doi:10.4149/neo_2021_210907N1276

Development and extensive analytical validation of deep amplicon sequencing for detecting *KRAS* and *NRAS* mutations in metastatic colorectal cancer samples

Rastislav SLAVKOVSKÝ[1,2], Jana STRÁNSKÁ[1,3], Veronika VĚNSKOVÁ[1], Sylwia JANČIK[1], Lucie KOTKOVÁ[1,3], Jarmila ŠIMOVÁ[4], Petr VOJTA[1], Patrik FLODR[5], Marie BARTOUŠKOVÁ[6,7], Hana ŠTUDENTOVÁ[6,7], Marián HAJDÚCH[1,2], Jiří DRÁBEK[1,2,*]

[1]Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic; [2]Laboratory of Experimental Medicine, University Hospital, Olomouc, Czech Republic; [3]Department of Neurology, University Hospital, Olomouc, Czech Republic; [4]CGB Laboratory, Inc., Ostrava, Czech Republic; [5]Department of Clinical and Molecular Pathology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic; [6]Department of Oncology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic; [7]Department of Oncology, University Hospital, Olomouc, Czech Republic

**Correspondence: jiri.drabek@upol.cz*

Received September 7, 2021 / Accepted November 2, 2021

The presence of wild-type RAS alleles, as determined by genotyping codons 12, 13, 59, 61, 117, and 146, is a prerequisite for personalized anti-EGFR treatment of metastatic colorectal cancer (mCRC) patients. Here we describe analytical validation of in-house developed massively parallel sequencing technology (MPS) in comparison to the *in vitro* diagnostics (IVD) certified qPCR method. DNA extracted from FFPE samples from CRC patients (n=703) and reference standards (n=33) were tested for *KRAS* and *NRAS* mutations in 6 codons of exons 2, 3, and 4 using deep amplicon sequencing (DAS) on a MiSeq benchtop sequencer (Illumina). Two different amplicon lengths and two different library preparation methods (long-RAS and short-RAS) were tested in order to evaluate their impact on DAS performance. In parallel, identical tumor DNA was tested by the following IVD assays: therascreen KRAS RGQ PCR Kit (Qiagen), cobas® KRAS Mutation Test (Roche Diagnostics), and SNaPshot assay (Thermo Fisher Scientific). Both DAS assays detected all the mutations present in reference standards and external quality control samples, except for the artificially generated *KRAS* codon 146 mutation. The DAS assays performed sufficient analytical specificity and sensitivity (≥0.95). The use of shorter amplicons prolonged the preparation steps but significantly improved the sequencing success rate of FFPE-derived DNA. RAS mutation frequencies in the Czech CRC patients were similar to previous reports, although rare mutations were also detected. DAS with short amplicons is a good strategy for routine assessment of somatic mutations in low-quality FFPE-derived DNA.

Key words: deep amplicon sequencing, next-generation sequencing, formalin-fixed paraffin-embedded, predictive diagnostics, KRAS, NRAS, colorectal carcinoma

The epidermal growth factor receptor (EGFR) is successfully targeted by drugs such as cetuximab or panitumumab to treat metastatic colorectal carcinoma (mCRC) [1]. Previous studies demonstrated a reverse association between *KRAS* mutations and patients' responses to the anti-EGFR therapy [2]. Since 2013, the European Medicines Agency (EMA) has stipulated testing six codons of *KRAS* and *NRAS* gene (codons 12 and 13 in exon 2, codons 59 and 61 in exon 3, and codons 117 and 146 in exon 4) for an indication of EGFR-targeted biological treatment for patients with mCRC [1]. Although there are other surrogate markers available [3], they have not achieved general acceptance.

Methods of detecting *KRAS* mutations are based on: sequencing (Sanger sequencing, pyrosequencing, and/or massively parallel sequencing) [4–6], restriction fragment length polymorphisms [7], denaturing high-performance liquid chromatography [8], primer hybridization [9], probe hybridization [10], amplicon high resolution melting analysis [11], primer extension assay [12], real-time quantitative PCR [13], and digital PCR [14]. We and others have previously shown that the results of different methods can disagree due to particular analytical parameters [15, 16].

Several works have described the feasibility of somatic mutations detection using MPS in various malignancies, such as non-small cell lung cancer [17], melanoma, and gastrointestinal malignancies [18]. In comparison to hereditary mutation detection, sequencing of somatic mutation requires higher read coverage due to the lower mutated allele propor-

tion in samples. It has been reported that >500× coverage was required to detect 5% mutation allele frequency (MAF) [18], or >100× coverage for 10% MAF detection [19]. In 10% to 30% of cases, sequencing fails due to an insufficient quantity of tissue available in the formalin-fixed paraffin-embedded (FFPE) blocks [20].

In this study, we tested whether modifications in library preparation can help effectively sequence dozens of samples in parallel, all with possible clinically relevant mutations in exons 2, 3, and 4 of *KRAS* and *NRAS*, with a mutant DNA detection limit on the background of wild-type alleles set to 5% or better. As result, we proposed two deep-amplicon sequencing (DAS) approaches, known as long-RAS and short-RAS.

Patients and methods

Samples and DNA extraction. For the validation cohort, we processed 736 FFPE mCRC tissues, including 703 patient samples from routine clinical diagnostics in two ISO15189-accredited laboratories and 33 tissues obtained as part of external (n=20) or inter-laboratory (n=13) quality assessment. Figure 1 shows the sample processing scheme and the various methods used. Beforehand, each sample was analyzed by an experienced pathologist. The tumor area was marked on hematoxylin and eosin-stained slides, and the tumor content percentage was noted. Only samples with 10% or higher tumor content were macro-dissected and genotyped. Our laboratory participated in the Colon External Quality Assessment Scheme (EQA; kras.eqascheme.org), an international blind proficiency test scheme, organized by The European Society of Pathology. Ten paraffin-embedded samples were provided in 2014 for *KRAS* and *NRAS* testing, including 8 real tumor samples (slides) and 2 samples from cell lines (tubes). We also participated in another external quality assessment, the 2014 EQA Colorectal Scheme, organized by the European Molecular Genetics Quality Network (EMQN) and the Centre for Genomic Medicine, St. Mary's Hospital, Manchester. For this scheme, we received another 10 samples.

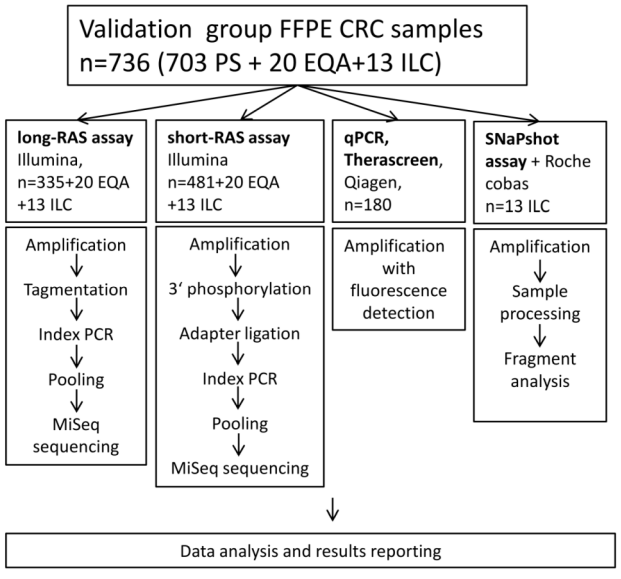


Figure 1. Sample processing scheme used in this study. Abbreviations: PS-patient samples; EQA-external quality assessment; ILC-inter-laboratory comparison. Note: Some of the samples were processed in parallel by multiple methods.

For standards, we used: the HCT116 cell line expressing $KRAS^{G13D/wt}$, the *NRAS* p.(Q61H) and p.(Q61L) FFPE reference standard (50% mutant allele frequency), and the *KRAS* multiplex FFPE reference standard (5%) containing 6 *NRAS* and *KRAS* mutations (all Horizon Diagnostics, Cambridge, UK). DNA was extracted using the cobas® DNA Sample Preparation Kit (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's protocol. For the sake of brevity, hereafter the mutation names are used without brackets (i.e., p.Q61H instead of p.(Q61H)). DNA sample concentrations were measured spectrophotometrically using NanoDrop ND 1000 (NanoDrop Technologies, Wilmington, NC, USA). If a sample was not analyzable by at least one method, or if low PCR amplification was observed (cycle threshold >32 in the majority of amplicons) then it was denoted as "low-quality". The standard DNA input was 1 µl of DNA (>30 ng/µl), or 2.5 µl for samples with low DNA concentration (1–30 ng/µl) for each PCR reaction, corresponding to >30 ng of DNA for standard samples and 2.5–75 ng for low concentration.

PCR design. *KRAS* and *NRAS*-specific primers were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) and the hg19 database (Supplementary Material 1 – Table S1).

Long-RAS assay, amplicon preparation. The long-RAS assay used amplicons ranging in length from 115 bp to 282 bp (average length 224 bp; Supplementary Material 1 – Table S1). All the variants within this paper are associated with transcript sequence and protein sequences as follows: NM_033360.4, NP_203524.1 for *KRAS*; NM_002524.5 and NP_002515.1 for *NRAS*.

Each PCR contained: 2.5 mM $MgCl_2$, 0.25 mM dNTPs, 0.5× Eva Green, primers at 0.2 µM each, and 1 U of Thermo-Start Taq polymerase in 1× buffer (Thermo Fisher Scientific, Waltham, MA, USA). The PCR program was run as follows: denaturation at 95 °C for 5 min, 50 amplification cycles (95 °C for 30 s, 62 °C for 30 s, 72 °C for 30 s), and melting from 60 °C to 95 °C (4 fluorescence acquisitions per second).

Short-RAS assay, amplicon preparation. The amplicons were 70- to 120 bp long (average length 94 bp) for the short-RAS assay (Supplementary Material 1 – Table S1). Each PCR contained: 2.5 mM $MgCl_2$, 0.25 mM dNTPs, 0.5× EvaGreen, primers at 0.2 µM each, and 1 U of Thermo-Start Taq polymerase in 1× buffer. The PCR program was run as follows: denaturation at 95 °C for 5 min, 45 amplification cycles (95 °C for 30 s, 64 °C for 30 s, 72 °C for 30 s), and melting from 60 °C to 95 °C (4 fluorescence acquisitions per second).

Library preparation by tagmentation, long-RAS assay. After amplicon purification (SPRIselect beads, 2:1 beads:amplicons, Beckman Coulter, Krefeld, Germany), the diluted pooled amplicons were tagmented (tagged by adaptor sequence and fragmented using transposase) using the Nextera® DNA XT kit (Illumina, San Diego, California, USA). Following the second round of indexing using a PCR premix from the Nextera® DNA XT kit (PCR, 16 cycles), amplicon purification and size selection were performed using SPRIselect beads.

Library preparation by ligation, short-RAS assay. After purification (QIAquick PCR purification kit, Qiagen, Hilden, Germany: or SPRIselect beads (2:1 beads:amplicons)), the pooled amplicons were diluted and ligated with adapters compatible with Nextera XT index primers (Supplementary Material 1 – Table S1) using T4 PNK and T4 ligase (both New England Biolabs, Ipswich, MA, USA). The products were then purified using SPRIselect beads. Following the second round of indexing using the FastStart High Fidelity PCR System (Roche Diagnostics), EvaGreen 0.5×, and the Nextera XT index kit, purification and size selection were performed using SPRIselect beads. The PCR program consisted of denaturation at 95 °C for 5 min, up to 28 amplification cycles (95 °C for 30 s, 55 °C for 30 s, 72 °C for 40 s). To avoid PCR artifacts, the amplification was stopped manually, usually after 20 cycles, when amplification curves suggested that the samples had reached an exponential phase.

Compared to fragmentation in long-RAS, the library preparation included one more purification step: a supplementary enzyme incubation (3' phosphorylation), and longer enzyme incubation times.

Quality control and sample pooling. For all the sequencing library preparations, concentrations of indexed amplicons were analyzed with the 2100 Bioanalyzer using the Agilent DNA 1000 Kit (Agilent Technologies, Santa Clara, CA, USA), or with the Qubit® 2.0 Fluorometer (Life Technologies Corporation, Carlsbad, CA, USA) using the Qubit® dsDNA HS Assay Kit (Life Technologies Corporation). All samples were normalized according to their concentrations and pooled.

Sequencing and data analysis with MiSeq. The pooled library was denatured with 0.2 M NaOH and diluted to a final concentration of 12 pM, according to the MiSeq System User Guide (Document # 15027617 v01, September). Library sequencing was performed using the MiSeq reagent kit v3, and 150 bp cycles (2×75 bp) sequencing kit (Illumina) with 22M-reads/flow-cell. Analysis was performed using Illumina MiSeq reporter 2.2.29 or above. It included alignment by BWA (Li and Durbin 2009), variant calling and annotation by Illumina Somatic Variant Caller. The output .vcf and .bam files were processed using Microsoft Office Excel VBA Macro (Supplementary Material 2) and IGV 2.3 software [21], respectively. Bedtools v2.25.0 [22] was used for reads coverage calculation. The QC (quality score) threshold was 30, and the detection limit during analysis was set to 1% mutation frequency. For the patient samples, the conclusion for results with at least 5% mutation frequency was "mutation detected", according to the threshold set in the previous publication [23]. For samples with frequency 1–5%, the genotyping was repeated to confirm previous results. At least 2 confirmations of the same result were required to draw a conclusion.

qPCR KRAS mutation tests and RAS SNaPshot assay. The *therascreen* KRAS RGQ PCR Kit (TS; Qiagen) was used to validate mutation detection in *KRAS* exon 2. For inter-laboratory comparison, the analysis of mutations in codons 12, 13 (exon 2), and 61 (exon 3) of the *KRAS* gene was performed using the cobas® KRAS Mutation Test (Roche Diagnostics). Analyses of mutations in codons 59 and 61 (exon 3), 117 and 146 (exon 4) of the *KRAS* gene, and in codons 12 and 13 (exon 2), 59 and 61 (exon 3), and 117 and 146 (exon 4) of the *NRAS* gene were performed using the primer extension method (SNaPshot assay). The assay consisted of PCR followed by a single-base extension reaction and used the commercially available ABI PRISM® SNaPshot™ Multiplex Kit (Life Technologies/Applied Biosystems, Foster City, CA, USA). Details of the method are provided in Supplementary Material 1.

Data collection, statistical analysis, and calculation of mutation frequency in CRC patient population. The mutation status results were collected using Google Sheets and Excel (Microsoft). Statistical analysis was performed and the analytical parameters of the methods were calculated using Microsoft Excel. The mutation frequency in the studied cohort for each *KRAS* and *NRAS* gene codon was calculated as follows: frequency (gene x, codon y) = number of patient samples positive for a mutation in codon y of gene x / total number of patient samples analyzed for anti-EGFR treatment prediction at the Institute of Molecular and Translational Medicine (IMTM), Olomouc.

Results

Performance of DAS RAS assays on reference materials. First, we verified the functionality of MiSeq based tests (long-RAS and short-RAS) for identifying different *KRAS* and *NRAS* mutations by using four commercially available standards. Mutations in 3 samples were correctly identified. In one reference sample containing multiple mutations, we detected 5 of 6 mutations present using the long-RAS assay (Table 1). Despite sufficient coverage of *KRAS* exon 4, we observed false-negative results for the *KRAS* p.A146T mutation. Upon investigation, we found that this was a reference DNA artifact because a molecular scar was artificially introduced into the DNA of the cell line. The sequence, with annotation of the genetic scar, was obtained and is available from the Horizon Discovery Ltd. (Supplementary Material 1 – Figure S0)

Quantification bias of the *KRAS* p.Q61H mutation towards higher values in the long-RAS assay (Table 1) was apparent.

Table 1. Detection of mutations in commercially available reference samples with known genotypes by two DAS methods using MiSeq (Illumina).

***Long*-RAS**						
Sample	**Expected Mutation**	**Expected MF**	**Resulting Mutation**	**No. of Runs**	**Average MF**	**SD MF**
REF1	*KRAS* p.G12D	5%	*KRAS* p.G12D	n=3	4.6%	0.6%
	KRAS p.G13D	5%	*KRAS* p.G13D	n=3	5.5%	1.3%
	KRAS p.Q61H	5%	*KRAS* p.Q61H	n=3	13.2%	1.5%
	KRAS p.A146T	5%	wt	n=3		
	NRAS p.G12V	5%	*NRAS* p.G12V	n=3	7.4%	0.6%
	NRAS p.Q61K	5%	*NRAS* p.Q61K	n=3	6.3%	1.0%
REF2	*NRAS* p.Q61L*	50%	*NRAS* p.Q61L	n=15	46%	3.1%
REF3	*KRAS* p.G13D	50%	*KRAS* p.G13D	n=13	44%	5%
REF4	*NRAS* p.Q61H*	50%	*NRAS* p.Q61H	n=5	82%	5.4%
Short-RAS						
Sample	**Expected Mutation**	**Expected MF**	**Resulting Mutation**	**No. of Runs**	**Average MF**	**SD MF**
REF1	*KRAS* p.G12D	5%	*KRAS* p.G12D	n=3	4.9%	0.7%
	KRAS p.G13D	5%	*KRAS* p.G13D	n=3	6.3%	1.1%
	KRAS p.Q61H	5%	*KRAS* p.Q61H	n=3	3.7%	0.6%
	KRAS p.A146T	5%	*KRAS* p.A146T	n=3	5.2%	0.7%
	NRAS p.G12V	5%	*NRAS* p.G12V	n=3	5.6%	0.8%
	NRAS p.Q61K	5%	*NRAS* p.Q61K	n=3	4.2%	0.6%
REF2	*NRAS* p.Q61L*	50%	*NRAS* p.Q61L	n=2	67%	NA
REF3	*KRAS* p.G13D	50%	*KRAS* p.G13D	n=9	53%	5%
REF4	*NRAS* p.Q61H*	50%	*NRAS* p.Q61H	n=9	86%	3.4%

Note: *Samples REF2 and REF4 were processed with the whole genome amplification procedure. Abbreviations: NA-not applicable; MF-allelic mutation frequency

Conversely, the short-RAS assay, which sequenced shorter amplicons around the predictive mutation hotspots, was able to detect the *KRAS* p.A146T mutation in the abovementioned standard. As shown in Table 1, the short-RAS assay identified all the mutations present and quantified them more precisely. On average, the absolute difference in measured mutation allelic frequency for the short-RAS assay did not exceed 1.3% from the declared value; the declared value was 5% (minimum 3.7%, maximum 6.3%, Table 1). After these promising results, we examined in detail the performance of the newly developed long- and short-RAS DAS assays.

External quality assessment. *RAS* genotyping results of samples from the first EQA study were correct in all cases except sample E9, wherein the genotype was determined as wild-type (wt) instead of *KRAS* p.A146T (Table 2A). After examining the reasons for the disagreement with the study organizers, we found that the reference standard carried the genetic scar. This modification caused primer incompatibility in mutant alleles, as had occurred previously. Thus, only unaffected regions of the *wt* allele were amplified. As expected, the short-RAS sequencing method correctly identified the mutation (Table 2A, sample E9').

The *RAS* genotyping results of samples from the EMQN study were correctly identified in all cases (Table 2B). In the first round of experiments, samples E19 and E20 disagreed with the official EQMN results. However, the results were determined as correct when compared with the TS qPCR method. After communicating with the study organizers, we discovered that we had received the *wt* section of heterogeneous tumor tissue. After making this correction, we concluded that we have reported correct results, from a methodological point of view. Repeated genotyping confirmed the *KRAS wt* genotype (Table 2B, samples E19' and E20').

Inter-laboratory comparison of DAS and SNaPshot assay. Using DAS, we analyzed the DNA samples with known *RAS* mutations identified by the SNaPshot assay and the cobas® *KRAS* Mutation Test. All 13 samples with different mutations were correctly genotyped (Table 2C) by DAS. One sample was initially genotyped as *KRAS* p.A146T by the SNaPshot assay and wt by DAS (Table 2C, sample I6). This was suggesting a false negative result, but retesting by SNaPshot produced the concordant *wt* genotype.

Validation experiments and reliability of DAS methods for routine RAS diagnostics. We compared *KRAS* exon 2 (codons 12 and 13) genotyping at 180 samples with both long-RAS and TS during routine diagnostics and achieved a high concordance between the methods (0.95, kappa value 0.91). The sensitivity and specificity of the long-RAS assay were calculated as 0.981 and 0.982, respectively (Table 3). The overview of non-concordant results is set out in Supplementary Material 1 – Table S0). The TS assay specificity was comparable to that of the long-RAS assay. However, the TS assay sensitivity was lower (0.967) than that of the long-RAS assay. For the long-RAS assay, we observed an approximate failure rate of 9%, due to the poor-quality tumor DNA

Table 2. Side to side comparison of method results.

Sample	DAS Result (*NRAS* and *KRAS* ex. 2, 3 ,4)[$]	MF (short / long-RAS)	*KRAS* TS CE-IVD Result (only *KRAS* codons 12/13 status is shown)	Correct Result of EQMN Study (not taking tumor heterogeneity into account)	Correct Result of EQMN Study (taking tumor heterogeneity into account)	Valid Result [yes/no]
Part A						
E1	*KRAS* p.G12D	25% / 20%	*KRAS* p.G12D	*KRAS* p.G12D		yes
E2	*NRAS* p.Q61K	44% / 46%	n.a.	*NRAS* p.Q61K		yes
E3	wt		wt	wt		yes
E4	wt		wt	wt		yes
E5	*KRAS* p.G13D	35% / 33%	*KRAS* p.G13D	*KRAS* p.G13D		yes
E6	wt		wt	wt		yes
E7	wt		wt	wt		yes
E8	wt		wt	wt		yes
E9	*KRAS* p.A146T/wt	47% / –	n.a.	*KRAS* p.A146T		no*
E10	*KRAS* p.G12C	39% / 34%	*KRAS* p.G12C	*KRAS* p.G12C		yes
Repetition - short-RAS assay is taken as main indicator of result						
E9'	*KRAS* p.A146T	45 % / -	n.a.	*KRAS* p.A146T		yes*
Part B						
E11	*KRAS* p.G12C	13% / 11%	*KRAS* p.G12C	*KRAS* p.G12C 15%	*KRAS* p.G12C 15%	yes
E12	*NRAS* p.Q61K	44% / 41%	n.a.	*NRAS* p.Q61L 50%	*NRAS* p.Q61K 50%	yes
E13	wt		wt	wt	wt	yes
E14	wt		wt	wt	wt	yes
E15	*KRAS* p.G13D	22% / 20%	*KRAS* p.G13D	*KRAS* p.G13D 20%	*KRAS* p.G13D 20%	yes
E16	*KRAS* p.G13D	52% / 44%	*KRAS* p.G13D	*KRAS* p.G13D 50%	*KRAS* p.G13D 50%	yes
E17	*KRAS* p.G12D	18% / 15%	*KRAS* p.G12D	*KRAS* p.G12D 20%	*KRAS* p.G12D 20%	yes
E18	wt		wt	wt	wt	yes
E19	wt		wt	*KRAS* p.G12D	wt	yes
E20	wt		wt	*KRAS* p.G12D	wt	yes
Repetition for confirmation of results						
E19'	wt		wt	*KRAS* p.G12D	wt	yes
E20'	wt		wt	*KRAS* p.G12D	wt	yes
Part C						
				Result CGB Lab Ostrava (SNaP Shot Assay/Cobas)		
I1	*KRAS* p.G12A	62% / 55%		*KRAS* p.G12A		yes
I2	*KRAS* p.G13C	22% / 21%		*KRAS* p.G13C		yes
I3	*KRAS* p.Q61H	13% / 22%		*KRAS* p.Q61H		yes
I4	*KRAS* p.K117N	30% / 32%		*KRAS* p.K117N		yes
I5	*KRAS* p.A146T	28% / 33%		*KRAS* p.A146T		yes
I6	wt			*KRAS* p.A146V/wt**		yes
I7	*NRAS* p.G12C	21% / 26%		*NRAS* p.G12C		yes
I8	*NRAS* p.G12D	31% / 32%		*NRAS* p.G12D		yes
I9	*NRAS* p.G12V	44% / 43%		*NRAS* p.G12V		yes
I10	*NRAS* p.G13R	16% / 20%		*NRAS* p.G13R		yes
I11	*NRAS* p.Q61K	31% / 45%		*NRAS* p.Q61K		yes
I12	*NRAS* p.Q61L	19% / 26%		*NRAS* p.Q61L		yes
I13	*KRAS* p.A146V	40% / not tested		*KRAS* p.A146V		yes

Notes: A, B) Detection of mutations in samples provided by two EQA studies. Correct genotypes were provided several weeks after reporting results. *Reasons for discordance explained in the text. C) Inter-laboratory comparison of RAS genotyping. Correct genotypes were known and provided before reporting results. **Sample was initially reported as mutated *KRAS;* however, after repeating the experiment, the genotype was corrected to wt. $ Both DAS methods provided the same results, unless indicated otherwise. The TS test detects only 7 common *KRAS* mutation; detection of other minor mutations is not applicable. Abbreviations: EQA-external quality assessment; NA-not analyzed; MF-allelic mutation frequency.

isolated from the FFPE tissues, which was significantly higher than the 2.2% failure rate of the qPCR-based TS assay. In order to overcome the problem of amplicon sequencing from highly degraded DNA, we modified the long-RAS assay by significant shortening of *RAS* amplicons (short-RAS assay). We also analyzed 42 samples using both short-RAS

Table 3. Analytical parameters of the DAS methods for genotyping *KRAS* exon 2 (codons 12 and 13).

Parameters of DAS Methods	Long-RAS	Short-RAS
Sensitivity	0.98	0.96
Specificity	0.98	1
Concordance with TS qPCR method	0.95	0.98
Failure rate	0.09	0.024
Number of samples	180	42

Notes: The therascreen RGQ PCR kit (TS), long-RAS and short-RAS methods were used to establish consensual results where two methods presented discordant results. Where not all the data were available, the TS PCR was used as the gold standard. Only non-failed analysis results were taken into account for specificity and sensitivity calculations. Samples with mutations not detectable by TS (mutations in *KRAS* exon 3, exon 4, and all *NRAS* mutations) were interpreted for the purposes of this particular analysis as *KRAS* exon 2 wt.

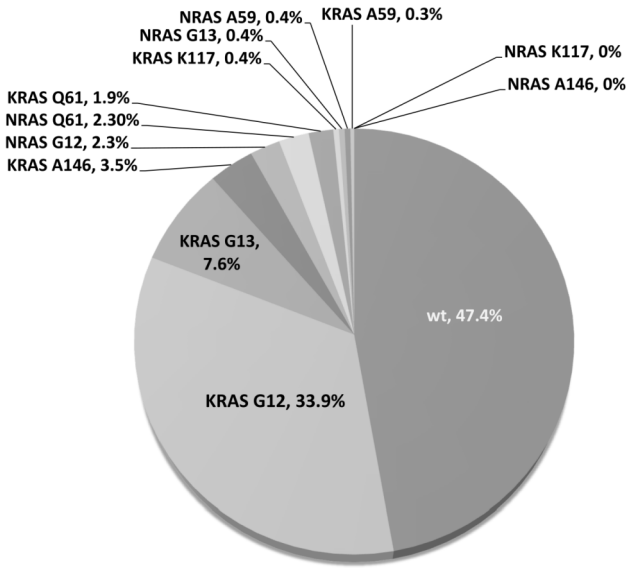


Figure 2. Prevalence of RAS mutations in mCRC. The proportion of samples with RAS mutations of all successfully analyzed samples (n=696) using TS, and/or long-RAS DAS, and/or short-RAS DAS. The total number of samples was 703, of which 7 samples were unanalyzable.

assay and TS. We observed discordance only in one sample, which was false negative by short-RAS. Analytical sensitivity and specificity for this limited set were calculated as 0.96 and 1.0 (Table 3). On a limited cohort of samples where the qPCR result was available, the short-RAS assay failure rate was lower (2.4%, n=42) than the long-RAS failure rate (9%, n=180). The overall short-RAS assay failure rate was 1.5% (7 failed, n=464). The overall long-RAS assay failure rate was 8.4% (29 failed, n=345).

Prevalence of RAS mutations in routine patient samples. Figure 2 gives an overview of the various mutations in tumor samples that were successfully genotyped. As expected, major mutations were present in *KRAS* codons 12 and 13 (cumulatively 39%, Figure 2). Minor mutations in other codons represented about 12% of cases, where the most frequent mutations were found in *KRAS* codons 61 and 146, and *NRAS* codons 12 and 61. We did not detect mutations in *NRAS* codons 117 and 146 (Figure 2).

Comparison of the short-RAS and long-RAS methods. We selected 16 samples (according to the availability of enough well-amplifiable DNA for all methods and the widest available spectrum of mutations) with known mutations to determine the concordance of the short-RAS method in comparison to the long-RAS method (Table 4A). In all 15 FFPE patient samples and 1 cell line sample, both methods produced the same genotyping results, with comparable quantification. The absolute difference in mutation quantification never exceeded 15%. The greatest difference in mutation quantification was observed in sample C2 – *NRAS* p.Q61H (33% for long-RAS vs. 19% for short-RAS).

Furthermore, we analyzed in detail the number of reads per different *RAS* amplicons in both assays for two selected sequencing runs (Figures 3A, 3B). In the short-RAS assay, the average number of reads per amplicon per sample varied from 40 K to 116 K, except for the *NRAS* exon 2 amplicon where we observed an increased number of reads (256 K and 292 K). On the other hand, in the long-RAS assay, a lower read number was observed for the *KRAS* exon 2 amplicon (5 K and 7.5 K) in comparison to other amplicons (84 K to 302 K, Figures 3A, 3B). In this respect, we considered the short-RAS assay as more robust and balanced in amplicon coverage.

We also compared the C_t values of two methods on patient FFPE samples to identify differences in PCR efficiency. There was a negligible difference (~0.5 cycles) in C_t values for intact DNA samples (cell line isolated DNA – sample C16, Table 4A). However, the average C_t value of the short-RAS method was smaller than for the long-RAS method, both for good-quality FFPE samples (~2 cycles, Table 4A) and samples with low-quality DNA that were not analyzable by the long-RAS method (~8 cycles, Table 4B).

Repeatability of replicates. To determine the precision of assays, we compared the mutated allele frequency of two sample replicates processed separately and sequenced in the same sequencing run. We observed a high coefficient of determination (R^2) for the short-RAS assay (0.98, Figure 4A), and also for the long-RAS assay (0.9, Figure 4B). To determine whether sequencing itself had some influence on the precision, we also analyzed the sequencing replicates (split after library preparation) in different sequencing runs. We observed almost identical results with R^2 ~1 (Figure 4C). We concluded that sequencing had a minimal impact on the variability of results, and library preparation is of utmost importance for the precision of the results.

Discussion

We developed and validated two deep sequencing methods for *KRAS* and *NRAS* genotyping, long-RAS and short-RAS, differing in amplicon length and library preparation (tagmentation vs. ligation).

Table 4. Comparison of two DAS methods (long-RAS vs. short-RAS) for detecting RAS mutations.

Sample	Long-RAS		Short-RAS		Long-RAS	Short-RAS	Difference
	Mutation	MF	Mutation	MF	Average Ct	Average Ct*	Delta Ct
Part A							
C1	*KRAS* p.A146T	21%	*KRAS* p.A146T	31%	33.7	29.5	4.2
C3	*KRAS* p.G13D	16%	*KRAS* p.G13D	20%	24.0	24.7	–0.7
C4	*NRAS* p.G12V	25%	*NRAS* p.G12V	20%	28.8	26.9	1.9
C5	*KRAS* p.K117N	46%	*KRAS* p.K117N	46%	28.5	26.5	1.9
C6	*KRAS* p.G12C	37%	*KRAS* p.G12C	44%	26.0	26	0
C7	*NRAS* p.Q61K	34%	*NRAS* p.Q61K	23%	28.3	26.9	1.4
C8	*KRAS* p.A146T	18%	*KRAS* p.A146T	22%	32.3	30.3	2
C9	*NRAS* p.G12S	42%	*NRAS* p.G12S	33%	30.7	27.5	3.2
C10	wt		wt		24.7	23.5	1.2
C11	*KRAS* p.G12V	19%	*KRAS* p.G12V	30%	28.5	24.7	3.8
C12	wt		wt		23.2	22.1	1.1
C13	*KRAS* p.G12V	3%	*KRAS* p.G12V	5%	33.0	29.4	3.6
C14	wt		wt		25.7	24	1.7
C15	wt		wt		28.7	26.1	2.6
C16	*KRAS* p.G13D	43%	*KRAS* p.G13D	52%	23.1	23.5	–0.4
						Average delta Ct	1.8
						SD delta Ct	1.5
						Paired t-test	p=0.0003
Part B							
S1	NA		wt		42.9	34.2	8.7
S2	NA		wt		38.8	35.9	2.9
S3	NA		*NRAS* p.Q61L	13%	38.0	29.7	8.3
S4	NA		*KRAS* p.G12D	36%	43.5	26.6	16.8
S5	NA		*KRAS* p.A146T	12%	41.6	35.4	6.1
S6	NA		*KRAS* p.G13D	43%	39.6	37.2	2.5
S7	NA		*KRAS* p.G13D	11%	37.1	31.8	5.3
S8	NA		wt		33.6	29	4.6
S9	NA		*KRAS* p.Q61H	52%	41.0	30.2	10.8
S10	NA		wt		40.8	28.4	12.4
S11	NA		wt		34.9	30.2	4.7
						Average delta Ct	7.6
						SD delta Ct	4.4
						Paired t-test	p=0.0002

Notes: A) Good-quality samples that were successfully genotyped by both methods. B) Bad-quality samples that were not analyzable by the long-RAS method. *Where the Ct value was not determined for certain amplicons, the value was set to 50 (i.e., total cycles number in qPCR reaction).
Abbreviations: MF-allelic mutation frequency; NA-not analyzable

We compared the ability of two different MPS approaches to detect somatic mutations in CRC, using the Illumina MiSeq platform. The DAS methods correctly genotyped reference standards and external quality control samples, thereby validating its reliability. However, we were unable to detect one of 6 mutations in one of the reference standards using the long-RAS method due to a genetic scar introduced during the engineering of this standard. This demonstrated the need for proper sequence characterization of the reference samples in addition to the codons with introduced mutations. In real patient samples, the presence of such a modification is not expected, unless the mutation is localized in close proximity to a translocation, deletion, or gene amplification.

Using reference standards with 5% mutant alleles, we showed that the short-RAS method provided less deviation from reference values than long-RAS. Our experiment design was unable to distinguish whether this was caused by a difference in the first gene-specific PCR amplification or by library preparation (tagmentation of adapters vs. adapter ligation).

We demonstrated the advantage of the short-RAS method for detecting mutations in low-quality samples. Overall, only 7 of 703 samples (approx. 1%) remained unanalyzable or inconclusive after use of the short-RAS method, compared to the approximate failure rate of 9% with the long-RAS method. This compares favorably with a published failure rate of 20% in a capture-based MPS panel [24], 10% in an

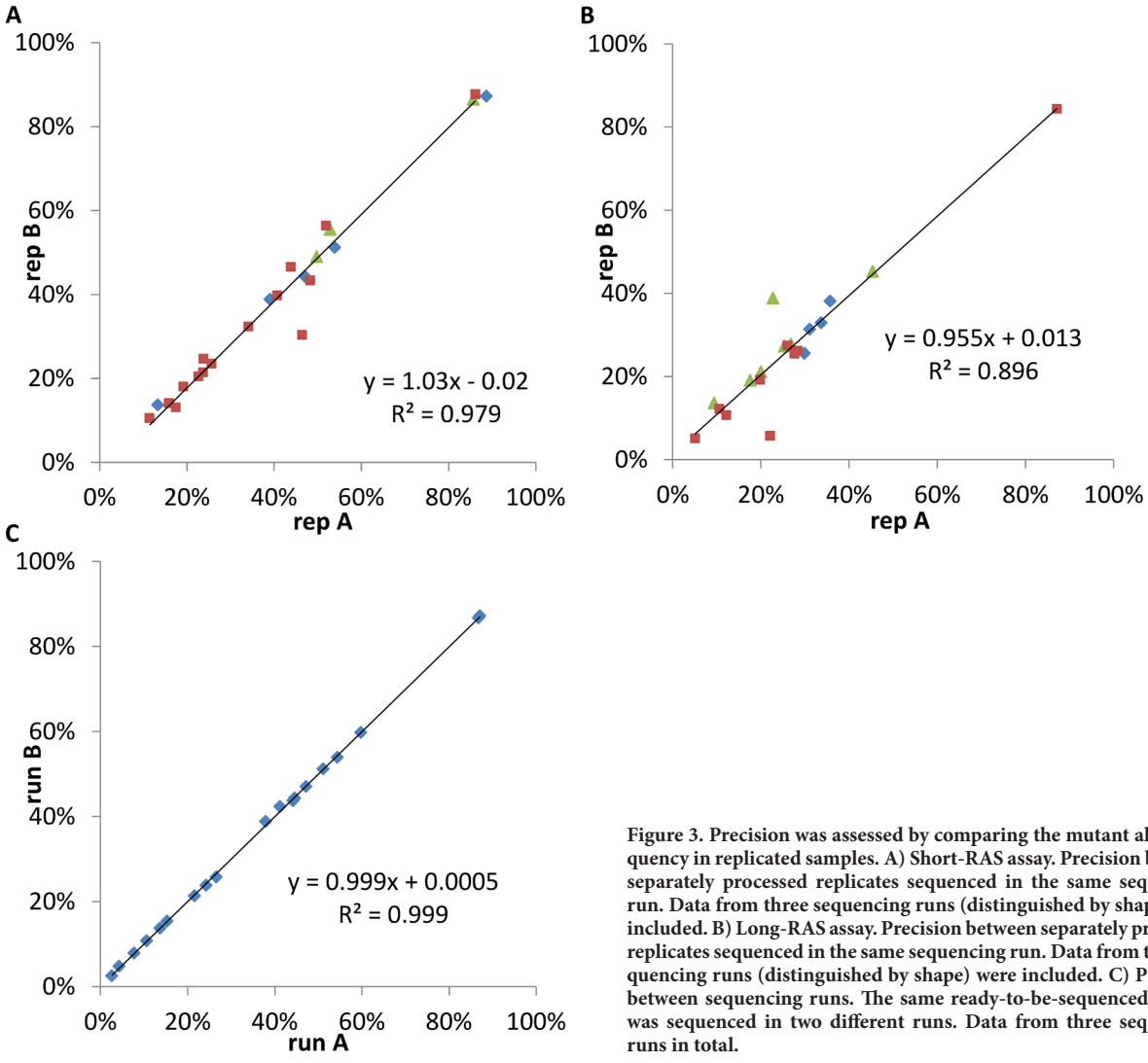


Figure 3. Precision was assessed by comparing the mutant allele frequency in replicated samples. A) Short-RAS assay. Precision between separately processed replicates sequenced in the same sequencing run. Data from three sequencing runs (distinguished by shape) were included. B) Long-RAS assay. Precision between separately processed replicates sequenced in the same sequencing run. Data from three sequencing runs (distinguished by shape) were included. C) Precision between sequencing runs. The same ready-to-be-sequenced sample was sequenced in two different runs. Data from three sequencing runs in total.

amplicon-based TruSeq Custom panel [20], and 5% failure rate in Ion Torrent cancer panel [25]. The main causes of failure were associated with insufficient tissue quantity, insufficient DNA quality, and failed library preparation. For the capture-based MPS panel, more than half of the specimens with highly degraded DNA, as measured by electrophoresis, caused the failures [24]. The culprit for failures in RAS testing is the sample preparation method, as tumor tissue is usually formalin-fixed and paraffin-embedded (FFPE). DNA isolated from FFPE samples is either severely degraded to fragments smaller than 300 bp, crosslinked to proteins, and/or depurinated by formic acid [26]. Though increasing the amounts of polymerase and DNA may sometimes alleviate DNA degradation [26], it is not always feasible due to economic and DNA concentration limitations.

We showed that shortening the amplicons for sequencing could dramatically improve DAS assay performance for low-quality samples. On the other hand, a disadvantage of the short-RAS method is that library preparation is more arduous than the long-RAS method. Nevertheless, further optimization of the library preparation procedure is possible, e.g. by including gene-specific primers already containing MPS adapters [27].

The lower C_t values (higher PCR efficiency) of the short-RAS method in low-quality samples is particularly appealing for difficult-to-analyze FFPE cancer samples. The short-RAS method uses adapter ligation, whereas the long-RAS method uses tagmentation, which further fragments the amplicon. Coverage resulting from ligation is uniform throughout the amplicon length, whereas coverage resulting from tagmentation is low at the amplicon ends, and gradually increases towards the middle of the amplicon. When the hot-spot mutation is not located at the amplicon ends, this phenomenon does not influence the quality of the long-RAS method.

Table 5. Comparison of cost parameters between detection methods, and examples of price calculations for genotyping different numbers of CRC samples using different genotyping platforms.

	TS *KRAS* QIAGEN		DAS Illumina Long-RAS			DAS Illumina Short-RAS	
Detection limit	0.8–6.5%		1% or adjustable			1% or adjustable	
Time consumption	1/2 day		2 days			2 days	
Simplicity of wet lab procedure	*		***			***	
Simplicity of data processing and analysis	*		***			***	
Modifiable, possible to test more genes in one run	No, *KRAS*, 2 codons only		Yes, several			Yes, several	
Maximum tested samples in one run, expected coverage	10 (2 codons)		96 (6 exons, 33 000×) §			96 (6 exons, 28 000×) §§	
No. of samples per run	**12**	**24**	**8**	**12**	**24**	**12**	**24**
Approx. materials costs EUR	1811	3623	1379	1542	2032	1121	1280
Approx. labor costs EUR	37	74	111	111	130	130	148
Total materials and labor costs in EUR	1848	3697	1490	1653	2162	1251	1428
Price per sample (EUR)	**149**	**149**	**180**	**133**	**87**	**101**	**57**
Exon hotspots analyzed	1	1	6	6	6	6	6
Price per exon/sample (EUR)	149	149	30	22	14	17	10

Notes: Only consumable materials costs and approximate labor costs were included in the calculations. No instrumentation or overhead costs were included. Difficulty scale: *simple → ****difficult; §use of 22M pair end (PE) read V3 flow cell, 7 amplicons; §§use of 22M PE read V3 flow cell, 8 amplicons.

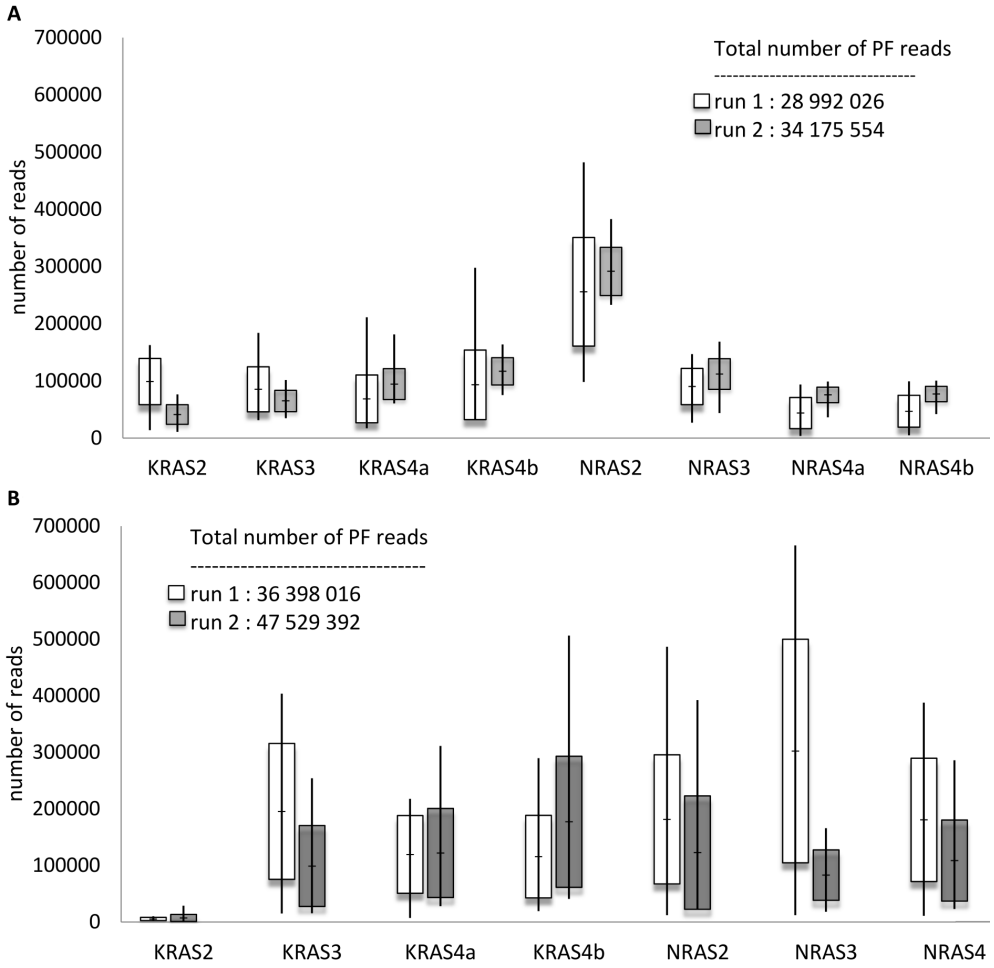


Figure 4. A typical number of reads per sample in different sequenced exon areas in (A) short-RAS assay and (B) long-RAS assay. The bar shows the average value ± standard deviation with maximum and minimum extremities shown by horizontal lines. White bars: run 1, gray bars run 2. A similar number of samples was analyzed within different sequencing runs (n=12–16, two technical replicates for each sample).

However, we observed a shift in the quantification of reference samples for the *KRAS* p.Q61H mutation. The reason for this shift is currently unknown. The coverage of different amplicons was more reliable in the short-RAS assay; this was another reason why this assay was selected as the primary assay for routine diagnostics in our laboratory.

The single nucleotide primer extension assay (SNaPshot) detected multiple mutations in a single reaction, which reduces costs, the amounts of patient sample DNA consumed, and sample handling requirements [28]. We showed that the designed multiplex SNaPshot assay could be used to further validate DAS results. However, targeting other hotspots in codons requires the design of appropriately sized probes, and the multiplexing level is limited to 10 SNPs (https://www.thermofisher.com/order/catalog/product/4323161#/4323161).

Our results confirmed previous findings [29] that MPS demonstrates high analytic sensitivity, and allows simultaneous detection of concomitant mutations and quantitative measurement of mutant allele frequencies. Moreover, we analyzed more than 700 CRC samples with DAS, surpassing the sample sizes of previous studies (310, 100, 114, and 441 samples, respectively) [25, 29–31]. To our knowledge, our study is the third one [32, 33] that comprehensively summarizes mutation frequencies in all possibly mutated clinically relevant predictive codons in *KRAS* and *NRAS* in a large cohort of colorectal tumor samples.

We identified *RAS* mutations in 53% of patients. This percentage is in line with 48% prevalence of *RAS* mutated tumors using an MPS approach in a U.S. population [29], 54% in a Japanese population using the Ion Torrent platform [30], and 56% in a multi-centric study using a combination of Sanger sequencing and MPS [32]. It was shown using targeted pyrosequencing that 49.4 % of tumors bore *KRAS* or *NRAS* mutations within the U.S. population. Thirty-nine percent of our tumor samples contained codon 12/13 *KRAS* mutations, as seen in other studies [32, 34].

We observed 11.5% of samples with minor *RAS* mutation, including codons 59/61/117/146 *KRAS* (6.1%) or *NRAS* (5.4%) mutations. This corresponds to a percentage of patients that would receive anti-EGFR therapy mistakenly after testing using *therascreen* KRAS RGQ PCR Kit. Vaughn et al. reported a 4% mutation frequency in *KRAS* codons 61 and 146 and 2.9% in *NRAS* codons 12, 13, and 61 using pyrosequencing [34]. Schirripa et al. observed 5.8% of samples with mutated *NRAS* (without *NRAS* codon 59) [35]. Peeters et al., using MPS, observed 5% of samples with *NRAS* (codons 12/13/61) and 2.5% with *KRAS* codon 61 mutations [32]. We detected three cases of the very rare mutation *KRAS* p.K117N. This mutation had already been described in colorectal cancer [36], and we estimated that the frequency of the mutation in colorectal cancer is 0.4%.

In addition to confirming previous findings [13, 25, 30], we showed that MPS methods are sufficiently robust with regards to an overall agreement, specificity, and sensitivity for mutations detected in tumor samples when compared to qPCR methods. A limitation of our study is that we used for comparison a qPCR assay that targets only *KRAS* exon 2 codons 12 and 13. The lower analytical sensitivity of qPCR-based *KRAS* genotyping was caused by a sample containing a rare p.G13C mutation that cannot be detected by the design of the particular qPCR assay used in this study.

An advantage of the MPS method over qPCR methods is an easy readjustment of a variable testing panel according to gene and mutation requirements (varied numbers of different genes and exons tested simultaneously, e.g., *PIK3CA* and *BRAF*, can be tested in one sequencing run together with *RAS*). Also, MPS can exploit scarce DNA material with high efficiency, while qPCR methods have limited multiplexing capabilities. Further testing of the target by qPCR requires a new DNA aliquot. The template can be pre-processed by whole genome amplification; however, this may bring unwanted bias [37].

Disadvantages of MPS include greater time consumption, and more complex protocol for sample handling and data analysis compared to qPCR (~2 days for long-RAS/~3 days for short-RAS vs. ~½ day for qPCR, Table 5). Nonetheless, we believe that these disadvantages may be overcome by the use of robotic [38] or microfluidic platforms [39] for the sequencing library preparation.

Table 5 provides approximate cost calculations for various platform options. The cost of MPS depends on the number of samples sequenced in one run. We calculated that at least 11 (long-RAS) or 8 (short-RAS) samples should be sequenced simultaneously to achieve reasonable cost-efficacy, e.g., less than €150 per sample, as it represents the approximate cost per sample for the *therascreen* KRAS RGQ PCR Kit (Table 5). Malapelle et al. reported a higher cost of genotyping materials (€187) when using the Ion Torrent cancer amplicon panel, which also contains RAS pathway genes [25]. This panel provides more information by covering hotspots in 22 genes; nonetheless, information from non-*RAS* genes is not currently required for anti-EGFR treatment, except for the *BRAF* p.V600E mutation, which was recently described as a useful predictive biomarker for combined anti-EGFR, anti-BRAF, and anti-MEK therapy in colorectal cancer [40]. The amplicon of *BRAF* exon 16 can be easily added to both short and long RAS assays. Therefore, in the next generation of our in-house DAS assay, *BRAF* genotyping will be included.

Yet, with an increasing number of samples and more genes and/or exons included in one sequencing run, the cost per sample and per gene decreases (Table 5). This represents a significant advantage, especially when using MiSeq (Illumina). The maximum amounts of tested samples and theoretical coverages calculated for a maximum number of samples are shown in Table 5. For MiSeq, this was calculated based on a flow-cell read capacity of 22 million.

The typical average read coverage for whole-genome sequencing is 30× to 50× and 100× for *whole-exome sequencing*, to expect 15× to 50× coverage of heterozygous

variant present at 50% allele frequency. Extrapolating this for deep amplicon sequencing, one would typically require 5,000× coverage to achieve "safe" coverage of 50 reads of mutated allele present at 1% frequency. From the clinical point of view, a reporting threshold of 5% mutation is a more realistic limit for reporting the presence of clinically relevant mutations, and this threshold was used for evaluating routine samples [23, 24]. As MPS allows an adjustable reporting threshold, in our study, the reporting threshold was deliberately chosen to be 5% with recommended coverage >1,000×. We were able to detect mutations below this threshold; however, these findings require repetition to confirm their validity. We have shown that the number of reads within the region of interest reached as much as one hundred thousand or more. This might seem like a waste of sequencing capacity. However, the abundant sequencing increases the robustness of the assay for poor-quality samples with poor amplification. Also, if required, the capacity could be assigned to another sequencing library, to achieve better overall cost-efficiency.

A limitation of this study is the fact that we did not use unique molecular (UMI) barcodes for labeling amplicons and suppressing PCR errors. This approach may further improve the detection limit of our MPS method to allow mutation testing from blood plasma, where a direct competitor would not be TheraScreen but BEAMing digital PCR technology [41]. Although it was not tested in this work, it is possible to add the UMI barcodes by adding several (8–12) random bases to the adapter that is ligated to DNA. However, adding UMI significantly increases the complexity of the bioinformatics pipeline that needs to be put in place. Our approach used simple but reliable tools available preinstalled on MiSeq sequencers.

To make the diagnostics procedure simple, in our work we did not normalize the mass input of DNA (nanograms) amount, nor did we study whether the higher input can increase the test reliability. Nowadays, it is clear that the DNA input dictates the test sensitivity, rather than just sequencing depth. At least 100 ng DNA (~33,000 genome copies) is required to achieve the LOD of 0.1%, where variant is read ~30×.

In summary, DAS is a technically competent alternative method to qPCR that can produce precise and accurate results for full-scale determination of somatic mutations involved in resistance to anti-EGFR targeted therapy. The cost of MPS is comparable to commercially available methods used in diagnostics. The duration of the method processing and data analysis is longer than qPCR methods, but within an acceptable range for clinicians.

Supplementary information is available in the online version of the paper.

Acknowledgments: Many thanks to Veronika Holinkova for technical assistance and STAR-group for proofreading. This research was supported by the Czech Ministry of Health Grant no. NV16-32198A. The research was approved by the ethical committee of Palacky University Olomouc and University Hospital Olomouc under number 85/15.

References

[1] RICHMAN SD, FAIRLEY J, BUTLER R, DEANS ZC. How close are we to standardised extended RAS gene mutation testing? The UK NEQAS evaluation. J Clin Pathol 2017; 70: 58–62. https://doi.org/10.1136/jclinpath-2016-203822

[2] JONKER DJ, O'CALLAGHAN CJ, KARAPETIS CS, ZALCBERG JR, TU DS et al. Cetuximab for the treatment of colorectal cancer. N Engl J Med 2007; 357: 2040–2048. https://doi.org/10.1056/NEJMoa071834

[3] KISS I, MLCOCHOVA J, BORTLICEK Z, POPRACH A, DRABEK J et al. Efficacy and toxicity of panitumumab after progression on cetuximab and predictive value of MiR-31-5p in metastatic wild-type KRAS colorectal cancer patients. Anticancer Res 2016; 36: 4955–4959. https://doi.org/10.21873/anticanres.11063

[4] ALTIMARI A, DE BIASE D, DE MAGLIO G, GRUPPIONI E, CAPIZZI E et al. 454 next generation-sequencing outperforms allele-specific PCR, Sanger sequencing, and pyrosequencing for routine KRAS mutation analysis of formalin-fixed, paraffin-embedded samples. Onco Targets Ther 2013; 6: 1057–1064. https://doi.org/10.2147/OTT.S42369

[5] TUONONEN K, MAKI-NEVALA S, SARHADI VK, WIRTANEN A, RONTY M et al. Comparison of Targeted Next-Generation Sequencing (NGS) and Real-Time PCR in the detection of EGFR, KRAS, and BRAF mutations on formalin-fixed, paraffin-embedded tumor material of Non-Small Cell Lung Carcinoma-superiority of NGS. Genes Chromosomes Cancer 2013; 52: 503–511. https://doi.org/10.1002/gcc.22047

[6] WANG JF, YANG HY, SHEN YC, WANG S, LIN DM et al. Direct sequencing is a reliable assay with good clinical applicability for KRAS mutation testing in colorectal cancer. Cancer Biomark 2013; 13: 89–97. https://doi.org/10.3233/cbm-130334

[7] ZHANG HD, SONG J, REN H, XU ZR, WANG XN et al. Detection of Low-Abundance KRAS Mutations in Colorectal Cancer Using Microfluidic Capillary Electrophoresis-Based Restriction Fragment Length Polymorphism Method with Optimized Assay Conditions. PLoS One 2013; 8: e54510. https://doi.org/10.1371/journal.pone.0054510

[8] KARIM B, FLORENCE C, KAMEL R, NADIA K, INES O et al. KRAS mutation detection in Tunisian sporadic colorectal cancer patients with direct sequencing, high resolution melting and denaturating high performance liquid chromatography. Cancer Biomark 2010; 8: 331–40. https://doi.org10.3233/cbm-2011-0222

[9] HAMFJORD J, STANGELAND AM, SKREDE ML, TVEIT KM, IKDAHL T et al. Wobble-enhanced ARMS Method for Detection of KRAS and BRAF Mutations. Diagn Mol Pathol 2011; 20: 158–165. https://doi.org/10.1097/PDM.0b013e31820b49e2

[10] BUXHOFER-AUSCH V, AUSCH C, ZEILLINGER R, OBERKANINS C, DANDACHI N et al. Duplex reverse-hybridization assay for the simultaneous detection of KRAS/BRAF mutations in FFPE-extracted genomic DNA from colorectal cancer specimens. Dis Markers 2013; 34: 171–177. https://doi.org/10.3233/DMA-120960

[11] GUEDES JG, VEIGA I, ROCHA P, PINTO P, PINTO C et al. High resolution melting analysis of KRAS, BRAF and PIK-3CA in KRAS exon 2 wild-type metastatic colorectal cancer. BMC Cancer 2013; 13: 169. https://doi.org/10.1186/1471-2407-13-169

[12] LURKIN I, STOEHR R, HURST CD, VAN TILBORG AAG, KNOWLES MA et al. Two Multiplex Assays That Simultaneously Identify 22 Possible Mutation Sites in the KRAS, BRAF, NRAS and PIK3CA Genes. PLoS One 2010; 5: e8802. https://doi.org/10.1371/journal.pone.0008802

[13] HINRICHS JWJ, VAN BLOKLAND WTM, MOONS MJ, RADERSMA RD, RADERSMA-VAN LOON JH et al. Comparison of next-generation sequencing and mutation-specific platforms in clinical practice. Am J Clin Pathol 2015; 143: 573–578. https://doi.org/10.1309/AJCP40XETVYAMJPY

[14] ARAN V, HERINGER M, DA MATA PJ, KASUKI L, MIRANDA RL et al. Identification of mutant K-RAS in pituitary macroadenoma. Pituitary 2021; 24: 746–753. https://doi.org/10.1007/s11102-021-01151-6

[15] JANCIK S, DRABEK J, BERKOVCOVA J, XU Y, STANKOVA M et al. A comparison of Direct sequencing, Pyrosequencing, High resolution melting analysis, TheraScreen DxS, and the K-ras StripAssay for detecting KRAS mutations in non small cell lung carcinomas. J Exp Clin Cancer Res 2012; 31: 79. https://doi.org/10.1186/1756-9966-31-79

[16] TACK V, LIGTENBERG AJL, TEMBUYSER L, NORMANNO N, VANDER BORGHT S et al. External quality assessment unravels interlaboratory differences in quality of RAS testing for anti-EGFR therapy in colorectal cancer. Oncologist 2015; 20: 257–262. https://doi.org/10.1634/theoncologist.2014-0382

[17] HAGEMANN IS, DEVARAKONDA S, LOCKWOOD CM, SPENCER DH, GUEBERT K et al. Clinical Next-Generation Sequencing in Patients with Non-Small Cell Lung Cancer. Cancer 2015; 121: 631–639. https://doi.org/10.1002/cncr.29089

[18] FISHER KE, ZHANG LS, WANG J, SMITH GH, NEWMAN S et al. Clinical Validation and Implementation of a Targeted Next-Generation Sequencing Assay to Detect Somatic Variants in Non-Small Cell Lung, Melanoma, and Gastrointestinal Malignancies. J Mol Diagn 2016; 18: 299–315. https://doi.org/10.1016/j.jmoldx.2015.11.006

[19] SIE D, SNIJDERS PJF, MEIJER GA, DOELEMAN MW, VAN MOORSEL MIH et al. Performance of amplicon-based next generation DNA sequencing for diagnostic gene mutation profiling in oncopathology. Cell Oncol 2014; 37: 353–361. https://doi.org/10.1007/s13402-014-0196-2

[20] FRANCZAK C, DUBOUIS L, GILSON P, HUSSON M, ROUYER M et al. Integrated routine workflow using next-generation sequencing and a fully-automated platform for the detection of KRAS, NRAS and BRAF mutations in formalin-fixed paraffin embedded samples with poor DNA quality in patients with colorectal carcinoma. PLoS One 2019; 14: e0212801. https://doi.org/10.1371/journal.pone.0212801

[21] ROBINSON JT, THORVALDSDOTTIR H, WENGER AM, ZEHIR A, MESIROV JP. Variant Review with the Integrative Genomics Viewer. Cancer Res 2017; 77: E31–E4. Httsp://doi.org/10.1158/0008-5472.Can-17-0337

[22] QUINLAN AR, HALL IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010; 26: 841–842. https://doi.org/10.1093/bioinformatics/btq033

[23] VAN CUTSEM E, LENZ HJ, KOHNE CH, HEINEMANN V, TEJPAR S et al. Fluorouracil, leucovorin, and irinotecan plus cetuximab treatment and RAS mutations in colorectal cancer. J Clin Oncol 2015; 33: 692–700. https://doi.org/10.1200/JCO.2014.59.4812

[24] AL-KATEB H, NGUYEN TT, STEGER-MAY K, PFEIFER JD. Identification of major factors associated with failed clinical molecular oncology testing performed by next generation sequencing (NGS). Mol Oncol 2015; 9: 1737–1743. https://doi.org/10.1016/j.molonc.2015.05.004

[25] MALAPELLE U, VIGLIAR E, SGARIGLIA R, BELLEVICINE C, COLAROSSI L et al. Ion Torrent next-generation sequencing for routine identification of clinically relevant mutations in colorectal cancer patients. J Clin Pathol 2015; 68: 64–68. https://doi.org/10.1136/jclinpath-2014-202691

[26] DIETRICH D, UHL B, SAILER V, HOLMES EE, JUNG M et al. Improved PCR performance using template DNA from formalin-fixed and paraffin-embedded tissues by overcoming PCR inhibition. PLoS One 2013; 8: e77771. https://doi.org/10.1371/journal.pone.0077771

[27] HADD AG, HOUGHTON J, CHOUDHARY A, SAH S, CHEN LJ et al. Targeted, high-depth, next-generation sequencing of cancer genes in formalin-fixed, paraffin-embedded and fine-needle aspiration tumor specimens. J Mol Diagn 2013; 15: 234–247. https://doi.org/10.1016/j.jmoldx.2012.11.006

[28] HEIDEMAN DAM, LURKIN I, DOELEMAN M, SMIT EF, VERHEUL HM et al. KRAS and BRAF mutation analysis in routine molecular diagnostics comparison of three testing methods on formalin-fixed, paraffin-embedded tumor-derived DNA. J Mol Diagn 2012; 14: 247–255. https://doi.org/10.1016/j.jmoldx.2012.01.011

[29] HALEY L, TSENG LH, ZHENG G, DUDLEY J, ANDERSON DA et al. Performance characteristics of next-generation sequencing in clinical mutation detection of colorectal cancers. Modern Pathol 2015; 28: 1390–1399. https://doi.org/10.1038/modpathol.2015.86

[30] SAKAI K, TSURUTANI J, YAMANAKA T, YONESHIGE A, ITO A et al. Extended RAS and BRAF mutation analysis using next-generation sequencing. PLoS One 2015; 10: e0121891. https://doi.org/10.1371/journal.pone.0121891

[31] UDAR N, LOFTON-DAY C, DONG J, VAVREK D, JUNG AS et al. Clinical validation of the next-generation sequencing-based Extended RAS Panel assay using metastatic colorectal cancer patient samples from the phase 3 PRIME study. J Cancer Res Clin Oncol 2018; 144: 2001–2010. https://doi.org/10.1007/s00432-018-2688-3

[32] PEETERS M, OLINER KS, PARKER A, SIENA S, VAN CUTSEM E et al. Massively parallel tumor multigene sequencing to evaluate response to panitumumab in a randomized phase III study of metastatic colorectal cancer. Clin Cancer Res 2013; 19: 1902–1912. https://doi.org/10.1158/1078-0432.CCR-12-1913

[33] WANG Y, LIU H, HOU Y, ZHOU X, LIANG L et al. Performance validation of an amplicon-based targeted next-generation sequencing assay and mutation profiling of 648 Chinese colorectal cancer patients. Virchows Arch 2018; 472: 959–968. https://doi.org/10.1007/s00428-018-2359-4

[34] VAUGHN CP, ZOBELL SD, FURTADO LV, BAKER CL, SAMOWITZ WS. Frequency of KRAS, BRAF, and NRAS mutations in colorectal cancer. Genes Chromosomes Cancer 2011; 50: 307–312. https://doi.org/10.1002/gcc.20854

[35] SCHIRRIPA M, CREMOLINI C, LOUPAKIS F, MORVILLO M, BERGAMO F et al. Role of NRAS mutations as prognostic and predictive markers in metastatic colorectal cancer. Int J Cancer 2015; 136: 83–90. https://doi.org/10.1002/ijc.28955

[36] BOLEIJ A, TOPS BBJ, ROMBOUT PDM, DEQUEKER EM, LIGTENBERG MJL et al. RAS testing in metastatic colorectal cancer: excellent reproducibility amongst 17 Dutch pathology centers. Oncotarget 2015; 6: 15681–15689. https://doi.org/10.18632/oncotarget.3804

[37] STRANSKA J, JANCIK S, SLAVKOVSKY R, HOLINKOVA V, RABCANOVA M et al. Whole genome amplification induced bias in the detection of KRAS-mutated cell populations during colorectal carcinoma tissue testing. Electrophoresis 2015; 36: 937–940. https://doi.org/10.1002/elps.201400136

[38] BORGSTROM E, LUNDIN S, LUNDEBERG J. Large scale library generation for high throughput sequencing. PLoS One 2011; 6: e19119. https://doi.org/10.1371/journal.pone.0019119

[39] TAN SJ, PHAN H, GERRY BM, KUHN A, HONG LZ et al. A microfluidic device for preparing next generation DNA sequencing libraries and for automating other laboratory protocols that require one or more column chromatography steps. PLoS One 2013; 8: e64084. https://doi.org/10.1371/journal.pone.0064084

[40] KOPETZ S, GROTHEY A, YAEGER R, VAN CUTSEM E, DESAI J et al. Encorafenib, Binimetinib, and Cetuximab in BRAF V600E-Mutated Colorectal Cancer. N Engl J Med 2019; 381: 1632–1643. https://doi.org/10.1056/NEJMoa1908075

[41] FRANCZAK C, WITZ A, GEOFFROY K, DEMANGE J, ROUYER M et al. Evaluation of KRAS, NRAS and BRAF mutations detection in plasma using an automated system for patients with metastatic colorectal cancer. PLoS One 2020; 15: e0227294. https://doi.org/10.1371/journal.pone.0227294

https://doi.org/10.4149/neo_2021_210907N1276

Development and extensive analytical validation of deep amplicon sequencing for detecting *KRAS* and *NRAS* mutations in metastatic colorectal cancer samples

Rastislav SLAVKOVSKÝ[1,2], Jana STRÁNSKÁ[1,3], Veronika VĚNSKOVÁ[1], Sylwia JANČIK[1], Lucie KOTKOVÁ[1,3], Jarmila ŠIMOVÁ[4], Petr VOJTA[1], Patrik FLODR[5], Marie BARTOUŠKOVÁ[6,7], Hana ŠTUDENTOVÁ[6,7], Marián HAJDÚCH[1,2], Jiří DRÁBEK[1,2,*]

Supplementary Information

Table S0: Examples of discordant or not analysable results as assayed by two methods: long-RAS DAS and TS. Where possible, sample analysis was repeated and/or samples were analysed by the third complementary method (short-RAS DAS). Final conclusion for every sample was made upon agreement in at least two analyses. NA: not analysable. Incorrect results were underlined.

ESM

ID	long-RAS DAS	long-RAS DAS repeated	short-RAS DAS	TS	TS repeated	conclusion	remark, explanation
D1	NA		wt	wt		wt	low quality DNA for long-RAS DAS
D2	NA		*KRAS* p.G12D 28%	*KRAS* p.G12D		*KRAS* p.G12D	low quality DNA for long-RAS DAS
D3	*KRAS* p.G12D 33%	*KRAS* p.G12D 37%		wt	*KRAS* p.G12D	*KRAS* p.G12D	technical/human error in TS procedure
D4	NA			*KRAS* p.G13D	*KRAS* p.G13D	*KRAS* p.G13D	low quality DNA for long-RAS DAS
D5	Wt		*KRAS* p.G13D 20%	*KRAS* p.G13D	*KRAS* p.G13D	*KRAS* p.G13D	technical/human error in long-RAS procedure
D6	Wt		wt	*KRAS* p.G12V (weak signal)		wt	very low content of mutation detected by TS
D8	*KRAS* p.G13S 5%		*KRAS* p.G12S 3%	wt		*KRAS* p.G13S	TS procedure is not able to detect *KRAS* p.G13S mutation
D9	*KRAS* p.G12C 64%, p.G12D 8%, p.G13D 4%	*KRAS* p.G12D 34%	*KRAS* p.G12D 37%	*KRAS* p.G12D		*KRAS* p.G12D	human error in long RAS DAS procedure or contamination with other samples
D10	Wt		wt	*KRAS* p.G12V (weak signal)		wt	very low content of mutation detected by TS
D11	*KRAS* p.G12V 27%		*KRAS* p.G12V 21%	*KRAS* p.G13D (p.G12V well NA)		*KRAS* p.G12V	technical/human error in TS procedure
D12	*NRAS* p.Q61L 13% *KRAS* p.G12D 2 %		*NRAS* p.Q61L 12 %	*KRAS* p.G12V (weak signal)	*KRAS* p.G12V (weak signal)	*NRAS* p.Q61L	very low content of mutation detected by TS
D13	Wt		wt	*KRAS* p.G12C (weak signal, p.G12V well NA)		wt	very low content of mutation detected by TS
D14	Wt		wt	*KRAS* p.G13D (weak signal, p.G12V well NA)		wt	very low content of mutation detected by TS
D15	*KRAS* p.G13C 8%			wt		*KRAS* p.G13C	TS procedure is not able to detect *KRAS* p.G13C mutation
D16	*KRAS* p.G12S 10%	wt		wt		wt	human error in long RAS DAS procedure or contamination with other samples
D17	*KRAS* p.G13D 18%	wt		wt		wt	human error in long RAS DAS procedure or contamination with other samples
D18	NA	NA	NA	wt		impossible to conclude	very low content of DNA of ultra-low quality

Fig. S0: Sequence of the genomic region around exon 4 of *KRAS* containing *KRAS* p.A146T variant in refere samples Horizon Diagnostics. Region contains genetic scar as indicated. Yellow color highlight indicate place of long-RAS assay primers overlapping the scar, green indicated place of short-RAS primers that are targeting region with the scar. The original document with sequence was obtained and is available from Horizon Discovery Ltd, 8100 Cambridge Research Park, Waterbeach, Cambridge, CB25 9TL, United Kingdc Published with the permit of Horizon Discovery.

Reference sequence: Ensembl transcript: KRAS-001 ENST00000311936

tattaacctattaggttatagttttactatacttctaggtatttgatcttttgagagagatacaaggtttctgtt
taaaaaggtaaagaaacaaaataactagtagaagaaggaaggaaaatttggtgtagtggaaactaggaattacat
tgttttctttcagccaaattttatgacaaaagttgtggacaggttttgaaagatatttgtgttactaatgactgt
gctataactttttttttctttcccagAGAACAAATTAAAAGAGTTAAGGACTCTGAAGATGTACCTATGGTCCTAG
TAGGAAATAAATGTGATTTGCCTTCTAGAACAGTAGACACAAAACAGGCTCAGGACTTAGCAAGAAGTTATGGAA
TTCCTTTTATTGAAACATCA **(G>A)** CAAAGACAAGACAGGtaagtaacactgaaataaatacagatctgtttt
ctgcaaaatcataactgttatgtcatttaatatatcagtttttctctcaattatgctatactagga**] [**aataaa
acaatatttagtaaatgtttttgtctcttgagagggcattgcttcttaatccagtgtccatggtactgcttttgg
ctttggtttctttctacattgaaaatttctcttcaattctgagcacatgttaacatttagaattcaagaggtggg
gatttttttttcccatggttacatatatatatatatatatatatatatatatatatatatatatatataaaga

Lowercase blue sequence indicates intron
Uppercase BLACK sequence indicates exon coding sequence
][indicates the site of engineering
(G>A) indicates the A146T mutation

Materials and Methods

Ultra-deep amplicon sequencing

Table S1: Primer and adaptor sequences used in this work. F or FW: forward primer, R or RW: reverse primer.

primer name	sequence 5' to 3'	amplicon size in bp
Adaptor sequences for Illumina compatible with Nextera index primers		
FW-adaptor-i5	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGT ; 5' PHOSPHORYLATED	
RW-adaptor-i5	CTGTCTCTTATACACATCTGACGCTGCCGACGA ; 5' PHOSPHORYLATED	
FW-adaptor-i7	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGT ; 5' PHOSPHORYLATED	
RW-adaptor-i7	CTGTCTCTTATACACATCTCCGAGCCCACGAGAC' ; 5' PHOSPHORYLATED	
Primers for long-RAS PCR reaction		
KRAS.exon 2 F	AGGCCTGCTGAAAATGACTG	115
KRAS.exon 2 R	TGGATCATATTCGTCCACAAAA	
KRAS.exon 3 F	TGGAGAAACCTGTCTCTTGGA	232
KRAS.exon 3 R	TGCATGGCATTAGCAAAGAC	
KRAS.exon4-b F	AGACACAAAACAGGCTCAGGA	229
KRAS.exon4-b R	AAGAAGCAATGCCCTCTCAAG	
KRAS.exon4-a F	AGTTGTGGACAGGTTTTGAAAGA	182
KRAS.exon4-a R	TGCTAAGTCCTGAGCCTGTT	
NRAS.exon2 F	GATGTGGCTCGCCAATTAAC	239
NRAS.exon2 R	GAATATGGGTAAAGATGATCCG	
NRAS.exon3 F	GGTGAAACCTGTTTGTTGGA	272
NRAS.exon3 R	AACCTAAAACCAACTCTTCCCA	
NRAS.exon4 F	CTGTACCCAGCCTAATCTTGTTTT	259
NRAS.exon4 R	CTGATGCAAACTCTTGCACAAAT	
Primers for short-RAS PCR reaction		
s_*KRAS*2_F	AGGCCTGCTGAAAATGACTG	75
s_*KRAS*2_R	ATCGTCAAGGCACTCTTGC	
s_*KRAS*3_F	TGGAGAAACCTGTCTCTTGGA	70
s_*KRAS*3_R	CTGGTCCCTCATTGCACTGTACT	
s_*KRAS*4A_F	CAGGACTTAGCAAGAAGTTATGGA	85
s_*KRAS*4A_R	CTGTATTTATTTCAGTGTTACTTACCTGTC	
s_*KRAS*4B_F	GGACTCTGAAGATGTACCTATGGTC	80
s_*KRAS*4B_R	GAGCCTGTTTTGTGTCTACTGTTCT	
s_*NRAS*2_F	AAATGACTGAGTACAAACTGGTGGT	90
s_*NRAS*2_R	CTACAAAGTGGTTCTGGATTAGCTG	
s_*NRAS*3_F	GTGAAACCTGTTTGTTGGACATACT	110
s_*NRAS*3_R	GATGGCAAATACACAGAGGAAGC	
s_*NRAS*4A_F	ACTCGGATGATGTACCTATGGTG	100
s_*NRAS*4A_R	CCGTAACTCTTGGCCAGTTC	
s_*NRAS*4B_F	TGCCAACAAGGACAGTTGAT	120
s_*NRAS*4B_R	TGCACAAATGCTGAAAGCTG	

SNaPshot Assay

Primer sequences used for PCR amplification of exons 2 and 3 of *NRAS* gene were according to Lurkin et al. 2010 [1]. PCR amplification of exon 4 of *NRAS* gene and exons 3 and 4 of *KRAS* gene were performed in a 25 µL reaction mix containing 1x PCR Master Mix (Thermo Fisher Scientific) with primers designed by tools available through the Primer 3 web interface. The length of PCR products was checked by electrophoresis on 3% agarose gel (primer sequence and PCR product length are in the Table S2). PCR products were treated with ExoI and FastAP™ Thermosensitive Alkaline Phosphatase (Thermo Scientific) at 37°C for 15 minutes followed by 80°C for 15 minutes. These seven SNaPshot primers used also contained an additional poly(dT) tail at their 5' end of the primers allowing their simultaneous detection (Table S3). SNaPshot primers were extended and designed manually according to the ABI PRISM SNaPshot Multiplex Kit protocol recommendations (Thermo Fisher Scientific).

SNaPshot analysis was performed using the Applied Biosystems® SNaPshot Multiplex Kit (Life Technologies) and previously described conditions [1] with slight modifications of the protocol. Optimal conditions for multiplex assays were determined empirically. SNaPshot reactions were performed in two separate reactions (panel one and panel two) in the final volume of 10 µl, containing 2 µl of purified PCR product. Conditions of multiplex single base extensions were following conditions: 30 cycles 96°C for 10 seconds and 60°C for 35 seconds. SNaPshot assay products were then treated with ExoI and FastAP™ Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific) and incubated at 37°C for 15 minutes followed by 80°C for 15 minutes, held at 4°C. Purified SNaPshot assay products were mixed with 8.75 µl of Hi-Di formamide and 0.25 µl GS120LIZ Size ladder (Thermo Fischer Scientific), denatured at 95°C for 5 minutes and analysed by capillary electrophoresis on 3130 Genetic Analyser with a 36 cm long capillaries and POP-7™ polymer (Thermo Fisher Scientific). The raw data from capillary electrophoresis were analysed on GeneMapper v4.1 Software (Life Technologies, Foster City, CA, USA).

Data analysis and interpretation was done according to the manufacturer's instructions panel parameters for automatic data analysis using GeneMapper Software version v4.1 (Life Technologies, Foster City, CA, USA). There are four possible alleles for each analysed genetic locus: the wild-type allele and three other possible variants. The position of these alleles, even the possible ones, in electrophoretogram can be automatically captured by the analysis software (Figure S1, S2). Parameters for each assay were initially established using SNaPshot® Primer Focus Kit data (Life Technologies) according to the manufacturer's recommendations. Interpretation of SNaPshot genotyping results was accomplished by automatic analysis of the raw data using the established panels and settings, followed by visual inspection of the spectra for all loci by at least two users. This assay can detect 5% - 10% of mutant DNA in the wild type background. Our experiments with reference standards (Horizon Diagnostics, Cambridge, UK) harbouring mutations of interest indicated the possibility of identifying at least 5% of mutated alleles in a background of wild-type DNA (unpublished data). The sensitivity of this assay is similar with data published previously [2].

SNaPshot assay references

[1] Lurkin I, Stoehr R, Hurst CD, van Tilborg AAG, Knowles MA, Hartmann A, et al. Two Multiplex Assays That Simultaneously Identify 22 Possible Mutation Sites in the *KRAS, BRAF, NRAS* and *PIK3CA* Genes. PLoS ONE. 2010 Jan;5(1).

[2] Magnin S, Viel E, Baraquin A, Valmary-Degano S, Kantelip B, Pretet JL, et al. A multiplex SNaPshot assay as a rapid method for detecting *KRAS* and *BRAF* mutations in advanced colorectal cancers. Journal of Molecular Diagnostics. 2011;13(5):485-92.

Table S2. Primer sequences for PCR amplification of exons 2, 3, 4 of *KRAS* and *NRAS* gene.

		Primer sequence 5'→ 3'	Amplicon size [bp]
NRAS exon 2	F	GGTGTGAAATGACTGAGTAC	137
	R	GGGCCTCACCTCTATGGTG	
NRAS exon 3	F	GGTGAAACCTGTTTGTTGGA	103
	R	ATACACAGAGGAAGCCTTCG	
NRAS exon 4	F	CAGGCATGAGCCACTGTACC	286
	R	CCAGAGTTAATCAACTGATGCAAA	
KRAS exon 3	F	CCAGACTGTGTTTCTCCCTT	155
	R	CACAAAGAAAGCCCTCCCCA	
KRAS exon 4	F	GACAAAAGTTGTGGACAGGT	328
	R	TAGCATAATTGAGAGAAAAACTG	

Table S3. SNaPshot extension primers for detection of mutations in the *KRAS* and *NRAS* gene by a single-base extension reaction using the commercially available ABI PRISM® SNaPshot™ Multiplex Kit

primer	Gene, exon	Position cDNA	Primer sequence 5'→ 3'	size [bp]	strand	Wild-Type
SNaPshot assay panel one						
1	*NRAS* ex2	*NRAS* c.34	T34 GTGCGCTTTTCCCAACACCAC	55	AntiS	C
2	*NRAS* ex2	*NRAS* c.35	T41 CTGGTGGTGGTTGGAGCAG	60	S	G
3	*NRAS* ex2	*NRAS* c.37	T46 GGTGGTGGTTGGAGCAGGT	65	S	G
4	*NRAS* ex2	*NRAS* c.38	T49 GTCAGTGCGCTTTTCCCAACA	70	AntiS	C
5	*NRAS* ex3	*NRAS* c.180	T54 GGACATACTGGATACAGCTGG	75	S	A
6	*NRAS* ex3	*NRAS* c.181	T58 CTCATGGCACTGTACTCTTCTT	80	AntiS	G
7	*NRAS* ex3	*NRAS* c.182	T63 GACATACTGGATACAGCTGGAC	85	S	A
8	*NRAS* ex3	*NRAS* c.183	T68 CTCTCATGGCACTGTACTCTTC	90	AntiS	T
7A	*NRAS* ex3	*NRAS* c.175	T60 GTTGGACATACTGGATACA	79	S	G
SNaPshot assay panel two						
8A	*NRAS* ex4	*NRAS* c.351	T65 GGTGCTAGTGGGAAACAA	83	S	G
9A	*NRAS* ex4	*NRAS* c.436	T70 GATTCCATTCATTGAAACCTCA	92	S	G
2A	*KRAS* ex3	*KRAS* c.175	T35 GTCTCTTGGATATTCTCGACACA	58	S	G
3A	*KRAS* ex3	*KRAS* c.176	T40 GCACTGTACTCCTCTTGACCT	61	AntiS	G
4A	*KRAS* ex4	*KRAS* c.351	T45 CTATGGTCCTAGTAGGAAATAA	67	S	A
5A	*KRAS* ex4	*KRAS* c.436	T50 GAATTCCTTTTATTGAAACATCA	73	S	G
6A	*KRAS* ex4	*KRAS* c.437	T55 GTTACTTACCTGTCTTGTCTTT	77	AntiS	G

Figure S1. Overview (electrophoretogram) of the SNaPshot primers used an indication of the peak colour in the cDNA position for wild-type sequence of analysed codons 12, 13, 59, and 61 in *NRAS* gene. Orange peaks – size standard: GS120LIZ Size ladder (Life Technologies)

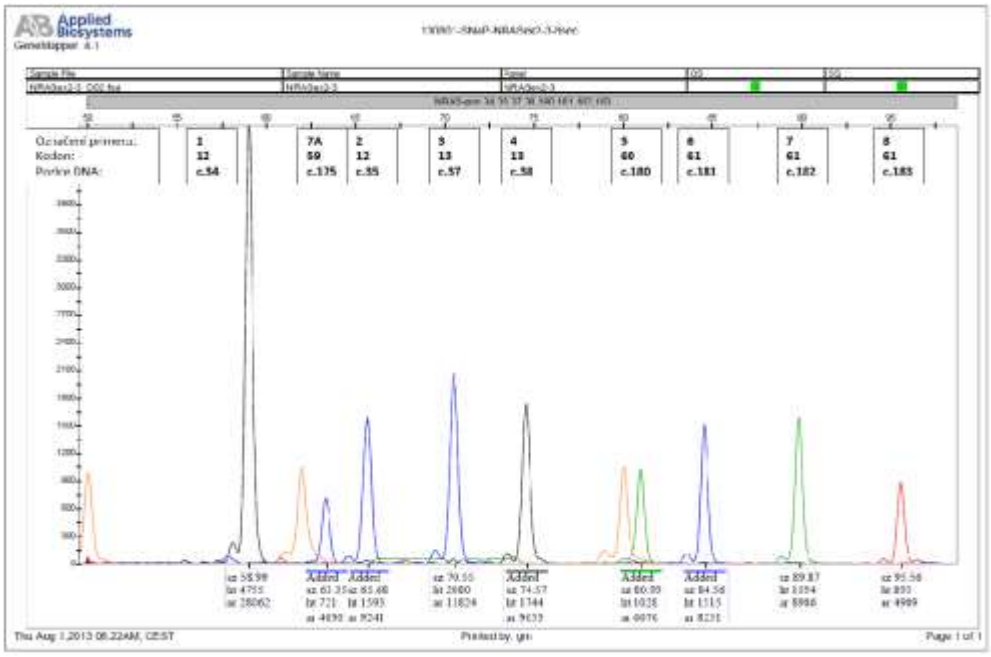
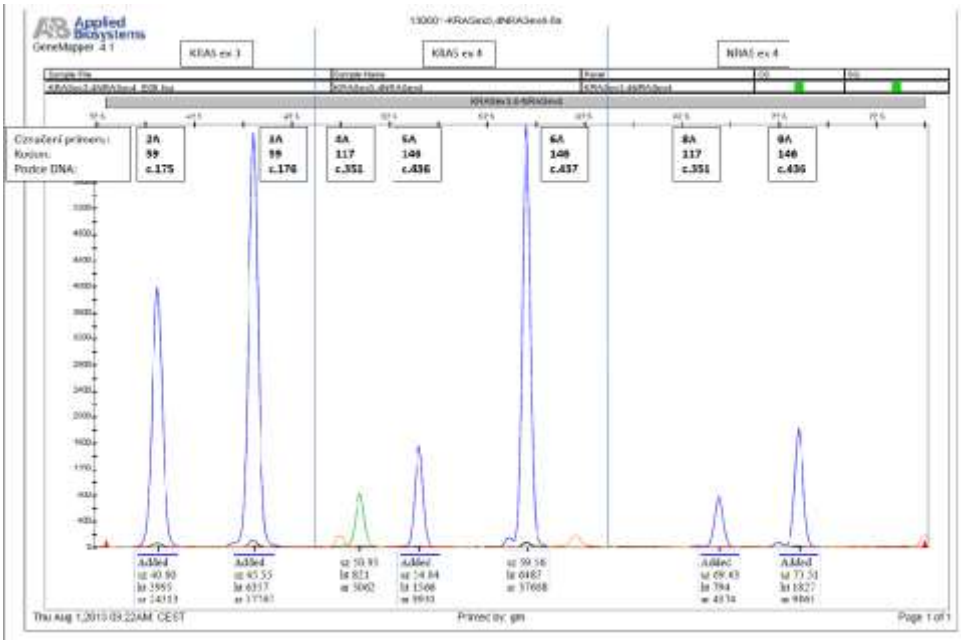


Figure S2. Overview (electrophoretogram) of the SNaPshot primers used an indication of the peak colour in the cDNA position for wild-type sequence of analysed codons 59, 117, and 146 of *KRAS* gene and codons 117 and 146 in *NRAS* gene.



World wide web references

EMEA (European Medicine Agency), a, WWW Accessed on 20th April 2015, Erbitux 5 mg/mL. solution for infusion, summary of product characteristics (SCP, Annex I) http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000558/WC500029119.pdf

EMEA (European Medicine Agency), b, WWW Accessed on 20th April 2015, Vectibix 20 mg/ml concentrate for solution for infusion, summary of product characteristics (SCP, Annex I) http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000741/WC500047710.pdf

FDA (U.S. Food and Drug Administration) web site, Cetuximab (Erbitux) and Panitumumab (Vectibix) Class Labeling Changes to anti-EGFR monoclonal antibodies, cetuximab (Erbitux) and panitumumab (Vectibix): *KRAS* Mutations (Accessed 20/4/2015). http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ucm172905.htm

Horizon diagnostics web site. *KRAS* (A146T/+) Cell line Engineering Summary (Accesed 21/4/2015). http://www.horizondx.com/hd-uploads/sequence_data/*KRAS*%20(A146T).pdf

Illumina web site. Illumina Customer Sequence Letter. (Accessed 21/4/2015). http://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/experiment-design/illumina-customer-sequence-letter.pdf