

TECHNICAL ADVANCE

Open Access



# Comprehensive routine diagnostic screening to identify predictive mutations, gene amplifications, and microsatellite instability in FFPE tumor material

Elisabeth M. P. Steeghs<sup>1†</sup>, Leonie I. Kroeze<sup>1†</sup>, Bastiaan B. J. Tops<sup>1,2</sup>, Leon C. van Kempen<sup>3</sup>, Arja ter Elst<sup>3</sup>, Annemiek W. M. Kastner-van Raaij<sup>1</sup>, Sandra J. B. Hendriks-Cornelissen<sup>1</sup>, Mandy J. W. Hermsen<sup>1</sup>, Erik A. M. Jansen<sup>4</sup>, Petra M. Nederlof<sup>5</sup>, Ed Schuurin<sup>3</sup>, Marjolijn J. L. Ligtenberg<sup>1,4†</sup> and Astrid Eijkelenboom<sup>1\*†</sup>

## Abstract

**Background:** Sensitive and reliable molecular diagnostics is needed to guide therapeutic decisions for cancer patients. Although less material becomes available for testing, genetic markers are rapidly expanding. Simultaneous detection of predictive markers, including mutations, gene amplifications and MSI, will save valuable material, time and costs.

**Methods:** Using a single-molecule molecular inversion probe (smMIP)-based targeted next-generation sequencing (NGS) approach, we developed an NGS panel allowing detection of predictive mutations in 33 genes, gene amplifications of 13 genes and microsatellite instability (MSI) by the evaluation of 55 microsatellite markers. The panel was designed to target all clinically relevant single and multiple nucleotide mutations in routinely available lung cancer, colorectal cancer, melanoma, and gastro-intestinal stromal tumor samples, but is useful for a broader set of tumor types.

**Results:** The smMIP-based NGS panel was successfully validated and cut-off values were established for reliable gene amplification analysis (i.e. relative coverage  $\geq 3$ ) and MSI detection ( $\geq 30\%$  unstable loci). After validation, 728 routine diagnostic tumor samples including a broad range of tumor types were sequenced with sufficient sensitivity (2.4% drop-out), including samples with low DNA input ( $< 10$  ng; 88% successful), low tumor purity (5–10%; 77% successful), and cytological material (90% successful). 75% of these tumor samples showed  $\geq 1$  (likely) pathogenic mutation, including targetable mutations (e.g. *EGFR*, *BRAF*, *MET*, *ERBB2*, *KIT*, *PDGFRA*). Amplifications were observed in 5.5% of the samples, comprising clinically relevant amplifications (e.g. *MET*, *ERBB2*, *FGFR1*). 1.5% of the tumor samples were classified as MSI-high, including both MSI-prone and non-MSI-prone tumors.

(Continued on next page)

\* Correspondence: [Astrid.Eijkelenboom@radboudumc.nl](mailto:Astrid.Eijkelenboom@radboudumc.nl)

<sup>†</sup>Elisabeth M. P. Steeghs, Leonie I. Kroeze, Marjolijn J. L. Ligtenberg and Astrid Eijkelenboom contributed equally to this work.

<sup>1</sup>Department of Pathology, Radboud university medical center, PO Box 9101, 6500, HB, Nijmegen, the Netherlands

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

**Conclusions:** We developed a comprehensive workflow for predictive analysis of diagnostic tumor samples. The smMIP-based NGS analysis was shown suitable for limited amounts of histological and cytological material. As smMIP technology allows easy adaptation of panels, this approach can comply with the rapidly expanding molecular markers.

**Keywords:** Predictive analysis, Next-generation sequencing, Mutation, Gene amplification, Microsatellite instability, FFPE, Melanoma, GIST, Colorectal carcinoma, Lung cancer

## Background

Personalized medicine becomes increasingly important in cancer treatment. Drugs targeting specific mutated genes or specific activated pathways are clinically available for several indications, e.g. for *EGFR*-mutated lung cancer [1]. Currently, many more targeted therapies are being evaluated in clinical trials and show promising results either alone or in combination with other drugs [2]. As molecular markers are constantly expanding, predictive analysis should be easily adaptable to future clinical need. In this era of personalized medicine, next generation sequencing (NGS) analysis using gene panels is increasingly becoming a standard diagnostic approach as interrogation of multiple molecular markers is required with only a limited amount of (mostly) formalin-fixed, paraffin-embedded (FFPE) tissue [3]. These markers include a spectrum of genetic alterations, ranging from small base pair alterations (e.g. point mutations, small deletions, or small insertions) to larger structural variants (e.g. translocations, amplifications, or deletions) affecting genes or large regions of chromosomes. Besides therapy decision, the presence or absence of specific genetic alterations can contribute to the differential diagnosis and provide relevant prognostic and predictive value [4]. In addition, mismatch repair deficiency, causing instability of repetitive DNA sequences known as microsatellites, was shown to be predictive for response to immune checkpoint blockade for a range of tumor types [5].

Sequencing is considered a high resolution approach for the detection of small genetic alterations. With the introduction of sensitive NGS techniques, mutations can reliably be detected in a very low number of cells. We have previously implemented and validated the single molecule Molecular Inversion Probes (smMIP) approach that includes molecular tagging to identify PCR duplicates. These duplicate reads are merged into consensus reads, which leads to the elimination of PCR and sequencing artifacts. In combination with the strand-specific analysis to distinguish genuine mutations from deamination artefacts, it is possible to detect point mutations, small deletions, small insertions, and complex mutations (indels) down to at least 1% mutant allele frequency, as well as to reliably exclude sequence variants with a mutant allele frequency > 3% [6]. In addition, existing gene panels are easily adaptable [6] and thereby this technology can comply with the growing clinical need. Taken together, the smMIPs technology allows screening

of multiple therapeutic targets with high sensitivity using FFPE tissue material in a routine diagnostic setting.

Detection of copy-number variations (CNVs) at specific genomic locations, e.g. amplification of *MET*, is frequently performed by fluorescence in situ hybridization (FISH). In addition, several genome-wide approaches are available, including (shallow) whole genome sequencing or (SNP) array. Interestingly, in the last few years it was shown that NGS-based approaches using defined gene panels are able to identify CNVs in parallel to sequence alterations [7–10]. Recommendations were published to guide the detection and reporting of copy number gains using gene-panel NGS data in a routine diagnostic setting [11].

Recent developments, showing clinical benefit of immune checkpoint blockade in microsatellite instability (MSI) positive tumors, increased the importance of tumor agnostic MSI detection [5]. MSI is characterized by spontaneous gains or losses of nucleotides in small nucleotide repeat regions (microsatellites) and reflects a state of genomic hypermutability. In diagnostic FFPE tissue, the microsatellite status is routinely investigated by a pentaplex PCR investigating five microsatellites [12]. This analysis was mainly validated in Lynch syndrome associated tumors, but little data is available regarding the reliability of this analysis in other tumor types [13, 14]. Studies focusing on the spectrum of unstable microsatellites across multiple cancers show different patterns of MSI among cancer types [15–17]. NGS analyses can be used to study multiple microsatellite loci in one test [13, 17–21].

Taken together, different molecular techniques (i.e. NGS, FISH, and pentaplex PCR) are often used for predictive analyses for cancer therapy. NGS gene panel analysis allows detection of these genomic aberrations in one assay, which will potentially save valuable time, material and costs. Therefore in the current study we aimed to develop a single smMIP-based NGS assay, which can detect sequence alterations, amplifications, and MSI. To comply with the increased number of therapeutic opportunities based on molecular aberrations, both in regular therapeutics and in clinical trial settings, our panel was designed to allow predictive screening for the four most requested molecular-screened indications, i.e. lung cancer, colorectal cancer, melanoma and gastro-intestinal stromal tumors (GIST), but is also useful for evaluation of a broad spectrum of other tumor types as well.

Molecular markers for initial therapy decisions as well as markers important for disease progression (i.e. resistance mechanisms) were included in the panel. In addition, we validated and implemented the detection of clinically relevant gene amplifications in compliance with recently published recommendations [11], and we included the analysis of 55 microsatellites for MSI detection across multiple cancer types.

**Methods**

**Panel design**

The Predictive Analysis for THERapy (PATH) panel should cover all clinically relevant predictive markers for molecular testing of lung cancer, colorectal cancer, melanoma, and GIST. To select predictive markers that have a consequence for targeted treatment-decision-making, clinicians and pharma companies were consulted, as well as the scientific societies (e.g. Dutch Association of Chest Physicians (NVALT), melanoma, and GIST-Consortium), and literature and clinical trial registries were screened for relevant genomic alterations for the specified cancer types. A combination of occurrence, predictive value, and clinical consequences guided the selection of a sequencing panel encompassing 13 kb (Table 1, Supplementary Table 1). Consensus of the panel was obtained with molecular biologists of three specialized hospitals.

**Sample preparation**

Samples were prepared as previously described [6]. In short, gDNA was isolated from FFPE tissue sections (generally 3 × 20 μm) using 5% Chelex-100 and 400 μg proteinase K followed by purification using NaAc and EtOH precipitation. DNA concentrations were measured using the Qubit Broad Range kit (Q32853; Thermo-Fisher, Waltham, MA). Previously, we reported correlations between a low DNA concentration, the number of years between tissue sampling and DNA isolation, and suboptimal smMIP-based NGS results [6]. However, no clear thresholds could be defined. As pre-sequencing analyses are time-consuming, which hampers the turn-around time of the majority of samples, the smMIP-based NGS analysis was performed on all samples without preselection based on DNA quality.

**Preparation of the smMIP pool**

smMIPs for novel targets were designed using the MIP pipeline as described previously [6, 22]. In brief, for mutational analyses preferentially each region had to be covered by two independent smMIPs targeting both strands (double tiling). Hotspots including surrounding regions or complete exons (including splice sites) in 29 genes were selected for mutational analyses. For gene amplifications at least 10 smMIPs per gene were used,

**Table 1** PATH panel

Mutation analysis	Amplification analysis	MSI
AKT1	ALK	55 MSI markers
AKT2	BRAF	
AKT3	EGFR	
ALK	ERBB2	
ARAF	FGFR1	
BRAF	FGFR2	
DDR2	FGFR3	
EGFR	KIT	
ERBB2	KRAS	
FGFR1	MDM2	
FGFR2	MET	
FGFR3	PDGFRA	
GNA11	PIK3CA	
GNAQ		
GNAS		
HRAS		
IDH1		
IDH2		
JAK2		
KIT		
KRAS		
MAP2K1		
MDM2		
MET		
MTOR		
NRAS		
PDGFRA		
PIK3CA		
POLE		
PTEN		
RAF1		
ROS1		
TP53		

which were spread throughout the gene without double tiling. For some genes the smMIPs that were designed for mutation detection fulfilled these criteria, whereas for other genes additional smMIPs were designed (Supplementary Table 1–2). Seven smMIPs for amplification analysis showed inefficient capture and sequencing of targeted regions. Therefore these smMIPs were replaced by newly designed smMIPs on adjusted target regions (Supplementary Table 2). To provide a sex control, smMIPs targeting X-chromosomal AMELX and Y-chromosomal AMELY were added to the panel. During smMIP design the presence of common SNPs (minor

allele frequency > 0.2%) under probe sequences was avoided. If unavoidable, smMIPs were designed to recognize the variant alleles, resulting in 76 additional smMIPs. For MSI detection, smMIPs targeting the five microsatellites of the pentaplex PCR analysis [23] were included, resulting in a total of 554 smMIPs. In a later stage, smMIPs for 50 additional microsatellite regions were designed and added to the sequencing panel (Supplementary Table 3–4).

The designed smMIPs (Integrated DNA Technologies (IDT), Leuven, Belgium) were pooled in an equimolar fashion. The smMIP pool was phosphorylated with 1  $\mu$ L of T4 polynucleotide kinase (M0201; New England Biolabs, Ipswich, MA) per 25  $\mu$ L of 100  $\mu$ mol/L smMIPs in ATP-containing T4 DNA ligase buffer (B0202; New England Biolabs, Ipswich, MA). Based on the obtained coverage for each smMIP, the pool was rebalanced by increasing the concentration of some of the smMIPs. The molecular ratio between gDNA and smMIPs was set to 1:3200 based on 100 ng gDNA input.

#### Automated smMIP library preparation and sequencing

Library preparation was performed as previously described [24]. The final purified libraries were denatured and diluted to a concentration of 1.2 pmol/L. Sequencing was performed on a NextSeq 500 (Illumina, San Diego, CA) according to the manufacturer's instructions (300 cycles Mid-Output or High-Output sequencing kit), resulting in 2  $\times$  150 bp paired-end reads.

#### Variant detection

Data analysis was performed as previously described [6, 24]. The first steps of data conversion were done automatically, i.e. BCL-to-FASTQ conversion and demultiplexing [24]. Demultiplexed FASTQ files were uploaded and analyzed in Sequence pilot version 4.4.0 (JSI medical systems, Ettenheim, Germany) [6], which included removal of *PIK3CA* and *PTEN* pseudogene reads from the alignment and subsequent analysis. After variant calling, all variants were manually inspected and curated. To improve detection of large deletions resulting in skipping of *MET* exon 14, we decreased the percentage of consecutive bases that have to match to the reference without a mismatch (30% instead of 50%) and the minimum total absolute coverage for both directions combined (10 instead of 20). To evaluate detection of large deletions resulting in skipping of *MET* exon 14, three artificial DNA sequences (IDT) were used.

#### Amplification analysis

An external baseline control series was generated from ten normal FFPE tissue samples, which were sequenced in several independent runs. The unique coverage per tumor sample was normalized using the median sequencing

depth of all amplicons in the sample, an approach adapted from Budczies et al [7]. This normalized coverage per gene per sample was divided by the mean coverage of the gene in the control series. The obtained value includes the relative unique coverage (or fold change) and can be divided by two to assess the total number of alleles that are present per genome equivalent. In addition, a significance score, the z-score, was calculated by dividing the difference between the normalized coverage of the sample and the control series by the standard deviation of the control series. To assess the number of alleles that are present in the tumor cells only, the relative coverage was corrected by tumor purity:  $(\text{relative coverage} * 2 - (1 - \text{fraction of neoplastic cells}) * 2) / \text{fraction of neoplastic cells}$  [11].

#### Microsatellite instability detection

mSINGS software was used for the detection of significantly altered read length distribution in sequencing reads covering microsatellite markers, as described by Salipante et al [20]. This tool is able to perform tumor-only MSI analysis without the need of a paired normal sample. In short, baseline reference values were generated from 20 normal FFPE tissue samples for the 55 microsatellite loci. Microsatellite status of unknown samples was assessed by comparing repeat length distribution for each locus to the baseline reference value. Per locus, the total number of alleles with different lengths was assessed. Repeat lengths were included if their read count exceeded 5% compared to the read count of the most frequently observed allele. The number of repeat lengths was compared to the number of the baseline. If the counted repeat lengths exceeded [mean number of alleles + (2  $\times$  SD)] the baseline value, a locus was scored as unstable. Finally, the mSINGS score was assessed by dividing the number of unstable loci per samples by the total number of evaluated loci.

#### OncoScan CNV array

80 ng DNA from FFPE samples was processed according to the manufacturer's instructions (ThermoFisher Cat. No. 902694). Data was visualized and analyzed using Chromosome Analysis Suite (ChAS) software version 3.2 (ThermoFisher).

#### Panel requirements

The requirements of the PATH panel were formulated prior to its validation and implementation, including i) > 90% of selected regions of interest should be double tiled, preferably recognizing both strands for hotspot regions. ii) Read depth distribution throughout the panel should be uniform; the average read depth for > 90% of the targets relevant for sequencing analysis may vary up to 1 order of magnitude. iii) The number of sample drop-outs should be similar or smaller compared to the

smMIPs cancer hotspot panel [6] and concordant results should be obtained. iv) For amplification analysis, 5 positive controls and 10 samples without any amplification should be correctly classified. v) For MSI detection, a series of 10 positive and 20 negative control samples should be correctly annotated.

## Results

### Panel validation – sequencing analysis

The smMIPs were pooled in an equimolar fashion and tested on 4 samples of archived FFPE tumor tissue. To boost coverage of underperforming smMIPs relevant for sequence analysis, the concentration of 55 smMIPs was increased (Supplementary Table 2). In addition, 7 smMIPs were replaced by newly designed smMIPs. 10 samples from archived tissue were sequenced using this rebalanced pool. 91% of relevant regions for mutation detection were covered by  $\geq 2$  smMIPs in both orientations. Unique read depth obtained from sequence analysis of these 10 samples resulted in a median coverage of 506 unique reads per region, representing an equal number of individually sequenced gDNA molecules. On average, unique coverage of 96.3% of relevant regions was within one order of magnitude (Fig. 1a). The coverage of 3 regions was lagging behind, i.e. the regions around *ARAF* codon 214, *ERBB2* exon 18, and *ERBB2* exon 21. These regions are relatively GC-rich (GC% *ARAF* region: 62%, *ERBB2* exon 18: 66%, *ERBB2* exon 21: 59%), which might partly explain why these regions are more difficult to amplify and sequence (Supplementary Table 5), as previously shown [6]. Since these regions are not frequently mutated, we accepted the lower coverage of these regions.

Subsequently, 57 tumor samples (NSCLC, colorectal carcinoma, melanoma, GIST, and a range of tumor types for which clinical trial options were evaluated (hereafter referred to as ‘miscellaneous’) were sequenced using this new pool (Fig. 1b) and compared with our routine diagnostics cancer hotspot panel [6]. Parallel analysis showed a comparable number of drop-outs ( $n = 5$ ) and a comparable coverage for the individual regions (Supplementary Fig. 1). The 5 drop-out samples could be explained by poor DNA quality, i.e. four FFPE samples were stored for  $> 8$  years and for one sample a cytological specimen with partly degraded cells was used. Noteworthy, another five FFPE samples stored for  $> 8$  years performed well. 75% (39/52) of the successfully evaluated cases harbored  $\geq 1$  (potentially) pathogenic mutation (Fig. 1c). 32 mutations were identified by both panels and showed concordant variant allele frequencies (Fig. 1d). The PATH panel identified 31 additional (potentially) pathogenic mutations, which were located in regions that were not present in the cancer hotspot panel (Fig. 1e).

In addition, we evaluated the detection of *MET* exon 14 skipping mutations, which can either be point mutations, deletions, or indel mutations that target the splice donor or acceptor site of exon 14, or the intronic polypyrimidine tract region [25–27]. In the Hs746T cell line, a c.3082 + 1G > T mutation at the splice donor site of *MET* (NM\_001127500.2), predicted to result in *MET* exon 14 skipping, was successfully detected by the PATH panel. The panel design includes smMIPs with large insert size specifically designed to allow amplification of large deletions at this locus. Three artificial DNA sequences containing a 46, 173, or 201 bp deletion affecting the splice acceptor site (based on Frampton et al [25]) were detected, suggesting that large deletions affecting one of the splice sites could be detected by the panel in a diagnostic setting. Indeed, after implementation in routine diagnostics, a splice donor site mutation was identified in *MET* intron 14 (c.3082 + 1G > A) as well as a 66 bp deletion (c.2942-4\_3003del) involving the *MET* exon 14 splice acceptor site. Based on these results and the different *MET* aberrations described [25, 26], we estimated that all point mutations and  $> 90\%$  of all deletions that can result in exon 14 skipping can be identified by gDNA sequencing using the PATH panel.

### Panel validation – amplification analysis

An external baseline control series was generated from ten normal FFPE tissue samples, with stable performance in multiple runs (Supplementary Fig. 2A). The efficacy of amplification analysis was evaluated in five positive control samples, which included a high level amplified *EGFR*, *MET*, and *ERBB2* (HER2) sample, and a median and low level amplified *ERBB2* sample. For sample normalization, median sequencing depth over all amplicons was used, as this is much less affected by high level amplifications compared to the mean and summed unique coverage (Supplementary Fig. 2B). All amplifications in the positive controls were confirmed, as visualized by a high relative coverage and z-score compared to the other genes in the samples (Fig. 2a). In addition, the values for the *ERBB2* amplified samples were proportional to the results obtained by *ERBB2* FISH analysis, yielding a ‘high’, ‘median’ and ‘low’ value. Moreover, an additional *EGFR* amplification was observed in one *ERBB2* positive control, which was confirmed by FISH analysis (not shown).

For reliable detection of amplifications in routine diagnostics, it is highly recommended to establish cut-offs for amplification detection [11]. The relative coverage of 46 samples from routine diagnostics were therefore compared to a control series (Fig. 2b, Supplementary Fig. 3A). Based on visual inspection, a threshold in relative coverage  $\geq 3.0$  was chosen for further validation, which corresponded to a z-score of  $> 6.4$  (Supplementary

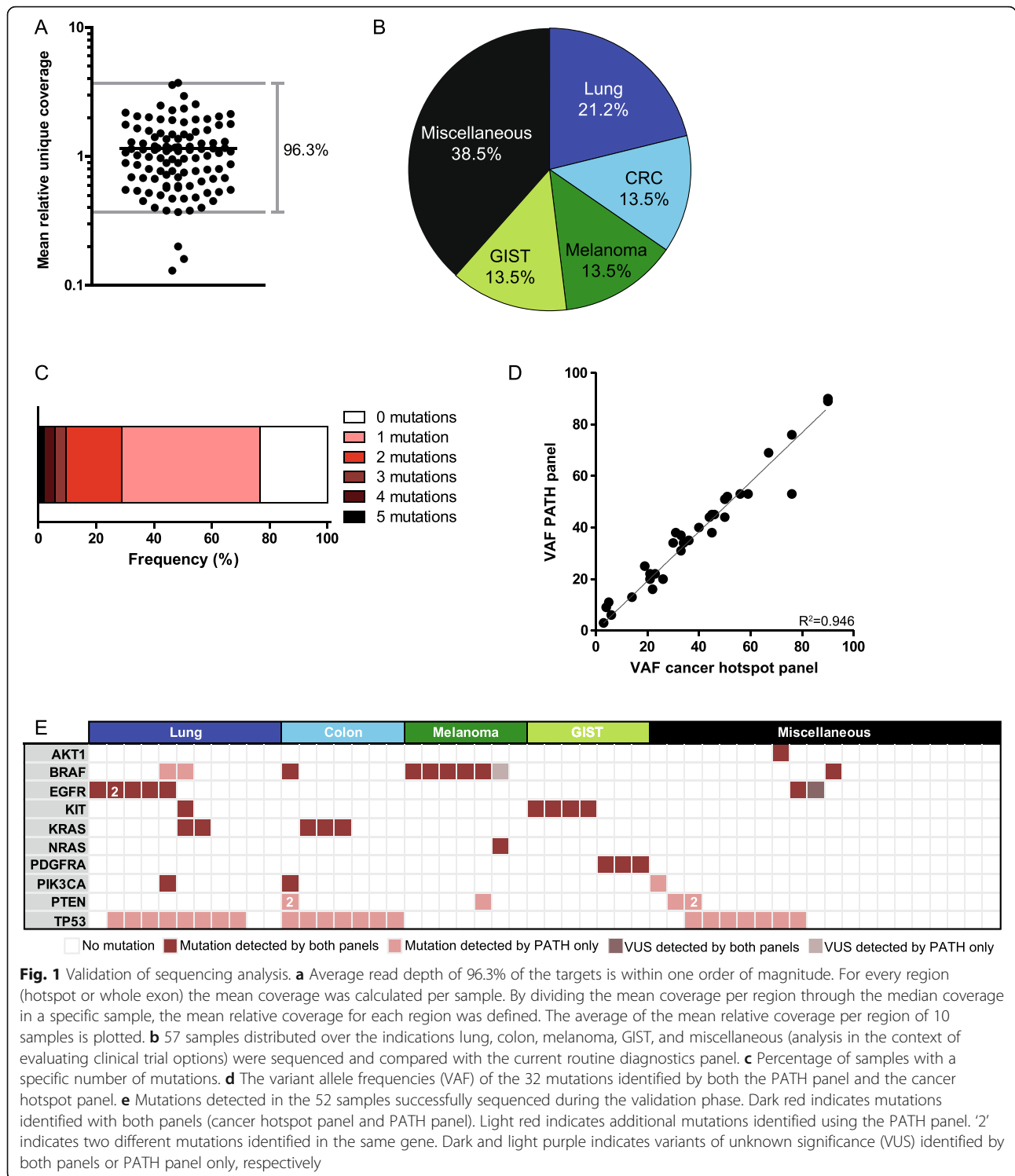
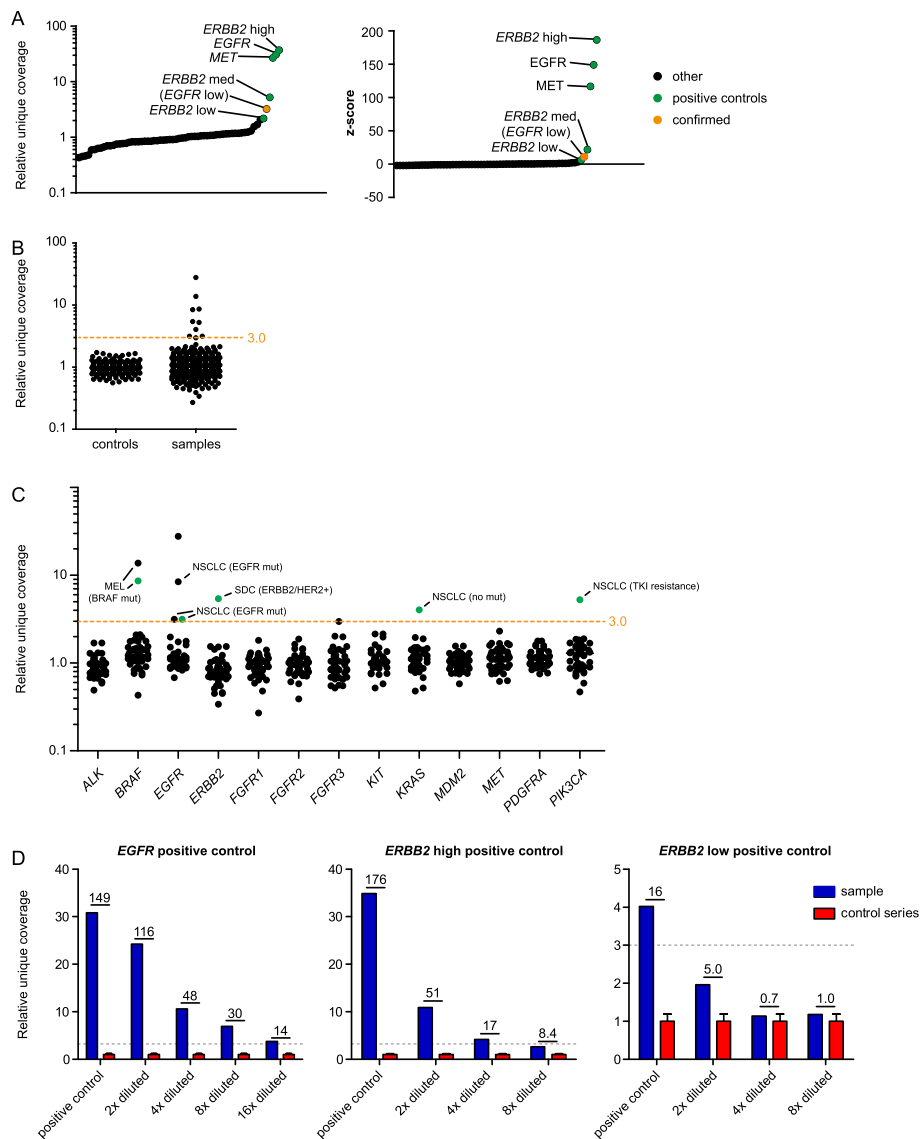


Fig. 3B). In addition, the lower limit of alleles that need to be present in the neoplastic cell component to reach the threshold of relative coverage  $\geq 3.0$  can be deduced based on the tumor purity of a sample (Supplementary Fig. 3C) [11]. Nine of 46 samples exceeded the threshold for one or more genes (Fig. 2c). The observed amplifications complied

with clinical specifications and additional molecular findings [28–33]: we detected an *EGFR* amplification in four samples (i.e. three *EGFR*-mutated lung tumors and one glioblastoma specimen), a *BRAF* amplification in two samples (both *BRAF*-mutated melanoma samples), a *PIK3CA* amplification in one sample (*EGFR*-mutated NSCLC case



**Fig. 2** Validation of detection of amplifications in smMIP-based NGS analysis on gDNA from clinical FFPE specimens. **a** Relative coverage and z-scores in 5 positive controls for *EGFR*, *MET* and *ERBB2* ( $n = 3$ : high, medium, and low level amplification) normalized to a normal tissue control series. Values of all 13 genes relevant for amplification detection (see Supplementary Table 1) are plotted for the five samples. Values were calculated per gene per sample and sorted by increasing value. The positive control values (one gene per samples) are depicted in green, values for all other genes (12 per samples) are shown in black. The additional detected *EGFR* amplification, shown in orange, was confirmed by FISH. **b** Grouped relative coverage in a series of 46 clinical tissue samples and 15 normal tissue controls. Values were calculated per gene per sample. **c** Relative coverage per gene in the series of 46 clinical samples, with additional clinical/molecular information (details in main text). The cut-off for validation (relative coverage  $\geq 3.0$ ) is shown by an orange line. Potential amplifications in green were validated by OncoScan array analysis (the others were not analyzed by OncoScan array). **d** Three positive control samples were diluted in gDNA isolated from normal tissue. Relative coverage (y-axis) and z-scores (above bars) compared with a normal tissue control series are shown. On the x-axis the dilution based on gDNA concentration is shown. *EGFR* positive control: unknown tumor purity. *ERBB2* high positive control: 70% tumor cells. *ERBB2* low positive control: 50% tumor cells

that showed progression upon tyrosine kinase inhibitor treatment), a *KRAS* amplification in one sample (NSCLC case without any identified mutations), and an *ERBB2* (HER2) amplification in one sample (salivary duct carcinoma). Amplifications in five samples were confirmed by OncoScan array analysis (Fig. 2c (green marker),

Supplementary Fig. 4). Copy number gains with a relative coverage between 2.0 and 3.0 are marked as ‘potentially amplified’ to allow follow-up analysis in case the gene amplification is clinically relevant for that particular case.

Cross-validation of the amplification analysis and the threshold in relative coverage was performed on a series

of samples containing amplifications of *FGFR1*, *KRAS*, *MDM2*, *MET*, or *PDGFRA* (Supplementary Table 6). The amplifications that were not detected by the PATH panel analysis were below the limit of detection due to a low tumor purity, or subclonal event as determined by FISH (Supplementary Table 6). A dilution series of three positive controls diluted in gDNA isolated from normal tissue further demonstrated the importance of both tumor purity and level of amplification for the sensitivity of the smMIP-based NGS analysis (Fig. 2d, Supplementary Table 6). High level amplified *EGFR* and *ERBB2* samples were detected up to 16 and 4 times dilution, respectively. However, the *ERBB2* amplification in the median amplified sample was only detectable in the undiluted sample. Taken together, limitations in sensitivity of the smMIP-based NGS analysis in samples with a low tumor purity and/or low level amplifications need to be taken into account. To specify these limitations, the number of required gene copies that can be reliably detected by the smMIP-based NGS analysis given the tumor purity should be documented.

Finally, the effect of the number of analyzed gDNA molecules on the amplification analysis was studied. In two diluted normal tissue specimens, the amount of gDNA did not affect the relative coverage of the 13 genes (Supplementary Fig. 5A). In addition, samples from the validation series were grouped in low (median amplicon coverage < 25 gDNA molecules) or high ( $\geq 25$  gDNA molecules) coverage (i.e. 25 gDNA molecules is set as lower limit of detection for variant analysis [6]). The group with the low unique coverage showed more outliers (Supplementary Fig. 5B), which is most likely due to technical variation due to suboptimal gDNA quality and warrants careful consideration to prevent false positive calls.

In conclusion, the PATH panel can be used to detect gene amplifications, which are considered reliable if they have a relative coverage  $\geq 3$ , an amplified signal clearly above the noise based on visual inspection of all genes and in all duplicate or triplicate analyses, and if the median unique coverage per amplicon is  $> 25$  unique gDNA molecules.

#### Panel validation - MSI detection

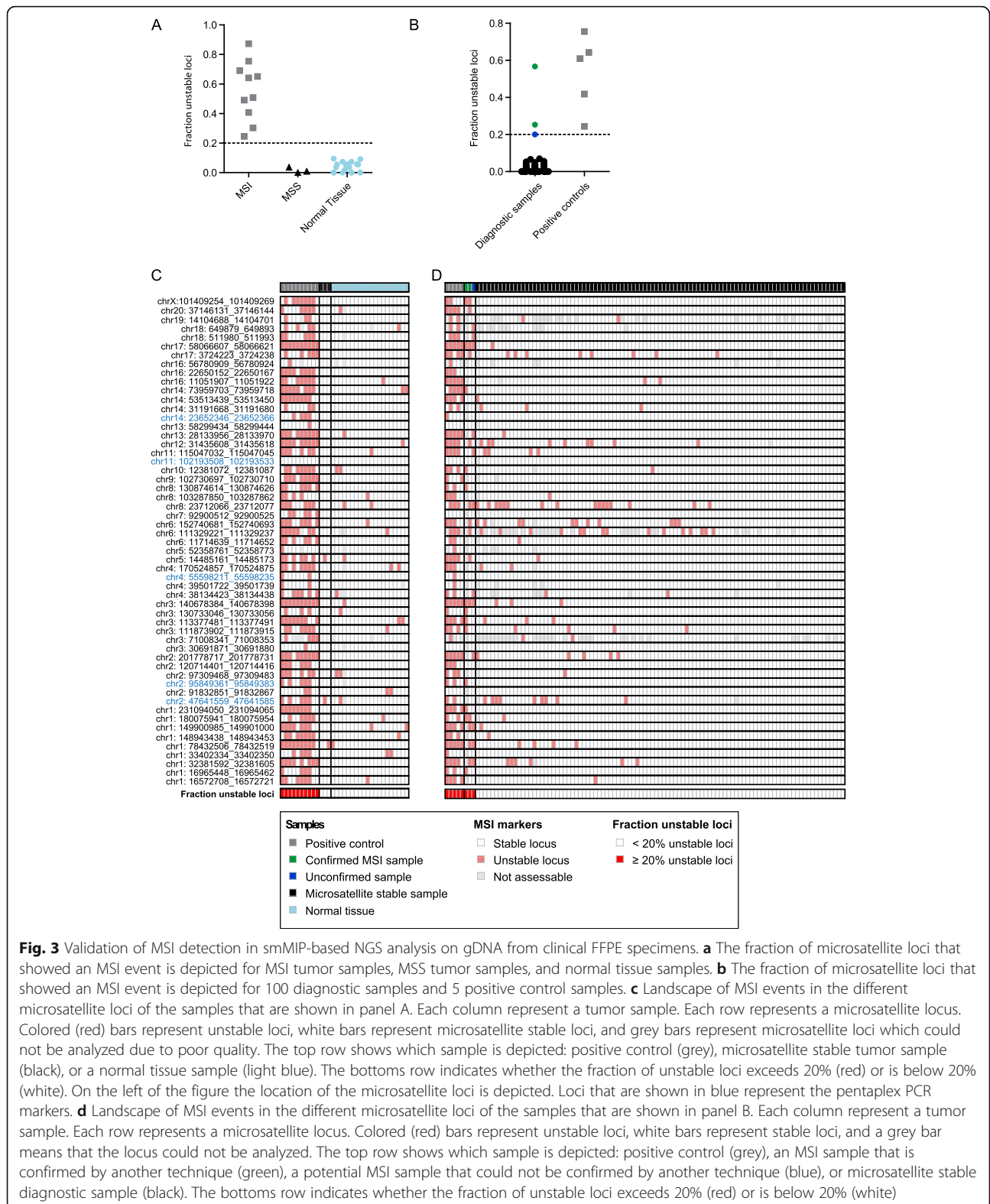
For MSI analyses, a total of 63 microsatellite repeats were selected that were recurrently unstable in a range of MSI unstable tumor types [15, 16]. After initial testing, 55 markers passed the minimal requirements, as also specified by the mSINGS software. To determine the sensitivity of the markers, we analyzed the microsatellite status of 20 normal tissue samples, 3 microsatellite stable (MSS) tumor samples, and 10 MSI-high tumor samples (i.e. colorectal, endometrial, prostate, and salivary duct carcinomas) as assessed by Genescan analysis of a pentaplex PCR [23] and/or immunohistochemistry

(IHC) of the mismatch repair genes. mSINGS software detected a significantly different read length in 22–87% of microsatellite markers for the MSI-high samples. Normal samples and MSS tumor samples showed 0–9% unstable loci (Fig. 3a, c). When applying the previously published cut-off of 20% unstable loci, a 100% sensitivity and specificity was reached [20]. The sample that scored the lowest fraction of unstable loci (22 and 27% in duplicate analysis) was an *MSH6* deficient prostate tumor, which showed a more subtle shift in the pentaplex PCR analysis compared to other positive controls (Supplementary Fig. 6). In three positive control samples (two endometrial tumors and the *MSH6*-deficient prostate carcinoma) none of the microsatellites included in the pentaplex PCR assay was classified as unstable by the smMIP-based NGS analysis, suggesting the sensitivity for these individual loci is lower than the pentaplex PCR due to the objective cut-off (mean number of alleles + (2 x SD)). Nevertheless, the cut-off of 20% unstable loci was also reached in these three samples due to instability of other loci, demonstrating the strength of multiple microsatellite makers.

The stability of the assay was examined by the evaluation of 5 MSI-high positive controls (2 colorectal carcinoma, 2 endometrial carcinoma, and 1 prostate cancer sample) in two independent sequencing runs. Reproducible results were obtained (Supplementary Fig. 7A). To challenge the threshold of 20% unstable loci, 3 positive control samples were diluted to mimic decreasing tumor load. Sample 1 harbored a high number of unstable loci and hence MSI was detected up to a neoplastic cell purity of 12.5% (Supplementary Fig. 7B). However, the fraction of unstable loci was lower in sample 2 and 3 and consequently MSI could not be detected below 35 and 30% neoplastic cell purity, respectively. These results show the importance of both the percentage of tumor cells and the fraction of unstable loci for the sensitivity of the MSI analysis. Based on these results, we decided to use a cut-off of  $\geq 30\%$  neoplastic cell purity to reliably report an MSS results.

Finally, the MSI analysis was evaluated in 100 samples for routine diagnostics. Addition of MSI smMIPs to the PATH panel resulted in near identical coverage of all targeted regions (Supplementary Fig. 7C) and did not affect the above described variant and amplification detection (data not shown). Besides the positive controls, three diagnostic samples (tumor purity  $> 80\%$ ) exceeded the threshold of 20% unstable loci (56, 25, 20%), which all three appeared to be lung cancer samples (Fig. 3b, d). IHC (loss of *MLH1* and *PMS2*) and pentaplex PCR analysis confirmed MSI in the sample with 56% unstable loci (Supplementary Fig. 7D). In the sample with 25% unstable loci (25 and 24% in duplicate analysis) subtle changes were observed in the PCR pentaplex analysis.





**Fig. 3** Validation of MSI detection in smMIP-based NGS analysis on gDNA from clinical FFPE specimens. **a** The fraction of microsatellite loci that showed an MSI event is depicted for MSI tumor samples, MSS tumor samples, and normal tissue samples. **b** The fraction of microsatellite loci that showed an MSI event is depicted for 100 diagnostic samples and 5 positive control samples. **c** Landscape of MSI events in the different microsatellite loci of the samples that are shown in panel A. Each column represent a tumor sample. Each row represents a microsatellite locus. Colored (red) bars represent unstable loci, white bars represent microsatellite stable loci, and grey bars represent microsatellite loci which could not be analyzed due to poor quality. The top row shows which sample is depicted: positive control (grey), microsatellite stable tumor sample (black), or a normal tissue sample (light blue). The bottom row indicates whether the fraction of unstable loci exceeds 20% (red) or is below 20% (white). On the left of the figure the location of the microsatellite loci is depicted. Loci that are shown in blue represent the pentaplex PCR markers. **d** Landscape of MSI events in the different microsatellite loci of the samples that are shown in panel B. Each column represent a tumor sample. Each row represents a microsatellite locus. Colored (red) bars represent unstable loci, white bars represent stable loci, and a grey bar means that the locus could not be analyzed. The top row shows which sample is depicted: positive control (grey), an MSI sample that is confirmed by another technique (green), a potential MSI sample that could not be confirmed by another technique (blue), or microsatellite stable diagnostic sample (black). The bottom row indicates whether the fraction of unstable loci exceeds 20% (red) or is below 20% (white)

No material was available for IHC analysis. For the third sample (19 and 21% unstable loci in duplicate analysis), none of the PCR pentaplex markers showed instability,

which was in agreement with stability of these same loci in the smMIP-based NGS analysis. The IHC analysis did not generate conclusive results for PMS2 and MLH1.

However, the markers for MSI evaluation by IHC and PCR pentaplex analysis were validated on Lynch syndrome associated tumors. Little data is available regarding the reliability of the IHC and PCR pentaplex analysis in lung cancer [14, 17, 20, 34–37]. Therefore, it remains elusive whether this sample should be considered as MSI-high.

Taken together, these observations prompted us to use a threshold of 30% unstable microsatellites to classify a sample as MSI-high. Samples that show < 15% unstable microsatellite loci are classified as MSS. For samples with 15–30% unstable microsatellites IHC of mismatch repair proteins and pentaplex PCR is used to clarify the MSI status.

### Analysis in routine diagnostics

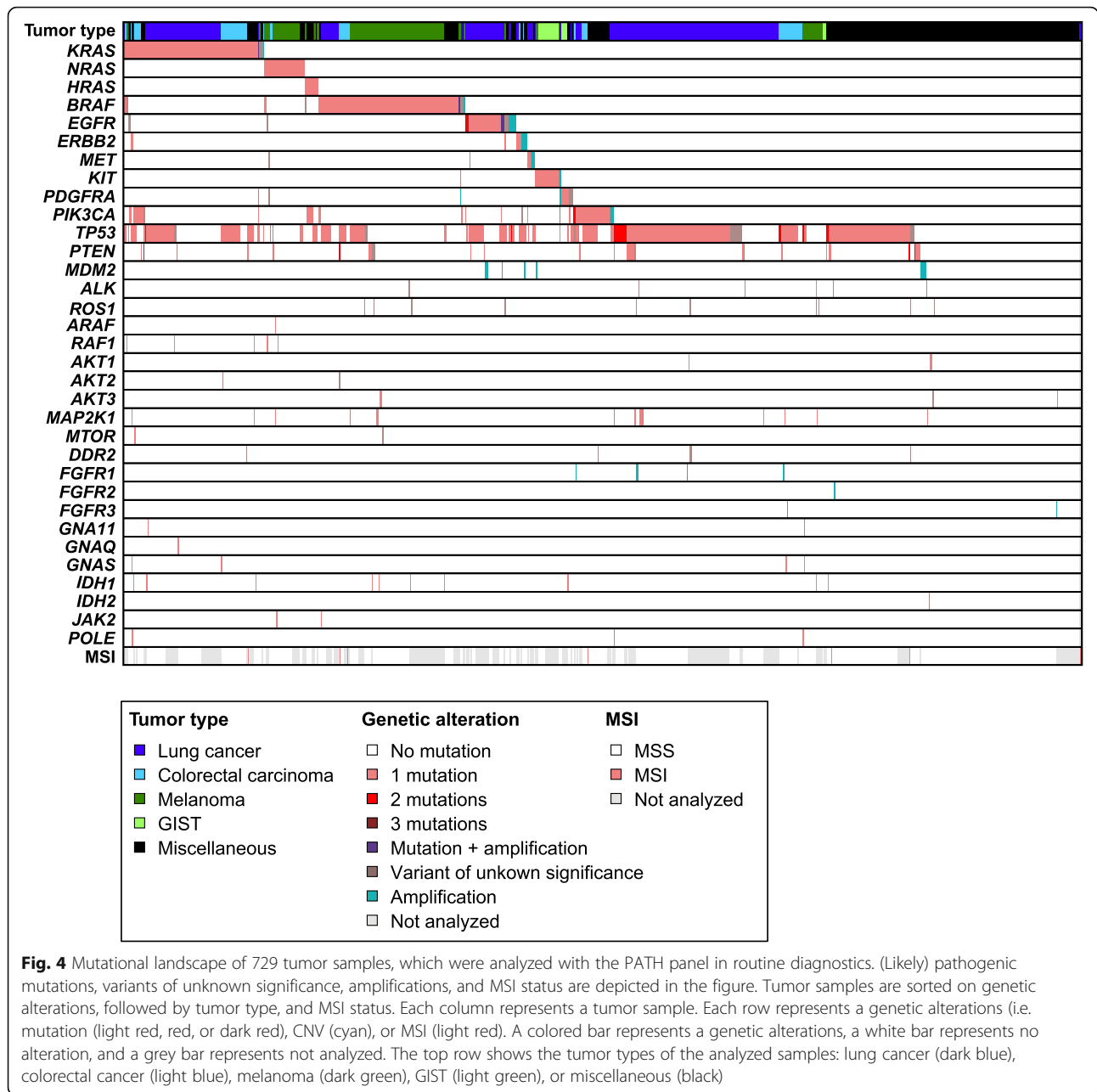
As all formulated requirements of the PATH panel were realized, the PATH panel (Table 1) was implemented as routine diagnostic procedure for predictive analyses for cancer therapy within the Radboudumc. In 2018, 745 routinely available diagnostic samples (35% lung cancer, 17% melanoma, 8% colorectal carcinoma, 3% GIST, and 37% miscellaneous tumors (a range of tumor types that is evaluated for clinical trial options)) were sequenced for variant detection and gene amplification (Supplementary Table 7). These samples were either analyzed in the context of first line treatment selection or detection of therapy resistance. MSI markers were added in a later stage and therefore MSI analysis was performed only in a subset of these samples ( $n = 478$ ). In 18 tumor samples sequencing analysis was not successful due to low quantity or poor quality of the DNA (2.4% sample dropout; Supplementary Table 7–8). In addition, MSI could not be excluded by the smMIP panel in 53 samples due to too low tumor purity (< 30%). 728 routine diagnostic tumor samples were sequenced with sufficient sensitivity (i.e. sufficient reads for reliable variant detection [11]), including samples with a low amount of DNA (< 10 ng; 28/32 samples successful), low tumor purity (5–10%; 10/13 samples successful), and cytological samples (56/62 samples successful). Both DNA quantity and quality influence the quality of the sequence analysis (Supplementary Table 8) and consequently, the minimum required DNA input amount will vary per sample. Therefore, no minimal input requirements were determined and all samples were subjected to smMIP-based NGS analysis.

The PATH panel showed its capability to detect splice site mutations (e.g. *MET*: c.3082 + 1G > A), large deletions (e.g. *MET*:c.2942-4\_3003del), large insertions (e.g. *KIT*: c.1721\_1768dup), and complex deletion-insertion mutations (e.g. *KIT*: c.1924\_1936delinsGAAGTCCTGAGTC; Supplementary Table 7). In total, 788 (likely) pathogenic mutations, and 94 variants of unknown significance were detected. Variant allele frequencies ranged from 2 to 100%. 75% (547/728) of the patients harbored  $\geq 1$  (likely) pathogenic mutations, including genes suitable for

targeted treatment options (e.g. *EGFR*, *BRAF*, *MET*, *ERBB2*, *KIT*, *PDGFRA*). Mutations in *KRAS*, *NRAS*, *HRAS*, *BRAF*, *EGFR*, *ERBB2*, *MET*, *KIT*, and *PDGFRA* were mostly mutually exclusive (Fig. 4, Supplementary Fig. 8). Mutations occurred in tumor types as expected: *KRAS* mutations were mainly observed in colorectal carcinoma (43%) [38] and lung cancer (22%) [1]; melanoma patients showed a high frequency of *BRAF* (56%) and *NRAS* (20%) mutations [39–41]; *EGFR* (12%), *ERBB2* (1%), and *MET* (1%) mutations could be mainly attributed to lung cancer samples [1]; almost all samples suspected of GIST showed a mutation (91%) in *KIT* (68%) or *PDGFRA* (23%) [42]. In the remaining tumor types, 57% of the samples harbored a (likely) pathogenic mutation and/or amplification, including amplifications of *ERBB2* and *EGFR*, and mutations in *BRAF* and *POLE*. In general, *TP53* was the most frequently mutated gene: 1 or 2 *TP53* mutations were observed in 46% (334/728) of the tumor samples [43].

Amplification analysis showed 44 gene amplifications in 5.5% (40/728) of the evaluated samples (Supplementary Table 7). Amplification of *MDM2* was detected most frequently (1.4%; 3 lung cancer, 1 melanoma, 2 liposarcoma, 2 urothelial carcinoma samples), which is reported to be associated with hyperprogressive disease after immunotherapy treatment [44]. Hence, this finding might contribute to the therapy decision. In addition, potential targetable amplifications were detected, i.e. high level amplifications of *ERBB2* [45–47] (colorectal carcinoma, salivary duct carcinoma, prostate cancer, ovarian cancer), *MET* [48, 49] (lung cancer, esophageal carcinoma), and *FGFR1* [50] (lung cancer, colorectal carcinoma).

The PATH panel detected MSI in 7 of the 478 (1.4%) analyzed tumor samples, including three MSI-prone tumors (i.e. 2 of 50 colorectal cancer, 1 of 6 endometrial cancer), and four non-MSI-prone tumors (i.e. 1 lung cancer, 1 prostate cancer, 1 salivary duct carcinoma, and 1 tumor of unknown primary origin; Supplementary Table 7). The number of unstable loci ranged from 41 to 76% and thereby all these tumors were classified as MSI-high tumors. Given the low frequency of MSI-high colorectal tumors, results were matched with available mismatch repair deficiency data. Concordant with the smMIP-based NGS result, none of the 23 tested colorectal tumors showed mismatch repair deficiency by immunohistochemical staining of MLH1, PMS2, MSH2 and MSH6 (Supplementary Table 7). In addition, a supplementary set ( $n = 13$ ) of mismatch repair deficient colorectal carcinomas as determined by IHC, was analyzed, confirming the assay sensitivity. All samples showed > 30% unstable loci and thus would have been classified as an MSI-high tumor by sequence analysis with the PATH panel (Supplementary Fig. 9). Taken together, identification of MSI in non-MSI-prone tumors shows the strength of combination analyses by one NGS panel. As MSI is the



first tumor-agnostic FDA-approved genetic biomarker that gives access to therapy, identification of MSI is of importance for any cancer type [5, 51].

**Discussion**

In the current study, a smMIP-based NGS approach was validated, which aimed at concurrent detection of genomic mutations, amplifications, and MSI in small amounts of histological tumor tissue or cytological material representing routinely available tumor samples. We provide detailed information on the validation of the procedure for predictive testing in current clinical practice and show that it is

feasible to reliably analyze different types of genomic aberrations using a small smMIP-based NGS panel on diagnostic materials. Based on this sensitive and reliable molecular diagnostic analysis, cancer patients can be stratified for targeted therapies and immunotherapy regarding currently available drugs. The combined analysis of different types of genomic aberrations within one analysis saves valuable time, material, and costs. In addition, a broad predictive analysis allows identification of genomic aberrations that are rarely considered by clinicians, e.g. MSI in a non-MSI-prone tumor or a *KIT* mutation in melanoma, which might be beneficial for treatment options [5, 51, 52].

Previously, simultaneous detection of mutations with either amplifications or MSI using targeted NGS was described [10, 13, 19, 53–55]. To our knowledge this is the first small smMIP-based NGS panel that allows concurrent analysis of mutations, amplifications, and MSI by numerous microsatellite loci. Although whole-genome sequencing offers the most comprehensive analysis, the lack of fresh frozen material, high costs and low amounts of tissues/neoplastic cells hinder its use in routine diagnostics. Targeted NGS approaches on the other hand offer lower costs, can yield higher coverage for regions of interest, offer a fast turnaround time, and allow analysis of FFPE specimens with suboptimal gDNA quality. Our panel uses a PCR-based target enrichment method, which is suitable for small gene panels and can handle low (< 10 ng) gDNA input. Upscaling our routine diagnostic panel from the previously published Cancer Hotspot Panel (247 smMIPs [6]) to the PATH panel (663 smMIPs) did not affect sensitivity. In addition, the smMIP panel can be easily adapted by adding smMIPs that target new regions to an already optimized smMIP pool [6], which is an advance in the constantly changing field of personalized medicines and molecular markers, as was demonstrated by the addition of microsatellite markers. Nevertheless, the PATH panel does not prevent additional predictive analyses. Immunohistochemistry of PD-L1, ALK and ROS1, or mRNA based analysis to detect relevant fusion genes involving *ALK*, *ROS1*, *RET*, or *NTRK* genes cannot be replaced by the PATH panel [56]. In addition, tumor mutational burden, associated with a favorable response to immune checkpoint inhibitors [57, 58] cannot be performed by a small NGS panel. Although large (commercial) NGS panels (> 1 Mb) can be used for this analysis instead of whole-genome sequencing or whole-exome sequencing, these panels are generally more expensive, time consuming (due to the complexity of the analysis and interpretation of results), and not appropriate for all diagnostic requests.

The PATH panel was designed to cover most targetable mutations and amplifications for lung cancer, colorectal cancer, melanoma, and GIST for first line treatment options, as well in the setting of therapy resistance (e.g. *EGFR*, *ALK* and *KIT* gatekeeper mutations). In addition, the panel can be used to evaluate possible treatment options for other cancer types in a named or compassionate use program. The 729 diagnostic tumor samples that were evaluated with the smMIP panel showed 788 (likely) pathogenic mutations, including mutations that give access to targeted treatment options (e.g. mutations in *EGFR*, *BRAF*, *MET*, *ERBB2*, *KIT*, *PDGFRA*). Several large deletions and insertions in *MET* and *KIT* were found. Moreover, clinically relevant amplifications were identified, including targetable amplifications (i.e. *MET*, *ERBB2*, *FGFR1*) and amplifications (*MDM2*) associated with

hyperprogressive disease upon immunotherapy treatment [33, 44–46, 48, 50, 59]. Targetable aberrations were not only detected in lung cancer, colorectal cancer, melanoma and GIST, but also in the other cancer types. Lastly, MSI was observed in both MSI-prone and non-MSI-prone tumor samples, which predicts response to immunotherapy [5, 51]. These results show that with the PATH panel a broad spectrum of actionable genetic alterations can be evaluated.

During validation, cut-offs for gene amplifications and MSI analysis were established. In addition, we have identified pitfalls of these analyses. Amplification analysis by NGS relies on sequencing coverage possibly combined with VAFs of germline polymorphisms (SNPs) [11]. We based our analysis on sequencing coverage. Before coverage outliers can be detected, a suitable normalization method has to be applied. The median coverage was shown to be a convenient normalization method, whereas the mean and summed coverage appeared less suitable. To detect coverage outliers, coverage in the sample can be compared to normal samples in the same sequencing run or an external reference pool [10, 60–62]. This latter one saves valuable costs and time (due to lack of availability of a normal sample for every tumor sample) and was shown to generate reliable results. Detection of outliers can be achieved by determining a threshold [54, 63, 64], calculating a *p*-value [7, 61], or bootstrap based estimation of the confidence intervals [65]. We chose for the threshold method and validated a relative coverage  $\geq 3.0$ . This method allows quantification of amplifications by deducing the number of alleles in the tumor cells from the relative coverage and tumor purity of the sample [11]. Accordingly the clinical relevance of amplifications can be established (i.e. low or high number amplification). In addition, the number of gene copies that needs to be present to be quantifiable by the NGS analysis can be deduced and should be considered to be included in the diagnostic report to specify restrictions of NGS-based detection of copy number gains [11]. Although these deductions depend on the estimated tumor purity, which is error-prone, the VAF of somatic variants can be used to support this estimation. If a low sensitivity is obtained or low-level amplifications are of interest, one could consider FISH analysis to detect a specific amplification or OncoScan array to comprehensively analyze copy number variations. Despite that these analyses are more sensitive to exclude the presence of amplifications, especially in samples with limited tumor load and to detect low level amplifications or amplifications in a subset of tumor cells with clusters, the smMIP-based NGS based analysis offers a cost-effective multiplex approach.

Like amplification analysis, MSI analysis was established without the need for a paired normal sample. The five markers that are routinely studied by the pentaplex PCR are mainly validated in Lynch syndrome associated

tumors. This approach is not necessarily applicable to NGS based analysis on different tumor types. With respect to the tumor type, different microsatellites might be affected. Indeed, studies focusing on the spectrum of unstable microsatellites across multiple cancer types showed different patterns of MSI [15–17]. The NGS based methodology also differs from the fragment length analysis based read out. In the smMIP-based NGS panel 55 microsatellite repeats were included that were reported unstable across multiple cancer types in NGS based analysis [15, 16]. By increasing the number of microsatellites that are evaluated, the smMIP-based NGS panel may be more sensitive for a broad range of tumor types compared to the traditional pentaplex PCR. If new predictive microsatellite loci are identified in future studies, smMIPs that target these microsatellite regions can be added to the smMIP pool. Tumor purity also contributes to the sensitivity of the assay. We determined a cut-off of at least 30% tumor cells to reliably exclude MSI and a fraction of unstable loci of over 30% to classify a sample as MSI. The relative low number of MSI-high samples in colorectal carcinoma urged additional validation which confirmed the assay sensitivity. The low frequency of MSI in the colorectal carcinoma samples in the diagnostic cohort, is in line with a low frequency of MSI-high tumors among metastasized colorectal cancer cases [66].

## Conclusion

More genomic information can be extracted from tumor samples than ever before, but less material is available for testing. In the current study, we show that reliable, sensitive, and simultaneous detection of mutations, gene amplifications, and MSI can be achieved in routinely available diagnostic tumor samples using an integrated smMIP-based NGS approach, which was designed in close collaboration with clinicians and molecular biologists of several specialized centers. This single smMIP-based NGS test reduces the number of analyses that are typically performed on tumor samples, uses only limited amount of material, and thereby simplifies the workflow for molecular cancer diagnostics. Moreover, the panel allows easy adaptation and can thereby comply with the rapidly expanding molecular markers.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12885-020-06785-6>.

**Additional file 1:** This files contains supplementary Fig. 1–9 and the supplementary figure legends.

**Additional file 2:** This files contains supplementary Table 1–8.

**Additional file 3:** This files contains the references of supplementary Table 1.

## Abbreviations

CNVs: Copy-number variations; FISH: Fluorescence in situ hybridization; GIST: Gastro-intestinal stromal tumors; IHC: Immunohistochemistry; MSI: Microsatellite instability; MSS: Microsatellite stable; NGS: Next-generation sequencing; PATH: Predictive Analysis for Therapy; smMIP: Single-molecule molecular inversion probe; VAF: Variant allele frequency; VUS: Variant of unknown significance

## Acknowledgments

We thank all technicians of the Molecular Diagnostic Group, all other members of the Laboratory of Tumor genetics, the Core Genome Analysis Laboratory, and all pathologists of our Pathology department for estimation of the tumor purity, isolation of DNA from FFPE tissue, and analyzing and reporting the NGS results.

## Authors' contributions

EMPS, LIK, BBJT, LCvK, AtE, PMN, ES, MJLL and AE were involved in the panel design. AWMKvR, SJBHC, and MJWH performed experiments. EAMJ designed the bio-informatics pipeline for the MSI analysis. EMPS, LIK, MJL and AE analyzed and interpreted the data, and drafted the manuscript. The manuscript was revised and approved by all authors.

## Funding

This work is part of the research program Personalised Medicine which is financed by the Netherlands Organization for Health research and Development (ZonMw, project number 846001001). The funder was not involved in the study design, data collection, analysis, interpretation, or writing of the manuscript.

## Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

## Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Local Ethical Committee of the Radboudumc (CMO 2016–2967). Informed consent was not applicable, as the identity of the human subjects cannot readily be ascertained directly or through identifiers linked to the subjects. The Local Ethical Committee of the Radboudumc approved the waiver of consent. Data from patients who objected against the use of their data for scientific research were excluded from the study.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no conflict of interest. BBJT is a member of the BMC Cancer editorial board, but has not played any role in editorial decisions made regarding this manuscript. Outside the submitted work MJLL has relations with AstraZeneca, Bayer, Bristol-Myers Squibb, Illumina, Janssen pharmaceuticals, Merck, and Nimagen; outside the submitted work LCvK has relations with Janssen pharmaceuticals, Bristol-Myers Squibb, Merck, nanoString, and is an advisor for Guidepoint; outside the submitted work ES has relations with AstraZeneca, Roche, Pfizer, Novartis, Bayer, BMS, BioCartis, Illumina, Agena Bioscience, CC Diagnostics, Janssen Cilag (Johnson&Johnson), Diaceutics, Bio-Rad, Abbott, Boehringer Ingelheim, Promega.

## Author details

<sup>1</sup>Department of Pathology, Radboud university medical center, PO Box 9101, 6500, HB, Nijmegen, the Netherlands. <sup>2</sup>Department of Pathology, Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands. <sup>3</sup>Department of Pathology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. <sup>4</sup>Department of Human Genetics, Radboud university medical center, Nijmegen, The Netherlands. <sup>5</sup>Department of Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands.

Received: 21 October 2019 Accepted: 25 March 2020

Published online: 07 April 2020

## References

- Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature*. 2018;553(7689):446–54.
- Weathers SPS, Gilbert MR. Toward personalized targeted therapeutics: An overview. *Neurotherapeutics*. 2017;14(2):256–64.
- Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin JM, Pfeifer J, Temple-Smolkin RL, Voelkerding KV, Nikiforova MN. Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommendation of the Association for Molecular Pathology and College of American pathologists. *J Mol Diagn*. 2017;19(3):341–65.
- Dietel M, Johrens K, Laffert MV, Hummel M, Blaker H, Pfitzner BM, Lehmann A, Denkert C, Darb-Esfahani S, Lenze D, et al. A 2015 update on predictive molecular pathology and its role in targeted cancer therapy: a review focussing on clinical relevance. *Cancer Gene Ther*. 2015;22(9):417–30.
- Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Lubner BS, Azad NS, Laheru D, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med*. 2015;372(26):2509–20.
- Eijkelenboom A, Kamping EJ, Kastner-van Raaij AW, Hendriks-Cornelissen SJ, Neveling K, Kuiper RP, Hoischen A, Nelen MR, Ligtenberg MJ, Tops BB. Reliable next-generation sequencing of formalin-fixed, paraffin-embedded tissue using single molecule tags. *J Mol Diagn*. 2016;18(6):851–63.
- Budczies J, Pfarr N, Stenzinger A, Treue D, Endris V, Ismael F, Bangemann N, Blohmer JU, Dietel M, Loibl S, et al. Ioncopy: a novel method for calling copy number alterations and mutations using amplicon sequencing data including significance assessment. *Oncotarget*. 2016;7(11):13236–47.
- Shen W, Szankasi P, Sederberg M, Schumacher J, Frizzell KA, Gee EP, Patel JL, South ST, Xu X, Kelley TW. Concurrent detection of targeted copy number variants and mutations using a myeloid malignancy next generation sequencing panel allows comprehensive genetic analysis using a single testing strategy. *Br J Haematol*. 2016;173(1):49–58.
- Singh RR, Patel KP, Routbort MJ, Aldape K, Lu X, Manekia J, Abraham R, Reddy NG, Barkoh BA, Veliyathu J, et al. Clinical massively parallel next-generation sequencing analysis of 409 cancer-related genes for mutations and copy number variations in solid tumours. *Br J Cancer*. 2014;111(10):2014–23.
- Hoogstraat M, Hinrichs JW, Besselink NJ, Radersma-van Loon JH, de Voijis CM, Peeters T, Nijman IJ, de Weger RA, Voest EE, Willems SM, et al. Simultaneous detection of clinically relevant mutations and amplifications for routine cancer pathology. *J Mol Diagn*. 2015;17(1):10–8.
- Eijkelenboom A, Tops BBJ, van den Berg A, van den Brule AJC, Dinjens WNM, Dubbink HJ, Ter Elst A, Geurts-Giele WRR, Groenen P, Groenendijk FH, et al. Recommendations for the clinical interpretation and reporting of copy number gains using gene panel NGS analysis in routine diagnostics. *Virchows Arch*. 2019;474(6):673–80.
- Suraweera N, Duval A, Reperant M, Vaury C, Furlan D, Leroy K, Seruca R, Iacopetta B, Hamelin R. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology*. 2002;123(6):1804–11.
- Waalkes A, Smith N, Penewit K, Hempelmann J, Konnick EQ, Hause RJ, Pritchard CC, Salipante SJ. Accurate pan-Cancer molecular diagnosis of microsatellite instability by single-molecule molecular inversion probe capture and high-throughput sequencing. *Clin Chem*. 2018;64(6):950–8.
- Luchini C, Bibeau F, Ligtenberg MJL, Singh N, Nottegar A, Bosse T, Miller R, Riaz N, Douillard JY, Andre F, et al. ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach. *Ann Oncol*. 2019;30(8):1232–43.
- Bonneville R, Krook MA, Kautto EA, Miya J, Wing MR, Chen H-Z, Reeser JW, Yu L, Roychowdhury S. Landscape of microsatellite instability across 39 cancer types. *JCO Precision Oncology*. 2017;1(1):1–15.
- Cortes-Ciriano I, Lee S, Park WY, Kim TM, Park PJ. A molecular portrait of microsatellite instability across multiple cancers. *Nat Commun*. 2017;8:15180.
- Hause RJ, Pritchard CC, Shendure J, Salipante SJ. Classification and characterization of microsatellite instability across 18 cancer types. *Nat Med*. 2016;22(11):1342–50.
- Kautto EA, Bonneville R, Miya J, Yu L, Krook MA, Reeser JW, Roychowdhury S. Performance evaluation for rapid detection of pan-cancer microsatellite instability with MANTIS. *Oncotarget*. 2017;8(5):7452–63.
- Nowak JA, Yurgelun MB, Bruce JL, Rojas-Rudilla V, Hall DL, Shivdasani P, Garcia EP, Agoston AT, Srivastava A, Ogino S, et al. Detection of mismatch repair deficiency and microsatellite instability in colorectal adenocarcinoma by targeted next-generation sequencing. *J Mol Diagn*. 2017;19(1):84–91.
- Salipante SJ, Scroggins SM, Hampel HL, Turner EH, Pritchard CC. Microsatellite instability detection by next generation sequencing. *Clin Chem*. 2014;60(9):1192–9.
- Zhu L, Huang Y, Fang X, Liu C, Deng W, Zhong C, Xu J, Xu D, Yuan Y. A novel and reliable method to detect microsatellite instability in colorectal Cancer by next-generation sequencing. *J Mol Diagn*. 2018;20(2):225–31.
- Boyle EA, O’Roak BJ, Martin BK, Kumar A, Shendure J. MiPGen: optimized modeling and design of molecular inversion probes for targeted resequencing. *Bioinformatics*. 2014;30(18):2670–2.
- Goel A, Nagasaka T, Hamelin R, Boland CR. An optimized pentaplex PCR for detecting DNA mismatch repair-deficient colorectal cancers. *PLoS One*. 2010;5(2):e9393.
- Neveling K, Mensenkamp AR, Derks R, Kwint M, Ouchene H, Steehouwer M, van Lier B, Bosgoed E, Rikken A, Tychon M, et al. BRCA testing by single-molecule molecular inversion probes. *Clin Chem*. 2017;63(2):503–12.
- Frampton GM, Ali SM, Rosenzweig M, Chmielecki J, Lu X, Bauer TM, Akimov M, Bufill J, Lee C, Jentz D, et al. Activation of MET via diverse exon 14 splicing alterations occurs in multiple tumor types and confers clinical sensitivity to MET inhibitors. *Cancer Discov*. 2015;5(8):850–9.
- Schrock AB, Frampton GM, Suh J, Chalmers ZR, Rosenzweig M, Erlich RL, Halmos B, Goldman J, Forde P, Leuenberger K, et al. Characterization of 298 patients with lung Cancer harboring MET exon 14 skipping alterations. *J Thorac Oncol*. 2016;11(9):1493–502.
- Onozato R, Kosaka T, Kuwano H, Sekido Y, Yatabe Y, Mitsudomi T. Activation of MET by gene amplification or by splice mutations deleting the juxtamembrane domain in primary resected lung cancers. *J Thorac Oncol*. 2009;4(1):5–11.
- Morinaga R, Okamoto I, Fujita Y, Arai T, Sekijima M, Nishio K, Ito H, Fukuoka M, Kadota J, Nakagawa K. Association of epidermal growth factor receptor (EGFR) gene mutations with EGFR amplification in advanced non-small cell lung cancer. *Cancer Sci*. 2008;99(12):2455–60.
- An Z, Aksoy O, Zheng T, Fan QW, Weiss WA. Epidermal growth factor receptor and EGFRvIII in glioblastoma: signaling pathways and targeted therapies. *Oncogene*. 2018;37(12):1561–75.
- Shi H, Moriceau G, Kong X, Lee MK, Lee H, Koya RC, Ng C, Chodon T, Scolyer RA, Dahlman KB, et al. Melanoma whole-exome sequencing identifies (V600E)B-RAF amplification-mediated acquired B-RAF inhibitor resistance. *Nat Commun*. 2012;3:724.
- Morgillo F, Della Corte CM, Fasano M, Ciardiello F. Mechanisms of resistance to EGFR-targeted drugs: lung cancer. *ESMO Open*. 2016;1(3):e000060.
- Wagner PL, Stiedl AC, Wilbertz T, Petersen K, Scheble V, Menon R, Reischl M, Mikut R, Rubin MA, Fend F, et al. Frequency and clinicopathologic correlates of KRAS amplification in non-small cell lung carcinoma. *Lung Cancer*. 2011;74(1):118–23.
- Alotaibi AM, Alqarni MA, Alnobi A, Tarakji B. Human epidermal growth factor receptor 2 (HER2/neu) in salivary gland carcinomas: a review of literature. *J Clin Diagn Res*. 2015;9(2):ZE04–8.
- Kim TM, Laird PW, Park PJ. The landscape of microsatellite instability in colorectal and endometrial cancer genomes. *Cell*. 2013;155(4):858–68.
- Pawlik TM, Raut CP, Rodriguez-Bigas MA. Colorectal carcinogenesis: MSI-H versus MSI-L. *Dis Markers*. 2004;20(4–5):199–206.
- Fong KM, Zimmerman PV, Smith PJ. Microsatellite instability and other molecular abnormalities in non-small cell lung cancer. *Cancer Res*. 1995;55(1):28–30.
- Shridhar V, Siegfried J, Hunt J, del Mar AM, Smith DI. Genetic instability of microsatellite sequences in many non-small cell lung carcinomas. *Cancer Res*. 1994;54(8):2084–7.
- Yaeger R, Chatila WK, Lipsyc MD, Hechtman JF, Cercek A, Sanchez-Vega F, Jayakumar G, Middha S, Zehir A, Donoghue MTA, et al. Clinical sequencing defines the genomic landscape of metastatic colorectal Cancer. *Cancer Cell*. 2018;33(1):125–36 e123.
- Hayward NK, Wilmott JS, Waddell N, Johansson PA, Field MA, Nones K, Patch AM, Kakavand H, Alexandrov LB, Burke H, et al. Whole-genome landscapes of major melanoma subtypes. *Nature*. 2017;545(7653):175–80.
- Lokhandwala PM, Tseng LH, Rodriguez E, Zheng G, Pallavajjala A, Gocke CD, Eshleman JR, Lin MT. Clinical mutational profiling and categorization of BRAF mutations in melanomas using next generation sequencing. *BMC Cancer*. 2019;19(1):665.

41. Reiman A, Kikuchi H, Scocchia D, Smith P, Tsang YW, Snead D, Cree IA. Validation of an NGS mutation detection panel for melanoma. *BMC Cancer*. 2017;17(1):150.
42. Corless CL, Barnett CM, Heinrich MC. Gastrointestinal stromal tumours: origin and molecular oncology. *Nat Rev Cancer*. 2011;11(12):865–78.
43. Baugh EH, Ke H, Levine AJ, Bonneau RA, Chan CS. Why are there hotspot mutations in the TP53 gene in human cancers? *Cell Death Differ*. 2018;25(1):154–60.
44. Kato S, Goodman A, Walavalkar V, Barkauskas DA, Sharabi A, Kurzrock R. Hyperprogressors after immunotherapy: analysis of genomic alterations associated with accelerated growth rate. *Clin Cancer Res*. 2017;23(15):4242–50.
45. Correa TS, Matos GDR, Segura M, Dos Anjos CH. Second-line treatment of HER2-positive salivary gland tumor: ado-Trastuzumab Emtansine (T-DM1) after progression on Trastuzumab. *Case Rep Oncol*. 2018;11(2):252–7.
46. Thorpe LM, Schrock AB, Erlich RL, Miller VA, Knost J, Le-Lindqwister N, Jujavarapu S, Ali SM, Liu JJ. Significant and durable clinical benefit from trastuzumab in 2 patients with HER2-amplified salivary gland cancer and a review of the literature. *Head Neck*. 2017;39(3):E40–4.
47. Sartore-Bianchi A, Trusolino L, Martino C, Bencardino K, Lonardi S, Bergamo F, Zagonel V, Leone F, Depetris I, Martinelli E, et al. Dual-targeted therapy with trastuzumab and lapatinib in treatment-refractory, KRAS codon 12/13 wild-type, HER2-positive metastatic colorectal cancer (HERACLES): a proof-of-concept, multicentre, open-label, phase 2 trial. *Lancet Oncol*. 2016;17(6):738–46.
48. Ou SH, Kwak EL, Siwak-Tapp C, Dy J, Bergethon K, Clark JW, Camidge DR, Solomon BJ, Maki RG, Bang YJ, et al. Activity of crizotinib (PF02341066), a dual mesenchymal-epithelial transition (MET) and anaplastic lymphoma kinase (ALK) inhibitor, in a non-small cell lung cancer patient with de novo MET amplification. *J Thorac Oncol*. 2011;6(5):942–6.
49. Lennerz JK, Kwak EL, Ackerman A, Michael M, Fox SB, Bergethon K, Lauwers GY, Christensen JG, Wilner KD, Haber DA, et al. MET amplification identifies a small and aggressive subgroup of esophagogastric adenocarcinoma with evidence of responsiveness to crizotinib. *J Clin Oncol*. 2011;29(36):4803–10.
50. Nogova L, Sequist LV, Perez Garcia JM, Andre F, Delord JP, Hidalgo M, Schellens JH, Cassier PA, Camidge DR, Schuler M, et al. Evaluation of BGJ398, a fibroblast growth factor receptor 1-3 kinase inhibitor, in patients with advanced solid tumors harboring genetic alterations in fibroblast growth factor receptors: results of a global phase I, dose-escalation and dose-expansion study. *J Clin Oncol*. 2017;35(2):157–65.
51. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science*. 2017;357(6349):409–13.
52. Meng D, Carvajal RD. KIT as an oncogenic driver in melanoma: An update on clinical development. *Am J Clin Dermatol*. 2019;20(3):315–23.
53. Kim JE, Chun SM, Hong YS, Kim KP, Kim SY, Kim J, Sung CO, Cho EJ, Kim TW, Jang SJ. Mutation burden and I index for detection of microsatellite instability in colorectal Cancer by targeted next-generation sequencing. *J Mol Diagn*. 2019;21(2):241–50.
54. Endris V, Penzel R, Warth A, Muckenhuber A, Schirmacher P, Stenzinger A, Weichert W. Molecular diagnostic profiling of lung cancer specimens with a semiconductor-based massive parallel sequencing approach: feasibility, costs, and performance compared with conventional sequencing. *J Mol Diagn*. 2013;15(6):765–75.
55. Trabucco SE, Gowen K, Maund SL, Sanford E, Fabrizio DA, Hall MJ, Yakirevich E, Gregg JP, Stephens PJ, Frampton GM, et al. A novel next-generation sequencing approach to detecting microsatellite instability and pan-tumor characterization of 1000 microsatellite instability-high cases in 67,000 Patient samples. *J Mol Diagn*. 2019;21(6):1053–66.
56. Vaughn CP, Costa JL, Feilotter HE, Petraroli R, Bagai V, Rachiglio AM, Marino FZ, Tops B, Kurth HM, Sakai K, et al. Simultaneous detection of lung fusions using a multiplex RT-PCR next generation sequencing-based approach: a multi-institutional research study. *BMC Cancer*. 2018;18(1):828.
57. Carbone DP, Reck M, Paz-Ares L, Creelan B, Horn L, Steins M, Felip E, van den Heuvel MM, Ciuleanu TE, Badin F, et al. First-line Nivolumab in stage IV or recurrent non-small-cell lung Cancer. *N Engl J Med*. 2017;376(25):2415–26.
58. Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, Minenza E, Linardou H, Burgers S, Salman P, et al. Nivolumab plus Ipilimumab in lung Cancer with a high tumor mutational burden. *N Engl J Med*. 2018;378(22):2093–104.
59. Liu S, Li S, Hai J, Wang X, Chen T, Quinn MM, Gao P, Zhang Y, Ji H, Cross DAE, et al. Targeting HER2 aberrations in non-small cell lung Cancer with Osimertinib. *Clin Cancer Res*. 2018;24(11):2594–604.
60. Pritchard CC, Salipante SJ, Koehler K, Smith C, Scroggins S, Wood B, Wu D, Lee MK, Dintzis S, Adey A, et al. Validation and implementation of targeted capture and sequencing for the detection of actionable mutation, copy number variation, and gene rearrangement in clinical cancer specimens. *J Mol Diagn*. 2014;16(1):56–67.
61. Ross DS, Zehir A, Cheng DT, Benayed R, Nafa K, Hechtman JF, Janjigian YY, Weigelt B, Razavi P, Hyman DM, et al. Next-generation assessment of human epidermal growth factor receptor 2 (ERBB2) amplification status: clinical validation in the context of a hybrid capture-based, comprehensive solid tumor genomic profiling assay. *J Mol Diagn*. 2017;19(2):244–54.
62. Grasso C, Butler T, Rhodes K, Quist M, Neff TL, Moore S, Tomlins SA, Reinig E, Beadling C, Andersen M, et al. Assessing copy number alterations in targeted, amplicon-based next-generation sequencing data. *J Mol Diagn*. 2015;17(1):53–63.
63. Pfarr N, Penzel R, Klauschen F, Heim D, Brandt R, Kazdal D, Jesinghaus M, Herpel E, Schirmacher P, Warth A, et al. Copy number changes of clinically actionable genes in melanoma, non-small cell lung cancer and colorectal cancer—a survey across 822 routine diagnostic cases. *Genes Chromosomes Cancer*. 2016;55(11):821–33.
64. Pfarr N, Penzel R, Endris V, Lier C, Flechtenmacher C, Volckmar AL, Kirchner M, Budczies J, Leichsenring J, Herpel E, et al. Targeted next-generation sequencing enables reliable detection of HER2 (ERBB2) status in breast cancer and provides ancillary information of clinical relevance. *Genes Chromosomes Cancer*. 2017;56(4):255–65.
65. Oliveira C, Wolf T. CNVPanelizer: Reliable CNV detection in targeted sequencing applications. R package version 1.8.0; 2016.
66. Koopman M, Kortman GA, Mekenkamp L, Ligtenberg MJ, Hoogerbrugge N, Antonini NF, Punt CJ, van Krieken JH. Deficient mismatch repair system in patients with sporadic advanced colorectal cancer. *Br J Cancer*. 2009;100(2):266–73.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

