) NRAS 37; (6) PIK3CA 1636; (7) APC 4348 and (8) APC 3340. **(B)** Detection of three mutations in a case of lung adenocarcinoma: *EGFR* E746_A750del (c.2235_2249del15) and *EGFR* T790M (c.2369C>T) (upper) and *TP53* R175H (c.524G>A) (lower). Upper image assays: (1) KRAS 34; (2) EGFR 2235_49del F; (3) EGFR 2369; (4) NRAS 181; (5) PIK3CA 1633; (6) CTNNB1 94 and (7) CTNNB1 121. Lower image assays: (1) NRAS 34; (2) PTEN 388; (3) CTNNB1 109; (4) PTEN 697; (5) PTEN 800delA; (6) NRAS 183; (7) TP53 524 and (8) TP53 916.



Supplementary Figure 8 (cont'd). Targeted mutational profiling impacts clinical management. (C) Distinct genotypes found in two tumor masses resected from a lung adenocarcinoma patient. Identification of the *KRAS* G12C (c.34G>T) mutation in the right lung resection (upper), and the *KRAS* G12A (c.35G>C) mutation in the left lung resection (lower). Of note, the proportion of tumor vs. normal cells was different in the two specimens (75% of tumor in the righ lung resection and 30-40% of tumor in the left lung resection), which partly explains the distincnt mutant vs. wild-type allele ratios observed in the two samples. Upper image assays: (1) KRAS 34; (2) EGFR 2235_49del F; (3) EGFR 2369; (4) NRAS 181; (5) PIK3CA 1633; (6) CTNNB1 94 and (7) CTNNB1 121. Lower image assays: (1) KRAS 35; (2) EGFR 2236_50del R; (3) PTEN 517; (4) TP53 733; (5) FLT3 2503; (6) PIK3CA 3139; (7) NOTCH1 4724 and (8) NOTCH1 4802.



Supplementary Figure 9. Comparison between SNaPshot and Sequenom MassARRAY genotyping methods. Wild-type genomic DNA (top) and total nucleic acid extracted from an FFPE lung adenocarcinoma specimen harboring the KRAS G12D mutation (bottom) were analyzed using SNaPshot and Sequenom MassARRAY. The arrow marks the mutant allele. Three assays are depicted for each method. **(A)** SNaPshot platform: automatic allele calling is based on a pre-established binning system that incorporates two sources of information: molecular weight (of the extension product) and color (of the fluorescently-labeled di-deoxynocleotide that is added onto each extension probe during the single base extension reaction). Assays: (1) KRAS 35; (2) EGFR 2236_50del R and (3) PTEN 517. **(B)** Sequenom MassARRAY method: allele calling is based on the distinct molecular weights of each extension product. In addition to the wild-type (w) and three potential mutant (m) signals, the spectral output of each Sequenom MassARRAY assay will also include a peak corresponding to the remaining unextended primer (u). Assays: (1) KRAS 35; (2) EGFR 2235_2249del R and (3) EGFR 2236_50del F. In our limited experience comparing the two systems, we found that the Sequenom MassARRAY performs well with FFPE-derived tissue and uses straightforward protocols. In our hands, the baseline background noise for the Sequenom MassARRAY was higher than with SNaPshot (but is is conceivable that the sensitivity could be increased with proper assay optimization, which we did not attempt). We believe that most laboratories would opt for the SNaPshot system mainly to avoid substantial capital expenses. In addition to those considerations, and because we are focusing on somatic mutations, we also preferred the low background of the SNaPshot spectra (note the low baseline in A) and the ability to distinguish allele variants based on their different colors, which we find useful when trying to identify low-level mutations. Of note, to test one sample with the SNaPshot assay presented in this study we use: 8 multiplexed panels, one chemistry and one extension reaction mix. The protocol designed by Sequenom scientists to test the same loci included: 14 multiplexed panels, two chemistries (IPLEX and hME) and four distinct extension reaction mixes, which would have been more labor intensive, more expensive, and would require ~75% more tumor tissue than our current method. While it is conceivable that experimental optimization would have achieved a better Sequenom protocol, we decided not to pursue such option.

	Cancer Genes												
Tumor type	**APC**	**BRAF**	**CTNNB1**	**EGFR**	**FLT3**	**JAK2**	**KIT**	**KRAS**	**NOTCH1**	**NRAS**	**PIK3CA**	**PTEN**	**TP53**
Breast	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	12%	0%	6%
	4%	3%	2%	<1%	0%	0%	0%	5%	2%	1%	25%	5%	55%
CMD	0%	0%	0%	0%	0%	40%	0%	0%	0%	0%	0%	0%	0%
	N/A	N/A	N/A	N/A	N/A	52%	10%	N/A	N/A	N/A	N/A	N/A	N/A
Colorectal	3%	3%	0%	0%	0%	0%	0%	33%	0%	10%	10%	0%	10%
	39%	11%	5%	<1%	0%	0%	1%	32%	2%	3%	14%	13%	42%
Lung	0%	0%	1%	17%	0%	0%	0%	21%	0%	1%	1%	0%	0%
	1%	2%	3%	26%	<1%	0%	0%	17%	1%	1%	3%	9%	64%
Melanoma	0%	45%	0%	0%	0%	0%	0%	0%	0%	18%	0%	0%	0%
	4%	42%	6%	1%	0%	0%	9%	2%	0%	20%	3%	18%	27%
Pancreatic	0%	0%	0%	0%	0%	0%	0%	48%	0%	0%	0%	0%	9%
	13%	3%	23%	<1%	0%	0%	0%	67%	0%	2%	6%	1%	68%
Prostate	0%	0%	10%	0%	0%	0%	0%	5%	0%	0%	5%	0%	0%
	7%	6%	7%	6%	0%	0%	0%	8%	0%	2%	2%	13%	80%

% values: top (our data); bottom (previous reports)

Supplementary Table 3

SAMPLE ID	ESTIMATED % TUMOR CELLS	% MUTANT = MUT*100/(WT+MUT)	MUTATION(S)	RESPONSE TO EGFR TKIs
NA09-261	N/A	30% (CTNNB1 S37F) 22% (EGFR E746_A750)	CTNNB1 S37F (110C>T) EGFR E746_A750 del in frm 15 (2235_49del)	UNKNOWN
NA09-137	*10-20%	7%	EGFR E746_A750 del in frm 15 (2235_49del)	UNKNOWN
NA09-184	60%	17%	EGFR E746_A750 del in frm 15 (2235_49del)	YES
NA09-237	N/A	59%	EGFR E746_A750 del in frm 15 (2235_49del)	YES
NA09-238	*60%	74%	EGFR E746_A750 del in frm 15 (2235_49del)	UNKNOWN
NA09-129	80%	14%	EGFR E746_A750 del in frm 15 (2236_50del)	YES
NA09-155	40%	5%	EGFR E746_A750 del in frm 15 (2236_50del)	UNKNOWN
NA09-165	10-20%	12%	EGFR L858R (2573T>G)	YES
NA09-187	N/A	13%	EGFR L858R (2573T>G)	UNKNOWN
NA09-206	40%	6%	KRAS G12A (35G>C)	NOT APPLICABLE
NA09-127	10-20%	41%	KRAS G12C (34G>T)	NOT APPLICABLE
NA09-150	N/A	35%	KRAS G12C (34G>T)	NOT APPLICABLE
NA09-162	N/A	57%	KRAS G12C (34G>T)	NOT APPLICABLE
NA09-188	30%	26%	KRAS G12C (34G>T)	NOT APPLICABLE
NA09-194	25-30%	33%	KRAS G12C (34G>T)	NOT APPLICABLE
NA09-207	60%	49%	KRAS G12C (34G>T)	NOT APPLICABLE
NA09-234	*10-20%	19%	KRAS G12C (34G>T)	NOT APPLICABLE
NA09-253	70-80%	66%	KRAS G12C (34G>T)	NOT APPLICABLE
NA09-117	30%	12%	KRAS G12D (35G>A)	NOT APPLICABLE
NA09-128	80%	45% (KRAS G12D) 44% (TP53 R248Q)	KRAS G12D (35G>A) TP53 R248Q (743G>A)	NOT APPLICABLE
NA09-135	80%	15%	KRAS G12V (35G>T)	NOT APPLICABLE
NA09-193	N/A	31%	KRAS G12V (35G>T)	NOT APPLICABLE
NA09-220	50%	9%	KRAS G12V (35G>T)	NOT APPLICABLE
NA09-232	20-30%	12%	KRAS G12V (35G>T)	NOT APPLICABLE
NA09-291	N/A	7%	KRAS G13D (38G>A)	NOT APPLICABLE
NA09-164	50%	10%	TP53 R273L (818G>T)	NOT APPLICABLE

N/A: not available

*Extremely limited tumor tissue

Supplementary Table 4

A

Multiplex PCR Primers	Primer Amount[1]
Panel I	
KRAS_ex2_(a1+a2)	2.4 pmol
EGFR_ex19_(a1+a2)	1.2 pmol
EGFR_ex20_(a1+a2)	0.6 pmol
NRAS_ex3_(a1+a2)	1.2 pmol
PIK3CA_ex10_(a1+a2)	0.6 pmol
CTNNB1_ex3_(a1+a2)	1.2 pmol
Panel II	
EGFR_exon19 _(a1+a2)	0.6 pmol
NRAS_exon2_(a1+a2)	0.6 pmol
BRAF_exon15_(a1+a2)	0.6 pmol
NRAS_exon3_(a1+a2)	0.6 pmol
PI3K exon2_(a1+a2)	0.6 pmol
TP53 exon7_(a1+a2)	0.6 pmol
CTTNNB1_exon3_(a1+a2)	1.2 pmol
Panel III	
EGFR_exon19_(a1+a2)	0.6 pmol
NRAS_exon2_(a1+a2)	0.6 pmol
EGFR_exon21_(a1+a2)	0.6 pmol
CTNNB1_exon3_(a1+a2)	0.6 pmol
PIK3CA_exon10_(a1+a2)	0.6 pmol
Panel IV	
KRAS_exon2_(a1+a2)	1.2 pmol
EGFR_exon19_(a1+a2)	0.6 pmol
PTEN_exon6_(a1+a2)	0.6 pmol
TP53_exon7_(a1+a2)	0.6 pmol
FLT3_exon20_(a1+a2)	0.6 pmol
PIK3CA_exon21_(a1+a2)	0.6 pmol
NOTCH1_exon26A_(a1+a2)	0.6 pmol
NOTCH1_exon26B_(a1+a2)	0.6 pmol
Panel V	
CTNNB1_exon3_(a1+a2)	1.2 pmol
KRAS_exon2_(a1+a2)	1.2 pmol
TP53_exon7_(a1+a2)	0.6 pmol
TP53_exon8_(a1+a2)	0.6 pmol
APC_exon16D_(a1+a2)	0.6 pmol
Panel VI	
CTNBB1_exon3_(a1+a2)	0.6 pmol
KRAS_exon2_(a1+a2)	1.2 pmol
EGFR_exon18_(a1+a2)	0.6 pmol
KIT_exon17_(a1+a2)	0.6 pmol
PIK3CA_exon21_(a1+a2)	0.6 pmol
PIK3CA_exon10_(a1+a2)	0.6 pmol
APC_exon16B_(a1+a2)	0.6 pmol
TP53_exon8_(a1+a2)	1.2 pmol
Panel VII	
PIK3CA_exon21_(a1+a2)	1.2 pmol
CTNBB1_exon3_(a1+a2)	0.6 pmol
JAK2_exon14_(a1+a2)	0.6 pmol
BRAF_exon15_(a1+a2)	0.6 pmol

Supplementary Table 6 (part 1)

NRAS_exon2_(a1+a2)	1.2 pmol
PIK3CA_exon10_(a1+a2)	0.6 pmol
APC_exon16C_(a1+a2)	1.2 pmol
APC_exon16A_(a1+a2)	1.2 pmol
Panel VIII	
NRAS_exon2_(a1+a2)	1.2 pmol
PTEN_exon5_(a1+a2)	1.2 pmol
CTNNB1_exon3_(a1+a2)	0.6 pmol
PTEN_exon7_(a1+a2)	1.2 pmol
NRAS_exon3_(a1+a2)	0.6 pmol
TP53_exon5_(a1+a2)	1.2 pmol
TP53_exon8_(a1+a2)	0.6 pmol

[1]The value indicated represents the amount of each forward and reverse primer added per 10 μl of multiplex PCR (X pmol of a1 and X pmol of a2).

B

Extension Primer	Primer Amount
Panel I	
KRAS34_extR	1.00 pmol
EGFR2235_49del#1_extF	0.56 pmol
EGFR2369_extR	1.25 pmol
NRAS181_extF	1.63 pmol
PIK3CA1633_extF	0.15 pmol
CTNNB1_94_extF	0.19 pmol
CTNNB1_121_extR	0.63 pmol
Panel II	
EGFR2235_49del#2_extR	0.10 pmol
NRAS38_extR	1.86 pmol
BRAF1799_extF	0.25 pmol
NRAS182_extF	0.13 pmol
PIK3CA263_extF	0.38 pmol
TP53_742_extF	1.25 pmol
CTNNB1_95_extR	0.44 pmol
CTNNB1_122_extR	0.25 pmol
Panel III	
EGFR2236_50del#1_extF	0.31 pmol
EGFR2573_extF	0.62 pmol
CTNNB1_133_extR	0.31 pmol
PIK3CA1624_extR	1.25 pmol
NRAS35_extF	0.19 pmol
Panel IV	
KRAS35_extF	0.13 pmol
EGFR2236_50del#2_extR	0.13 pmol
PTEN517_extF	0.75 pmol
TP53.733_extR	3.13 pmol
FLT3_2503_extF	0.25 pmol
PIK3CA3139_extR	0.08 pmol
NOTCH1_4724_extR	3.13 pmol
NOTCH1_4802_extR	6.25 pmol

Supplementary Table 6 (part 2)

Panel V	
CTNNB1_110_extF	0.38 pmol
KRAS38_extF	0.44 pmol
CTNNB1_134_extR	0.23 pmol
TP53_743_extF	1.25 pmol
TP53_817_extF	1.25 pmol
APC4666_67insA_extF	2.25 pmol
Panel VI	
CTNNB1_98_extF	0.15 pmol
KRAS37_extF	0.13 pmol
EGFR2155_extF	0.38 pmol
KIT2447_extF	0.96 pmol
PIK3CA3145_extR	1.25 pmol
PIK3CA1637_extF	0.25 pmol
APC4012_extF	1.86 pmol
TP53_818_extF	0.75 pmol
Panel VII	
PIK3CA3140_extR	0.18 pmol
CTNNB1_101_extF	0.15 pmol
JAK2_1849_extF	0.38 pmol
BRAF1798_extF	0.18 pmol
NRAS37_extR	3.00 pmol
PI3K1636_extF	1.00 pmol
APC4348_extF	0.13 pmol
APC3340_extF	0.88 pmol
Panel VIII	
NRAS34_extR	0.50 pmol
PTEN388_extF	2.50 pmol
CTNNB1_109_extF	0.63 pmol
PTEN697_extR	0.63 pmol
PTEN800delA_extF	1.25 pmol
NRAS183_extR	1.25 pmol
TP53_524_extF	2.50 pmol
TP53_916_extR	0.25 pmol

Supplementary Table 6 (part 3)

Oligonucleotide name[1]	Sequence	Amount[2]
A.ctrl_APC4348C>T	GTACTTCTCACTTGGTTTGAGCTGTTTGAGAAAAA	40 pmol
A.ctrl_APC4666_67 insA	AATCAATAGTTTTTTTCTGCCTCTTTCTCTAAAAA	3 pmol
A.ctrl_BRAF1798G>A	GAGATTTCATTGTAGCTAGACCAAAATCACAAAAA	3 pmol
A.ctrl_CTNNB1_94G>A	ATTCCAGAGTTCAGGTAAGACTGTTGCTGCAAAAA	3 pmol
A.ctrl_CTNNB1_94G>C	ATTCCAGAGTGCAGGTAAGACTGTTGCTGCAAAAA	3 pmol
A.ctrl_CTNNB1_94G>T	ATTCCAGAGTACAGGTAAGACTGTTGCTGCAAAAA	3 pmol
S.ctrl_CTNNB1_95A>G	GCAGCAACAGTCTTACCTGGGCTCTGGAATCCATTCTGGT	30 pmol
A.ctrl_CTNNB1_98C>G	ATGGATTCCACAGTCCAGGTAAGACTGTTGAAAAA	3 pmol
A.ctrl_CTNNB1_101G>A	ATGGATTTCAGAGTCCAGGTAAGACTGTTGAAAAA	3 pmol
A.ctrl_CTNNB1_101G>T	ATGGATTACAGAGTCCAGGTAAGACTGTTGAAAAA	10 pmol
A.ctrl_CTNNB1_109T>G	GTGGCACCAGCATGGATTCCAGAGTCCAGGAAAAA	3 pmol
A.ctrl_CTNNB1_110C>A	GTGGCACCATAATGGATTCCAGAGTCCAGGAAAAA	3 pmol
A.ctrl_CTNNB1_110C>G	GTGGCACCACAATGGATTCCAGAGTCCAGGAAAAA	3 pmol
A.ctrl_CTNNB1_110C>T	GTGGCACCAAAATGGATTCCAGAGTCCAGGAAAAA	3 pmol
S.ctrl_CTNNB1_122C>T	AATCCATTCTGGTGCCACTATCACAGCTCCTTCTCTGAGT	30 pmol
S.ctrl_CTNNB1_133T>C	TGCCACTACCACAGCTCCTCCTCTGAGTGGTAAAGGCAAT	3 pmol
A.ctrl_EGFR2155G>T	GCACCGGAGCACAGCACTTTGATCTTTTTGAAAAA	3 pmol
A.ctrl_KIT2447A>T	TTCTTGATGACTCTGGCTAGACCAAAATCAAAAAA	3 pmol
A.ctrl_KRAS35G>C	CCTACGCCAGCAGCTCCAACTACCACAAGTAAAAA	10 pmol
A.ctrl_KRAS37G>T	TCTTGCCTACGCAACCAGCTCCAACTACCAAAAAA	3 pmol
S.ctrl_NOTCH1_4724T>C	GAGGCTGGCGGCCGGCACGCCGGTGGTGGTGGTGCTGATG	1 pmol
A.ctrl_NOTCH1_4802T>C	GAAGACCACGTTGGTGTGCGGCACGCGGCTGAGCTCCCGC	3 pmol
S.ctrl_NRAS34G>A	ACTGGTGGTGGTTGGAGCAAGTGGTGTTGGGAAAAGCGCA	3 pmol
S.ctrl_NRAS35G>C	ACTGGTGGTGGTTGGAGCAGCTGGTGTTGGGAAAAGCGCA	3 pmol
A.ctrl_NRAS35G>C	TGCGCTTTTCCCAACACCAGCTGCTCCAACCACCACCAGT	3 pmol
S.ctrl_NRAS37G>T	GGTGGTGGTTGGAGCAGGTTGTGTTGGGAAAAGCGCACTG	1 pmol
S.ctrl_NRAS38G>A	GGTGGTGGTTGGAGCAGGTGATGTTGGGAAAAGCGCACTG	3 pmol
S.ctrl_NRAS38G>T	GGTGGTGGTTGGAGCAGGTGTTGTTGGGAAAAGCGCACTG	2 pmol
A.ctrl_NRAS182A>C	TACTCTTCTGGTCCAGCTGTATCCAGTATGAAAAA	5 pmol
A.ctrl_NRAS182A>G	TACTCTTCTCGTCCAGCTGTATCCAGTATGAAAAA	3 pmol
S.ctrl_NRAS183A>C	CATACTGGATACAGCTGGACACGAAGAGTACAGTGCCATG	3 pmol
S.ctrl_NRAS183A>T	CATACTGGATACAGCTGGACATGAAGAGTACAGTGCCATG	3 pmol
A.ctrl_PIK3CA1636C>A	TCTTTCTCCTTCTCAGTGATTTCAGAGAGAAAAAA	3 pmol
S.ctrl_PIK3CA3140A>T	AAACAAATGAATGATGCACTTCATGGTGGCTGGACAACAA	3 pmol
S.ctrl_PIK3CA3145G>A	AATGAATGATGCACATCATAGTGGCTGGACAACAAAAATG	10 pmol
A.ctrl_PTEN388C>G	ACACCAGTTCCTCCCTTTCCAGCTTTACAGAAAAA	1 pmol
A.ctrl_PTEN388C>T	ACACCAGTTCATCCCTTTCCAGCTTTACAGAAAAA	3 pmol
S.ctrl_TP53_733G>A	TAACAGTTCCTGCATGGGCAGCATGAACCGGAGGCCCATC	3 pmol
A.ctrl_TP53_817C>T	GCACAAACACACACCTCAAAGCTGTTCCGTAAAAA	3 pmol

[1]S.ctrl primers were designed as the coding strand (sense) for validation of reverse orientation assays. A.ctrl primers were designed as the non-coding strand (antisense) for validation of forward orientation assays.

[2]Amount of mutation control oligonucleotide used for SNaPshot assay validation.

Supplementary Table 7

Sequencing Primer Name	Sequence	Annealing Temp	$MgCl_2$
APC_ex16A_Seq_a1 M13	TGTAAAACGACGGCCAGTGAGCACTGATGATAAACACCTCAA	58°C[2]	40 nmol
APC_ex16A_Seq_a2 M13	CAGGAAACAGCTATGACCATAGGCTGATCCACATGACGTT		
CTNNB1_ex3_Seq_a1	GATTTGATGGAGTTGGACATGG	60 ºC	50 nmol
CTNNB1_ex3_Seq_a2	TGTTCTTGAGTGAAGGACTGA		
EGFR_ex18_Seq_a1 M13	TGTAAAACGACGGCCAGTCTGAGGTGACCCTTGTCTCTG	65 ºC	40 nmol
EGFR_ex18_Seq_a2 M13	CAGGAAACAGCTATGACCTACAGCTTGCAAGGACTCTGG		
EGFR_ex19_Seq_a1 M13	TGTAAAACGACGGCCAGTGGTAACATCCACCCAGATCAC	65 ºC	40 nmol
EGFR_ex19_Seq_a2 M13	CAGGAAACAGCTATGACCTGAGCAGGGTCTAGAGCAGAG		
EGFR_ex20_Seq_a1 M13	TGTAAAACGACGGCCAGTCGAAGCCACACTGACGTGC	65 ºC	40 nmol
EGFR_ex20_Seq_a2 M13	CAGGAAACAGCTATGACCCTCCTTATCTCCCCTCCCCG		
EGFR_ex21_Seq_a1 M13	TGTAAAACGACGGCCAGTTCTTCCCATGATGATCTGTCC	65 ºC	40 nmol
EGFR_ex21_Seq_a2 M13	CAGGAAACAGCTATGACCCCTGGTGTCAGGAAAATGCT		
JAK2_ex12_Seq_a1	GCAGCAAGTATGATGAGCAAGCTTTC	65 ºC	40 nmol
JAK2_ex12_Seq_a2	CAGATGCTCTGAGAAAGGCATTAG		
PIK3CA_ex21_Seq_a1 M13	TGTAAAACGACGGCCAGTCATACATTCGAAAGACCCTAGCC	58 ºC[2]	40 nmol
PIK3CA_ex21_Seq_a2 M13	CAGGAAACAGCTATGACCATGGATTGTGCAATTCCTATGC		
TP53_ex5_ Seq_a1	CTTGTGCCCTGACTTTCAAC	64 ºC	40 nmol
TP53_ex5_ Seq_a2	ACCAGCCCTGTCGTCTCTC		
TP53_ex6_ Seq_a1	AGGCCTCTGATTCCTCACTG	62 ºC	50 nmol
TP53_ex6_ Seq_a2	ACTGACAACCACCCTTAACC		
TP53_ex7_ Seq_a1	TCATCTTGGGCCTGTGTTATC	58 ºC	50 nmol
TP53_ex7_ Seq_a2	GAAATCGGTAAGAGGTGGGC		
TP53_ex8_ Seq_a1	TTTCCTTACTGCCTCTTGCTTC	58 ºC	50 nmol
TP53_ex8_ Seq_a2	GGAAAGGTGATAAAAGTGAATCTG		
M13_Seq_a1	TGTAAAACGACGGCCAGT	N/A	N/A
M13_Seq_a2	CAGGAAACAGCTATGACC		

Supplementary Table 8